

University of Wollongong - Research Online

Thesis Collection

Title: Mass spectrometry of acylated peptides and proteins

Author: Jenny Anne Vazquez

Year: 2010

Repository DOI:

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Research Online is the open access repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au



RESEARCH ONLINE

University of Wollongong
Research Online

University of Wollongong Thesis Collection

University of Wollongong Thesis Collections

2010

Mass spectrometry of acylated peptides and proteins

Jenny Anne Vazquez
University of Wollongong

Recommended Citation

Vazquez, Jenny Anne, Mass spectrometry of acylated peptides and proteins, Doctor of Philosophy thesis, School of Chemistry, University of Wollongong, 2010. <http://ro.uow.edu.au/theses/3151>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact Manager Repository Services: morgan@uow.edu.au.



RESEARCH ONLINE

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

MASS SPECTROMETRY OF ACYLATED PEPTIDES AND PROTEINS

A thesis submitted in fulfilment of the requirements for the award of the degree

DOCTOR OF PHILOSOPHY

from

UNIVERSITY OF WOLLONGONG

by

JENNY ANNE VAZQUEZ, BA, BSc(Hons), Grad.Dip.Ed.

SCHOOL OF CHEMISTRY

2010

CERTIFICATION

I, Jenny Anne Vazquez, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Chemistry, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Jenny Anne Vazquez

5th May 2010

TABLE OF CONTENTS

CERTIFICATION	i
LIST OF TABLES	iv
LIST OF FIGURES	vi
LIST OF SCHEMES.....	x
ABBREVIATIONS	xi
PUBLICATIONS.....	xiii
ABSTRACT.....	xiv
ACKNOWLEDGMENTS	xvi
1 General Introduction	1
1.1 Proteins and Peptides	1
1.1.1 Modified Proteins and Peptides	2
1.2 Analysis of Proteins and Peptides.....	6
1.2.1 Analysis of Acylated Proteins and Peptides	11
1.3 Mass Spectrometry	12
1.3.1 Ion Sources	13
1.3.2 Mass Analysers.....	17
1.4 Mass Spectrometry of Peptides and Proteins.....	22
1.5 Mass Spectrometry of Modified Proteins and Peptides.....	25
1.6 Aims.....	27
2 Methods.....	29
2.1 Mass Spectrometry	29
2.1.1 Electrospray Ionisation Q-o-TOF MS/MS	29
2.1.2 MALDI Q-o-TOF MS/MS.....	29
2.1.3 MALDI TOF MS and MALDI PSD MS	30
2.1.4 Electrospray Ionisation-Fourier Transform Ion Cyclotron-Resonance MS/MS (ESI-FTICR MS/MS)	31
2.1.5 Materials Required for MS Analysis	32
2.2 Preparation of Acylated Peptides.....	32
2.2.1 Synthesis	33
2.2.2 High Performance Liquid Chromatography (HPLC) Purification	33
2.2.3 Materials for Peptide Acylation.....	34
2.3 Preparation of Monodisperse Polymers.....	34
2.4 Enzymatic Digestions	35
2.4.1 Procedure	35
2.4.2 Materials for Enzymatic Digestions	35
2.5 Liquid Chromatography / Mass Spectrometry (LC/MS).....	35
2.5.1 Sample Preparation and Instrumentation.....	35
2.5.2 Materials for LC/MS.....	36
2.6 Theoretical Mass Calculations.....	36
3 Analysis of Acylated Peptides by Mass Spectrometry.....	38
3.1 Introduction.....	38
3.1.1 Ghrelin: An Appetite Controlling Acylated Peptide.....	39

3.1.2	Mass Spectrometry of Acylated Peptides and Proteins	41
3.2	Results and Discussion	43
3.2.1	Tandem Mass Spectrometry of <i>O</i> -Acylated Neurokinin A	44
3.2.2	Tandem Mass Spectrometry of <i>O</i> -acylated Eledoisin	54
3.2.3	Tandem Mass Spectrometry of Ghrelin	59
3.2.4	Tandem Mass Spectrometry of Ghrelin Tryptic Peptide T1-11	65
3.2.5	Tandem Mass Spectrometry of Tyrosine(4) and Tyrosine(13) <i>O</i> -acylated Renin Substrate Tetradecapeptide	71
3.2.6	Tandem Mass Spectrometry of <i>S</i> -acylated Glutathione	81
3.2.7	Tandem Mass Spectrometry of <i>N</i> -acylated Substance P	88
3.3	Conclusions	91
4	Sequence Determination for a Potential HIV Immune Response	
	Antagonist	96
4.1	Introduction	96
4.1.1	Peptide and Lipopeptide Vaccines	96
4.1.2	Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS)	98
4.1.3	HIV type-1 Nef Protein	102
4.1.4	Aims	104
4.2	Results and Discussion	105
4.2.1	MS Analysis of the Synthetically <i>N</i> -acylated Nef ₁₋₆ Mixture	105
4.2.2	MS/MS of <i>N</i> -glycine Palmitoylated Nef ₁₋₆	108
4.2.3	MS/MS of Nex 1 and Nex 2	115
4.2.4	Tryptic Digestion of Nex 1 and Nex 2	129
4.3	Conclusions	138
5	Characterisation of Monodisperse Polymers by Mass Spectrometry	140
5.1	Introduction	140
5.1.1	Monodisperse Polymers	140
5.1.2	Characterisation of Monodisperse Polymers	141
5.2	Results and Discussion	143
5.2.1	Analysis of Peracylated Ubiquitin by Mass Spectrometry	143
5.2.2	Analysis of Peracylated Bovine Carbonic Anhydrase II by Mass Spectrometry	182
5.2.3	General Discussion	227
5.3	Conclusions	228
6	General Conclusions	230
	APPENDICES	234
	REFERENCES	250

LIST OF TABLES

Table 1.1	Nomenclature for acyl modifications.....	6
Table 2.1	User in-put sequences and user defined amino acids.....	37
Table 3.1	Characteristic and useful sequence ions identified for <i>O</i> -acylated neurokinin A.....	51
Table 3.2	Characteristic and useful sequence ions identified for ghrelin.	64
Table 3.3	Characteristic and useful sequence ions identified for ghrelin tryptic peptide T1-11.....	70
Table 3.4	Characteristic and useful sequence ions identified for <i>O</i> -acylated renin substrate tetradecapeptide.	77
Table 3.5	Characteristic and useful sequence ions identified for <i>S</i> -acylated glutathione.	84
Table 3.6	Marker ions observed in the ESI-Q-o-TOF MS/MS spectra of acylated peptides.	93
Table 4.1	Proposed identities of species present in the acylated Nef ₁₋₆ synthetic mixture.	107
Table 4.2	Assignment of ESI-FTICR MS/MS results for (pam)Nef ₁₋₆	113
Table 4.3	Calculation of the mass difference between Nex1 and Nex 2.	118
Table 4.4	Calculations to determine the mass of amino acid-7.	123
Table 4.5	Summary of the assignment of ESI-FTICR MS/MS results for Nex 1 and Nex 2.	127
Table 4.6	Peptides resulting from the tryptic digestion of Nex 1 and Nex 2.	130
Table 4.7	Assignments for ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 combined tryptic peptides T1-9.	132
Table 4.8	Assignments for ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 combined tryptic peptides T1-6.	136
Table 5.1	Acyl groups used to perfunctionalise ubiquitin and bovine carbonic anhydrase.	141
Table 5.2	Summary of Intact MS observed for functionalised ubiquitin.....	147
Table 5.3	Peptide assignments for the GluC enzymatic digest of perfunctionalised ubiquitin.	152
Table 5.4	Peptide assignments for the AspN enzymatic digest of perfunctionalised ubiquitin.	166
Table 5.5	HPLC-ESI-Q-o-TOF MS/MS analysis of glutarylated and PEGylated ubiquitin containing sub-stoichiometric functionalisation of ubiquitin.	176
Table 5.6	Identification of possible sites containing truncated PEG addition to ubiquitin.	182

Appendix A	Assignment of MALDI-Q-o-TOF MS/MS results for (pam)Nef ₁₋₆	254
Appendix B	Assignment of ESI-Q-o-TOF MS/MS results for (pam)Nef ₁₋₆	255
Appendix C	Assignment of MALDI-Q-o-TOF MS/MS results for Nex 1 and Nex 2....	256
Appendix D	Assignment of ESI-Q-o-TOF MS/MS results for Nex 1 and Nex 2.....	258
Appendix E	Assignment of ESI-FTICR MS/MS results for Nex 1 and Nex 2.....	260
Appendix F	Peptide Assignments for the AspN Enzymatic Digest of Peracylated Bovine Carbonic Anhydrase II.....	263
Appendix G	Peptide Assignments for the GluC Enzymatic Digest of Perpegylated Bovine Carbonic Anhydrase II.....	268

LIST OF FIGURES

Figure 3.1	HPLC trace for the purification of <i>O</i> -octanoylated eledoisin.....	44
Figure 3.2	MALDI PSD mass spectra of neurokinin A (a) and <i>O</i> -octanoyl-neurokinin A (b).....	45
Figure 3.3	MALDI-Q-o-TOF CID MS/MS analysis of neurokinin A (C.E 110 eV) (a) and <i>O</i> -octanoyl-neurokinin A (C.E 110 eV) (b).....	46
Figure 3.4	ESI-Q-o-TOF CID MS/MS analysis of neurokinin A (C.E. 25 eV; 2+ precursor ion) (a) and <i>O</i> -octanoyl-neurokinin A (C.E. 30 eV) (b).....	47
Figure 3.5	Collision energy profiles for the precursor ion and product ions of: <i>O</i> -octanoyl neurokinin A (a) and <i>O</i> -myristoyl neurokinin A (b).....	52
Figure 3.6	MALDI PSD mass spectra of eledoisin (a) and <i>O</i> -octanoyl eledoisin (b). ..	56
Figure 3.7	MALDI-Q-o-TOF CID MS/MS of eledoisin (C.E. 110 eV) (a) and <i>O</i> -octanoyl eledoisin (C.E. 110 eV) (b).....	57
Figure 3.8	ESI-Q-o-TOF CID MS/MS of eledoisin (C.E. 25 eV; 2+ precursor ion) (a) and <i>O</i> -octanoyl eledoisin (C.E. 30 eV) (b).....	58
Figure 3.9	MALDI PSD mass spectrum of ghrelin.....	60
Figure 3.10	ESI-Q-o-TOF CID MS/MS of des-acyl ghrelin.	62
Figure 3.11	ESI-Q-o-TOF CID MS/MS of ghrelin.....	63
Figure 3.12	Collision energy profile for the precursor ion and product ions of ghrelin. .	65
Figure 3.13	MALDI PSD mass spectrum of the ghrelin tryptic peptide T1-11.....	67
Figure 3.14	MALDI-Q-o-TOF CID MS/MS of the ghrelin tryptic peptide T1-11. Bold-face type indicates ions containing the acyl moiety.	68
Figure 3.15	ESI-Q-o-TOF CID MS/MS of the ghrelin tryptic peptide T1-11	69
Figure 3.16	Collision energy profile for the precursor ion and product ions of: ghrelin tryptic peptide T1-11.	71
Figure 3.17	ESI-Q-o-TOF MS/MS of native and <i>O</i> -acyl-renin substrate tetradecapeptide	73
Figure 3.18	Collision energy profiles for the precursor ion and product ions of: tyr(13) <i>O</i> -octanoyl renin substrate (a); tyr(13) <i>O</i> -myristoyl renin substrate (b); tyr(13) <i>O</i> -palmitoyl renin substrate (c); tyr(4) <i>O</i> -octanoyl renin substrate (d) (3+ precursor ions).....	79
Figure 3.19	Collision energy profile for the fragmentation of native and Tyr(13) <i>O</i> -acylated renin substrate tetradecapeptide	80
Figure 3.20	ESI-Q-o-TOF MS/MS of native and <i>O</i> -acyl-glutathione	82
Figure 3.21	Collision energy profiles for the precursor ion and product ions of: <i>S</i> -octanoyl glutathione (a); <i>S</i> -myristoyl glutathione (b); <i>S</i> -palmitoyl glutathione (c).....	86
Figure 3.22	Collision energy profile for the fragmentation of native and <i>S</i> -acylated glutathione.	87

Figure 3.23	MALDI-Q-o-TOF MS/MS of substance P (C.E.110 eV) (a) and <i>N</i> -octanoylated substance P (C.E. 110 eV) (b).	89
Figure 3.24	ESI-Q-o-TOF CID MS/MS of substance P (C.E.25 eV) (a) and <i>N</i> -octanoylated substance P (C.E. 25 eV) (b).	90
Figure 4.1	MALDI-TOF MS of the acylated Nef ₁₋₆ synthetic mixture. Inset: expansion of the <i>m/z</i> region showing Nex 1 and Nex 2 ions.	106
Figure 4.2	MALDI-Q-o-TOF MS/MS of (pam)Nef ₁₋₆ , <i>m/z</i> 900.6 precursor ion.	109
Figure 4.3	ESI-Q-o-TOF MS/MS of (pam)Nef ₁₋₆ , <i>m/z</i> 450.81 precursor ion.	110
Figure 4.4	ESI-FTICR MS/MS of (pam)Nef ₁₋₆ <i>m/z</i> 450.8 precursor ion.	111
Figure 4.5	MALDI-Q-o-TOF MS/MS of Nex 1 and Nex 2.	117
Figure 4.6	Combined ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 using a precursor isolation value of <i>m/z</i> 578.7 with a low resolution setting, allowing for the isolation of both Nex species.	120
Figure 4.7	ESI-FTICR MS/MS of Nex 1 and Nex 2; <i>m/z</i> 579 precursor ion, Q2 CAD at 10 eV, 1000 ms external accumulation.	122
Figure 4.8	ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 tryptic peptides T1-9, <i>m/z</i> 666, low resolution.	131
Figure 4.9	ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 [T1-6+2H] ²⁺ , <i>m/z</i> 463/465. Inset ESI-Q-o-TOF MS/MS of [Nef ₁₋₆ + 2H] ²⁺ , <i>m/z</i> 450.	135
Figure 5.1	MALDI-TOF MS analysis of peracetylated ubiquitin prior to treatment with lithium hydroxide a) and after lithium hydroxide treatment b).	145
Figure 5.2	MALDI-TOF MS analysis of native ubiquitin a), peractylated ubiquitin b), perglutarylated ubiquitin c) and perPEGylated ubiquitin d).	146
Figure 5.3	HPLC-ESI-Q-o-TOF MS/MS analysis of AspN digested perPEGylated ubiquitin: TIC a), combined mass spectrum for the peptide analysis b). ..	150
Figure 5.4	HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 1-16. Peracetylated ubiquitin a).	154
Figure 5.5	HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 25-34. Peracetylated ubiquitin a), perbenzoylated ubiquitin b).	158
Figure 5.6	HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 35-51. Peracetylated ubiquitin a), perbenzoylated ubiquitin b).	161
Figure 5.7	HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 1-20.	168
Figure 5.8	HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 21-31.	168
Figure 5.9	HPLC-ESI-Q-o-TOF MS/MS analysis of perbenzoylated ubiquitin, AspN peptide 32-38.	170
Figure 5.10	HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 39-51.	172

Figure 5.11	HPLC-ESI-Q-o-TOF MS/MS analysis of perglutarated ubiquitin, AspN peptide 58-76.	173
Figure 5.12	HPLC-ESI-Q-o-TOF MS/MS analysis of the sub-stoichiometrically PEGylated ubiquitin GluC peptide 25-34 a). Selected ion chromatogram for the combined 2+ and 3+ states of substoichiometrically PEGylated and perPEGylated ubiquitin AspN peptide 25-34 b).	177
Figure 5.13	HPLC-ESI-Q-o-TOF MS/MS analysis of the unmodified ubiquitin AspN peptide 39-51 a). Selected ion chromatogram for the combined 2+ and 3+ states of the native and PEGylated ubiquitin AspN peptide 39-51 b).	178
Figure 5.14	TIC for the HPLC-ESI-Q-o-TOF MS/MS analysis of AspN digested perPEGylated BCA.	185
Figure 5.15	HPLC-ESI-Q-o-TOF MS/MS analysis of native BCA AspN enzymatic peptide 150-159.	190
Figure 5.16	HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated BCA AspN enzymatic peptide 150-159.	192
Figure 5.17	HPLC-ESI-Q-o-TOF MS/MS analysis of pertrifluoroacetylated BCA AspN enzymatic peptide 150-159.	194
Figure 5.18	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 150-159.	196
Figure 5.19	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA, GluC enzymatic peptide 1-13.	201
Figure 5.20	HPLC-ESI-Q-O-TOF MS/MS analysis of perPEGylated BCA, containing asp-10, AspN enzymatic peptide 1-9 a), containing asp-10, AspN enzymatic peptide 10-17 b).	204
Figure 5.21	HPLC-ESI-Q-o-TOF MS/MS analysis of BCA AspN enzymatic peptide 18-30 a), BCA AspN enzymatic peptide 18-30 containing a succinimide intermediate at position-23 b).	205
Figure 5.22	HPLC-ESI-Q-o-TOF MS/MS analysis of BCA AspN enzymatic peptide 61-69 containing asp-61.	208
Figure 5.23	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 74-79 a) and sub-stoichiometrically PEGylated BCA Asp N enzymatic peptide 74-79 b).	211
Figure 5.24	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 163-172.	213
Figure 5.25	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 31-39.	217
Figure 5.26	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 41-52.	217
Figure 5.27	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 106-116.	219
Figure 5.28	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 117-128.	220

Figure 5.29	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 137-149.....	221
Figure 5.30	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 204-212.....	224
Figure 5.31	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 204-232.....	225
Figure 5.32	HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 237-259.....	226

LIST OF SCHEMES

Scheme 1.1	Structures of an amino acid (a) and of a two consecutive amino acids, showing the amide (or peptide) bond (b).....	2
Scheme 1.2	Schematic of a typical ESI source (not to scale).	14
Scheme 1.3	Schematic of a typical MALDI-TOF mass spectrometer (not to scale).	16
Scheme 1.4	Nomenclature for the fragmentation of peptides by tandem mass spectrometry.	25
Scheme 2.1	Schematic of the QqQ FTICR mass spectrometer used in this thesis.	31
Scheme 2.2	Reaction scheme for the acylation of side-chain residues.	33
Scheme 3.1	Sequence of human ghrelin.	40
Scheme 4.1	Annotated sequence of (pam)Nef ₁₋₆	114
Scheme 4.2	Structure of the palmitoyl moiety.	114
Scheme 4.3	Structure for a) oleoyl and b) stearoyl moieties.....	119
Scheme 4.4	Structure of ester-linked kynurenine residue (denoted X throughout text).	124
Scheme 4.5	Annotated sequence of Nex 1 and Nex 2.....	128
Scheme 4.6	Tryptic digestion of Nex 1 and Nex 2.	130
Scheme 4.7	MS/MS of tryptic peptide 1-9 from Nex 1 and Nex 2.....	133
Scheme 4.8	MS/MS of tryptic peptide T1-6 from Nex 1 and Nex 2..	137
Scheme 5.1	Amino acid sequence for Ubiquitin (human and bovine).....	143
Scheme 5.2	Sequence coverage maps for the GluC digestion of perfunctionalised ubiquitin.	151
Scheme 5.3	Sequence coverage maps for the AspN digestion of perfunctionalised ubiquitin.	165
Scheme 5.4	Amino acid sequence for bovine carbonic anhydrase.	183
Scheme 5.5	Sequence coverage maps for the AspN digestion of perfunctionalised bovine carbonic anhydrase II.	186
Scheme 5.6	Sequence coverage maps for the GluC digestion of perfunctionalised bovine carbonic anhydrase II.	188
Scheme 5.7	Deamidation of Asparagine, via a succinimide intermediate.	199

ABBREVIATIONS

Symbol	Meaning
a	acetyl
ACN	acetonitrile
ACTH	adrenocorticotrophic hormone, clip 18–39
AIDS	acquired immune deficiency syndrome
b	benzoyl
BCA	bovine carbonic anhydrase
BUDA	Boston University Data Analysis (software)
°C	degrees Celsius
C.E.	collision energy
CAD	collision activated dissociation
CE	capillary electrophoresis
CID	collisionally induced dissociation
D	dimensional (as in 2D-gel electrophoresis)
Da	Daltons
DNA	deoxyribose nucleic acid
e	exponent; $\times 10^n$
ECD	electron-capture dissociation
ESI	electrospray ionisation
eV	electron volts
f	trifluoropropionoyl
FTICR	Fourier transform-ion cyclotron resonance
g	glutaroyl
GPI	glycosylphosphatidyl inositol (anchor)
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography
i	iodoacetyl
i.d.	internal diameter
IDA	intelligent data acquisition TM
IPA	isopropanol
IUPAC	International Union of Pure and Applied Chemistry
kDa	kilodaltons
kV	kilovolts
kyn	kynurenine
LC	liquid chromatography
LIT	linear ion trap
<i>M</i>	molar (mol/L)
[m+H] ⁺	protonated or positively charged ion
<i>m/z</i>	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionisation

mg	milligram
MHC	major histocompatibility complex
min	minute
mL	millilitre
mM	millimolar
MS	mass spectrometry or mass spectrum(a)
MS/MS	tandem mass spectrometry (mass spectrometry/mass spectrometry)
MS ⁿ	multi-stage mass analysis
myr	tetradecano- or myrist- acyl group
nm	nanometres
NMR	nuclear magnetic resonance
NNRTI	nonnucleoside reverse transcriptase inhibitor
o	orthogonal (as in Q-o-TOF)
oco	octano- or capryl- acyl group
ole	n-octadeceno- or ole- acyl group
p	methoxypolyethylene glycol propionoyl
pam	hexadecano- or palmit- acyl group
PEG	used to represent methoxypolyethylene glycol propionate
pH	-log[H ⁺]
pmol	picomole
ppm	parts per million
PSD	post source decay
Q or q	quadrupole
QIT	quadrupole ion trap
R	general symbol used to represent a functional group (e.g. for an amino acid sidegroup.)
RF	radio frequency
RNA	ribose nucleic acid
SDS-PAGE	sodium dodecylsulphate polyacrylamide gel-electrophoresis
SIV	simian immunodeficiency virus
SORI	sustained-off resonance ionisation
ste	octadecano- or stear- acyl group
SWIFT	stored waveform inverse Fourier transform
T	denotes peptide resulting from tryptic digestion
T	Tesla
TFA	trifluoroacetic acid
TIC	total ion current
TOF	time-of-flight
u	unified atomic mass unit
μL	microlitre
μm	micrometre
UV	ultra-violet
V	volts
VIS	visible

PUBLICATIONS

1. Yang J., Gitlin I., Krishnamurthy VM., **Vazquez JA.**, Costello C, Whitesides GM. Synthesis of monodisperse polymers from proteins. *Journal of the American Chemical Society.* **2003** 125:12392-3.

My contribution towards this publication was in providing molecular mass measurements of the monodisperse polymers and identifying a spurious acylation. This work is presented in Chapter 5 of this thesis.

2. Gudiksen KL., Urbach AR., Gitlin I., Yang J., **Vazquez JA.**, Costello CE., Whitesides GM. Influence of the Zn(II) cofactor on the refolding of bovine carbonic anhydrase after denaturation with sodium dodecyl sulfate. *Analytical Chemistry.* **2004** 76:7151-61.

My contribution towards this publication was in providing molecular mass measurements of the native and denatured protein, bovine carbonic anhydrase, and identifying sites of deamidation in the native and denatured protein forms. This work is presented, in part, in Chapter 5 of this thesis.

3. Van Rhijn, I.; Young, D. C.; De Jong, A.; **Vazquez, J.**; Cheng, T.-Y.; Talekar, R.; Barral, D. C.; Leon, L.; Brenner, M. B.; Katz, J. T.; Riese, R.; Ruprecht, R. M.; O'Connor, P. B.; Costello, C. E.; Porcelli, S. A.; Briken, V.; Moody, D. B. CD1c bypasses lysosomes to present a lipopeptide antigen with 12 amino acids. *Journal of Experimental Medicine.* **2009**, 206, 1409-1422.

My contribution to this paper was the characterisation of the lipopeptide antigen. The sequencing of this peptide was essential to the progression of the research which led to the publication of this paper. This work is presented in Chapter 4 of this thesis.

4. Moody, B.D., Rhijn, I.V., Young, D.C., Costello, C.E. (2008) Patent: Methods and Compositions for Immunomodulation. USA: 20080226587

My contribution to this patent was the sequencing of the 12-mer lipopeptide (described in publication 3) which is one of two compositions contained this patent. This work is presented in Chapter 4 of this thesis.

5. **Vazquez, JA.**, Berg, EA., Panepinto MJ., Catherine E. Costello, CE. Mass Spectrometry of Acylated Peptides. (In preparation.)

The content of this paper is presented in Chapter 3 of this thesis.

ABSTRACT

The majority of naturally occurring proteins are modified in some manner, with many biological systems requiring these modifications in order to function properly. Acylation is one such type of modification. For example, the neuropeptide ghrelin, which plays a critical role in appetite stimulation, is octanoylated on its ser-3 residue. Many proteins are palmitoylated at one or more cysteine residues, with the lipid moiety essential for membrane binding. It is likely that many biological systems rely on protein acylations, and it would be beneficial to develop techniques that allow a facile detection of naturally occurring acylations.

Synthetically acylated peptides and proteins (a class of monodisperse polymer) have many potential uses. Synthetically acylated peptides and proteins are being developed for use as therapeutic agents and also as chemical standards. Synthetically acylated peptides, for example, have the potential for use as vaccines for diseases such as hepatitis and HIV. The ability to successfully characterise these types of semi-synthetic molecules is imperative in their development process. In this work, mass spectrometry is explored as a means of analysing acylated peptides and proteins.

A number of synthetically acylated peptides were examined using a range of mass spectrometry techniques in order to identify characteristic fragmentations. Acylated peptides were fragmented using electrospray ionisation or matrix-assisted laser desorption/ionisation combined with collisionally induced dissociation tandem mass spectrometry, and also using matrix-assisted laser desorption/ionisation post source decay mass spectrometry. Acylated peptides were observed to fragment in a similar manner to their unacylated counterparts, with the degree of fragmentation observed dependant on the length of the acyl chain (i.e. higher collision energies were required to

illicit the same degree of fragmentation with increasing chain length). The presence of an acylation on an N-terminal amino acid allowed the formation of a b_1 product ion, not normally observed in the spectrum of unmodified peptides. A number of useful marker ions for acylation at serine, tyrosine and cysteine residues were observed, including, acyl carbenium ions and acylated immonium ions. The neutral loss of the acyl moiety was also commonly observed. The tandem mass spectrometry conditions required to produce marker ions were explored.

Two lipopeptides, synthesised as by-products in the production of N-terminally acylated HIV protein Nef₁₋₆ (acyl-GGKWSK), were characterised. These lipopeptides were found to be twelve residues long, with an unusual ether-linked kynurenine at position seven, and contained either a stearyl or oleoyl moiety at their N-terminal glycine residues (ste/ole-GGKWSK-O-kyn-SKWSK). The successful characterisation of these molecules has allowed continued investigations into their use as immunological agents.

A range of novel monodisperse polymers, acylated at their lysine residues, were analysed intact or subsequent to enzymatic digestion, using high performance liquid chromatography and tandem mass spectrometry. Analogous tandem mass spectra were observed irrespective of the acylating agent, however, chromatographic elution times were found to be dependant on the identity of the acyl moiety. Acylated lysine carbenium ions were observed in the tandem mass spectra, 17 u lower in mass than the calculated value for an acylated lysine immonium ion.

This work demonstrates the ability of mass spectrometry to enable the characterisation of a wide range of acylated peptides and proteins. The use of a variety of mass spectrometry and commonly employed analysis techniques (such as enzymatic digestion) further aid in the characterisation of acylated peptides and proteins.

ACKNOWLEDGMENTS

I would first and foremost like to thank my family. My husband, Santiago for his patience and my three children, James, Emily and Annabelle for being themselves.

I thank my initial supervisor Margaret Sheil, who instilled in me a passion for mass spectrometry; the late Larry Hick, who showed me the ropes of an electrospray quadrupole and John Korth for valuable instruction on MALDI-TOF.

I would like to thank all those in the following groups and institutions who have helped me during the duration of my PhD studies:

- University of Wollongong, School of Chemistry and School of Biological Sciences. In particular, the Mass Spectrometry Group.
- Pfizer, Ann Arbor, Michigan. Where I fine-tuned my LC/MS skills.
- Boston University, Mass Spectrometry Resource. Cathy Costello kindly allowed me to join her group, with the project of developing mass spectrometry for the analysis of acylated peptides.
- Collaborators at Brigham and Women's Hospital, Boston. D. Branch Moody and his group who asked me to sequence an interesting pair of acylated peptides.
- Collaborators at Harvard University, Boston. George Whitesides group. In particular, Jerry Yang and Katie Gudiksen, who had prepared a range of acylated proteins.

Thanks go to my friend Sarah Cook for proofreading my thesis and to my Mum for proofreading the reference list.

And final thanks go to Will Price, who supported me through the final hurdles of preparing this thesis.

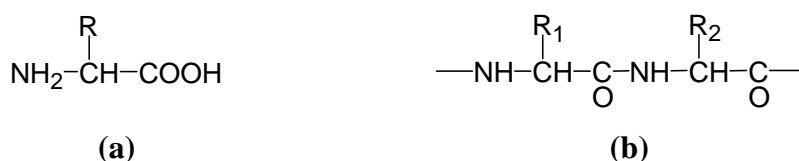
1 General Introduction

The ability to characterise naturally occurring and synthetic chemicals is fundamental to modern biology and chemistry. A complete characterisation of the biomolecules involved in biological processes is imperative in order to understand the mechanisms involved in these processes. Where synthetic molecules are intended for use as therapeutic agents it is important to accurately know the structure and to be able to confirm the purity of a sample. If used as chemical standards for mass measurement or as elution markers, it is also necessary to have a high degree of certainty in the chemical composition. This thesis involves establishing the method of mass spectrometry as a means of characterising proteins and peptides acylated by lipids and small molecules. These types of molecules occur in biological tissues and can also be synthesized for potential use as pharmaceuticals or chemical standards.

1.1 Proteins and Peptides

Proteins are biomolecules consisting of carbon, hydrogen, nitrogen, oxygen and sulphur atoms. Proteins are polymers comprised of L- α -amino acid residues, connected by amide bonds (also referred to as peptide bonds). The sequence of a protein, or polypeptide, is determined by a cell's deoxyribonucleic acid (DNA).^{1, 2} This amino acid sequence is referred to as the primary structure of the protein. Following transcription the protein folds into a particular conformation, determined by a number of noncovalent interactions such as hydrogen bonding, Van Der Waals forces and hydrophobic and hydrophilic interactions. The protein initially forms alpha helices and beta sheets³⁻⁵ (secondary structure) which undergo further interactions to give a tertiary protein structure.⁶ Disulphide bonds between cysteine residues and ionic salt bridges aid in maintaining the three-dimensional structure of the molecule.^{2, 7} Several protein subunits

can join together to give a quaternary structure, referred to as a protein complex.^{8, 9} Conformational changes can occur to a protein's structure as it performs its biological functions.¹⁰ The amino terminus of the protein is referred to as the N-terminus and the carboxyl terminus, as the C-terminus. There are twenty naturally occurring amino acids; the chemical properties of an amino acid are determined by the functionality of its side chain.¹¹ The structures of an amino acid and of two amino acids joined by an amide, or peptide, bond are shown in Scheme 1.1. The term 'peptide' is generally used to refer to polypeptides with fewer than forty amino acids in the chain or to polypeptides which have no tertiary structure.



Scheme 1.1 Structures of an amino acid (a) and of a two consecutive amino acids, showing the amide (or peptide) bond (b). Proline is an exception to the general amino acid structure, as one of the hydrogens on the nitrogen is replaced with a bond to the amino acid side chain.

1.1.1 Modified Proteins and Peptides

The majority of naturally occurring proteins and peptides are modified either co- or post-translationally in some form.^{12, 13} Some proteins become modified as a result of protein aging or disease.^{14, 15} Modifications play a critical role in the structure and/or function of a protein or peptide.¹³ Many modifications are of a reversible nature, and act as a switch for certain biological mechanisms.¹³ Proteins and peptides may also be modified synthetically for uses such as pharmaceuticals or chemical standards.^{16, 17}

Common types of protein modification include small molecule additions, such as phosphorylation and acetylation; and large molecular modifications such as glycosylations and lipid modifications.¹³ It is not uncommon for proteins to have either more than one modification of the same type, or to have multiple modifications of different types. For example, myelin protein zero contains both a glycosylation and an acylation.¹⁸ Protein modifications may also involve a chemical change to individual amino acids, such as oxidation, deamidation, or isomerisation.¹⁴

Protein phosphorylation is a reversible covalent protein modification that occurs on serine, threonine or tyrosine residues.¹⁹ Phosphorylation is found in the activation and inactivation of various enzymes. In some cases protein phosphorylation occurs prior to degradation.²⁰ As phosphorylation is associated with many cell processes, abnormal phosphorylation is found, as a corollary, to be a cause or consequence of many diseases.²¹

A number of proteins have been found to be acetylated, most notably the histones.²² N-acetylation is a permanent modification, with a great variety of N-terminal sequences being modified in this way.²³ Histones, amongst other proteins, are acetylated on the side-chain amino group of lysine residues. Lysine acetylation is reversible, with predominant enzymes being histone acetyltransferase and histone deacetylase,^{23, 24} which also act on a wide range of non-histone proteins.^{25, 26} Other proteins, such as the tubulins, have their own unique acetylation/deacetylation system.²³ Acetylation is associated with protein stability, in particular, by stabilising the N-terminus (in the case of N-acetylation) and with the regulation of protein-DNA interactions, as in the case of the histone proteins.^{13, 22} In the case of DNA-interacting histones the role of acetylation is to reduce the charge of the tail portion of the protein molecule, thus altering the

manner in which the protein interacts with DNA.²³ Many other cellular functions are also associated with acetylation.^{13, 22, 27} Abnormal acetylation, being linked to many cellular functions is also associated with many diseases, such as Huntington's disease,²⁸ inflammatory lung diseases,²⁹ leukaemia, epithelial cancers, fragile X syndrome and Rubenstein-Taybu syndrome.³⁰

Glycosylation involves the enzyme mediated attachment of glycans (linked saccharides) to proteins, lipids and other organic molecules.³¹ Glycosylation can occur co- and post-translationally, with attachment being via a nitrogen of asparagine or arginine (N-linked), an oxygen of serine, threonine, or tyrosine (O-linked) and occasionally a carbon of a tryptophan side chain (C-linked). Glycans are considerably larger than phosphate and acetyl moieties, and as such, have a considerable impact on protein conformation and solubility.³² There is a large family of inherited disorders associated with the mis-assembly of glycans.³³ Disorders range from severe multisystemic cases to disorders restricted to specific organs. Diseases associated with errors in the assembly of glycans for both N-glycosylation and O-glycosylation have been identified.³⁴ Glycation, involves the random attachment of saccharides (fructose and glucose) to proteins. This is an unmediated process, and as such, has a deleterious effect on the function of biomolecules. Diabetes is one chronic disease associated with which glycation.^{35, 36}

In 1951 Folch and Lees were the first to isolate lipid modified proteins (proteolipids) from mammalian tissues.³⁷ The identity of the lipid moieties, however, was not described until the early 1970s.³⁸ Since then, several hundred proteins have been found to be lipidated.³⁹ Lipidation of specific proteins is common to eukaryotic cells.⁴⁰ There are four major forms of lipid modification of proteins: acylation; prenylation (farnesyl; and geranyl-geranyl);⁴¹ attachment of cholesterol^{42, 43} and the attachment of

glycosylphosphatidyl inositol (GPI) anchors.⁴⁴ Prenyl and GPI groups are found at the carboxy-terminus of proteins and are generally considered to be of a permanent nature. GPI anchors are attached via a glycan linkage.⁴⁵ A class of diseases, associated with the mis-folding of proteins, is due in some instances to errors in lipidation.^{46, 47} This class of diseases includes genetic diseases such as cystic fibrosis and phenylketonuria, and are caused by impaired or aberrant protein folding. As a result the protein either undergoes degradation by intracellular proteases before a functional conformation is acquired or an accumulation of protein forms aggregates. A missing prenylation on the protein Rab17A has been implicated in the retinal degeneration disease choroideremia.^{48, 49} Compared to errors in glycosylation, far fewer diseases have been linked to missing or inaccurate lipidation of proteins. Throughout the literature, acyl modified proteins are referred to as both proteolipids and lipoproteins. The term “lipoprotein”, however, is also used in reference to non-covalently linked lipid-protein complexes.⁵⁰

Protein or peptide acylation occurs when an amino acid residue is modified by an acyl functional group via an ester, thioester or amine linkage.⁴⁰ The most common acylations found in nature to date are myristoylation via an N-terminal amine linkage and palmitoylation via a thio-ester linkage at cysteines.⁵⁰ Ester linkages to serine, tryptophan and threonine amino acid residues have also been described.^{40, 51} Table 1.1 shows the structure and nomenclature for some of the common naturally occurring acyl modifications.

Table 1.1 Nomenclature for acyl modifications[†]

Symbol	Structure O=C-[R]-CH ₃ 	Delta mass (u)	Systematic Name ^a	Common Name ^b	Abbreviation
C8:0	-[CH ₂] ₆ -	126	octano-	capryl-	oco
C14:0	-[CH ₂] ₁₂ -	210	tetradecano-	myrist-	myr
C16:0	-[CH ₂] ₁₄ -	238	hexadecano-	palmit-	pam
C18:0	-[CH ₂] ₁₆ -	266	octadecano-	stear-	ste
C18:1(n)	-[CH ₂] ₁₄ -CH=CH- ^c	264	n-octadeceno-	ole-	ole

[†] Based on IUPAC lipid nomenclature⁵². The acyl moiety is also synonymously described as a “fatty acid”.⁵⁰

^a Ending in ‘yl’ for an acyl group or ‘ate’ for an ester group.

^b Ending in ‘oyl’ for an acyl group or ‘ate’ for an ester group.

^c The position of the double bond is shown at C-15 for illustrative purposes only.

Synthetically altered proteins, also referred to as monodisperse polymers¹⁷ are proteins which have acylations and other similar modifications on various amino acid residues, such as lysine and tyrosine.⁵³ Many potential uses have been suggested, ranging from pharmaceuticals to standards with controlled physico-chemical properties.^{54, 55} Acylation has the potential to increase the stability and lifetime of protein-based products and also offers an increased number of delivery options, in particular via the skin or mucosal membranes.⁵⁶⁻⁵⁸ Other applications can be in the purification of recombinant proteins and protein fusion applications.⁵⁹

1.2 Analysis of Proteins and Peptides

Throughout the twentieth century several techniques have been employed to characterised peptides and proteins. These techniques include chromatography, Edman sequencing, nuclear magnetic resonance spectroscopy, crystallography and various forms of electrophoresis. Some of these techniques involve simply the separation and

detection of proteins, whilst others allow structural elucidation. Early forms of protein characterisation involved the separation of samples on chromatographic paper, with the proteins being visualised after reacting with ninhydrin, which turns a deep purple colour.⁶⁰ This technique has since evolved into various forms of liquid chromatography. More specific techniques have been developed to analyse the modifications found on proteins, such as radio-labelling and site-directed mutagenesis.

Various forms of liquid chromatography have long been used in the analysis of proteins and peptides. Liquid chromatography (LC) involves partitioning of analyte ions between a mobile and a stationary phase. High performance liquid chromatography (HPLC) involves separation of analytes on a column that is packed with a stationary phase with the mobile phase being forced through the column under pressure.⁶¹ A common packing for HPLC is octadecylsilyl, C18, with an aqueous based mobile phase. This arrangement of mobile-stationary hydrophobicity is referred to as “reversed phase” liquid chromatography.⁶² HPLC is widely used for the separation of peptides and proteins, both on a preparative and an analytical scale.^{63, 64} The eluent from a HPLC column is often monitored by a UV-VIS or fluorescence detector. HPLC can provide identifying information when the retention times of an analyte are compared to reference compounds (as is the case in Edman sequencing, described below) or when linked to a mass spectrometer, enabling sequencing of peptides as they elute from the column (see section 1.3 below).

There are two major forms of electrophoresis used in the analysis of proteins: gel and capillary. Both forms rely on the differential migration of proteins in an electric field. Gel electrophoresis was first developed in the 1930s, replacing paper and thin-layer electrophoresis, due to its improved separation and higher loading capacity.⁶⁵ The most widely used form of gel electrophoresis for protein analysis is referred to as SDS-

PAGE.^{66, 67} This method involves first treating the proteins with a large excess of sodium dodecylsulphate (SDS). The protein mixture is then separated on a polyacrylamide gel under the influence of an electric field (referred to as polyacrylamide gel-electrophoresis, PAGE). This results in a separation of the proteins essentially by molecular weight. The separated proteins are then visualized using a stain, such as Coomassie Blue,⁶⁸ or by transferring to a blot (for example, a Western Blot) where the blotted proteins can be probed by various antibodies, to test for the presence of specific proteins,⁶⁹ or for use of a sensitive general stain, such as silver stain.⁷⁰ Protein phosphorylations can also be detected on blotted proteins using specific anti-phospho antibodies.⁷¹

A two-dimensional form of gel electrophoresis was developed in the mid-1970s by O'Farrell.⁷² This technique usually involves separation of the proteins based on their individual isoelectric points, through isoelectric focusing on a narrow gel. In the second dimension proteins are generally separated by mass by embedding the first thin gel at the top of an SDS-PAGE gel. Some indication of the presence of protein modifications can be seen on developed gels, where charge ladders can be observed for protein species varying in charge, due to modifications such as phosphorylation or acetylation.⁷²

SDS-PAGE and two-dimensional SDS-PAGE have been very useful analytical tools in the biological laboratory for the purposes of visualising the presence or absence of particular proteins. These techniques also give the user an estimate of the molecular weight and/or the isoelectric point of the proteins analysed. These techniques are limited, however, in their ability to give an accurate molecular weight and to accurately identify posttranslational modifications.

Capillary electrophoresis (CE, sometimes referred to as capillary zone electrophoresis) was developed in the 1960s.⁷³ This technique involves separation of electrically charged analytes in a conductive liquid medium contained in a narrow-bore tube, under the influence of an electric field. Proteins are separated based on their size-to-charge ratio, in a similar manner to SDS-PAGE. Proteins are usually detected upon elution by a UV-VIS detector, although can be coupled to a mass spectrometer.^{74, 75} CE is at least equivalent to HPLC in sensitivity and speed of separation and has superior resolution.⁷⁶ Advantages of CE over HPLC are the ability to separate large and/or hydrophilic protein and the minimal requirement for denaturing solvents.⁷⁷

Edman sequencing is a technique used to determine the amino acid sequence of a protein. This technique was discovered by Pehr Edman in the late 1940s.⁷⁸ The technique involves the reaction of the NH_2 of the N-terminal amino acid with the compound phenylisothiocyanate, which forms a cyclic phenylthiocarbomoyl derivative. Under acidic conditions this modified terminal amino acid is cleaved, without affecting the remaining amino acid chain. The identity of the amino acid is subsequently determined using liquid chromatography or electrophoresis. This technique has been automated to increase the utility of the method,⁷⁹ however the technique has several major limitations. Firstly, as the method relies on reaction at the N-terminal, if the N-terminal amine group is modified post-translationally in some manner, the sequencing cycle is blocked.⁸⁰ Secondly, posttranslational modifications to amino acid residues cause changes to the retention times, leading to erroneous assignments or difficulties in identification.⁸¹ Thirdly, by the nature of the analysis, Edman sequencing cannot readily accommodate protein mixtures, and requires a purified sample. This could be obtained, for example, by prior separation involving electrophoresis or chromatography. A further disadvantage of Edman sequencing is the protracted cycle time, of

approximately 30 minutes per amino acid.⁸² Edman sequencing is a relatively sensitive technique, requiring approximately 10-100 picomoles of purified sample.⁷⁹

The techniques of X-ray crystallography (also known as X-ray diffraction) and protein nuclear magnetic resonance (protein NMR) provide structural information regarding protein conformation.⁸³ X-ray crystallography is most widely known for its use in deciphering the double-helical structure of DNA.⁸⁴ X-ray crystallography involves the analysis of diffraction patterns resulting from the targeting of a sample with X-rays. This technique is very useful in determining the secondary, tertiary and quaternary structure of proteins and was used in the discovery of beta-sheet and helical formations in proteins.^{3, 4} A major drawback of X-ray crystallography is that it requires the protein, or protein complex, of interest to be present in a highly purified crystal form. Generally, a relatively large amount of sample (approximately 10 mg) is required to find optimum crystallisation conditions.⁸⁵

Protein NMR was pioneered by Wüthrich who along with Ernst developed the first two-dimensional NMR experiments and later established the nuclear Overhauser effect as a means for measuring distances within proteins.⁸⁶ Protein NMR results in similar and complementary information to X-ray crystallography on the structures of proteins.⁸³ NMR spectroscopy relies on the nuclear magnetic spin of individual atoms. Many isotopes, such as carbon-12 and oxygen-16 have no net spin. Hydrogen, or proton NMR can be utilised for protein NMR, however a simpler NMR spectrum can be produced if proteins are labelled with carbon-13 or nitrogen-15 isotopes, which each have a net nuclear spin of $\frac{1}{2}$. The sample is prepared in a thin-walled glass tube and placed inside the NMR spectrometer. A radio frequency is applied and response of the analyte nuclei can be measured. The technique gives information on the identity and number of neighbouring atoms, both linked by a covalent bond, and proximal in space. An

10

advantage of NMR is that it enables the solution-phase examination of a protein's conformation.⁸⁷ Two major disadvantages of protein NMR include the requirement for a highly purified sample, and a sample size in the order of nanomoles. There are difficulties in analysing naturally produced proteins for both these reasons, and for the added reason that optional isotopic labelling is only possible with expressed proteins.

1.2.1 Analysis of Acylated Proteins and Peptides

The analysis of acylated proteins and peptides involves the identification of the acyl moiety and of the site of acylation. Current techniques for these purposes include radio-labelling, site-directed mutagenesis and the chemical removal/chromatographic identification of the acyl component(s).

Radio-labelling involves the addition of tritium-labelled fatty acids to culture mediums, which are thereby added to the acylated protein via the cells' usual biochemical processes. After separation, usually by gel-electrophoresis, the tritium labelled protein bands can be visualised.^{88, 89} Site-directed mutagenesis involves substituting the amino acid thought to be acylated with a non-reactive amino acid using DNA technology.³⁹ If the position of acylation is correctly identified, a subsequent change in protein function is observed.

Identification of acyl components has historically required the chemical removal of the acyl moiety from the protein. Amine-linked acyl components can be removed from proteins by treatment with concentrated hydrochloric acid, whilst ester and thio-ester components can be released by base hydrolysis. The released components can then be converted to an ester form and analysed by gas chromatography.⁵⁰

None of these techniques are suited to, nor particularly sensitive for identifying both the type and position of specific acylations. Whilst radio-labelling is a sensitive technique

and allows for the identification of acylated proteins produced in an *in vitro* environment, this technique does not, by itself, allow for the identification the specific acylation sites. In addition, this method can only be used where culture media are being employed. It is not always appropriate in the study of diseases which involve animal tissue. Site-directed mutagenesis is a useful tool in identifying individual sites of acylation, but is not of particular utility in identifying novel acylations in biological systems. There is also the added complication that whilst the substituted amino acid may not be acylated, it may be a critical residue for forming or sustaining a proteins conformation; thereby giving a false positive when the proteins function is inhibited. The removal of acyl components from proteins is useful in identifying these particular components, but gives no specific information as to which amino acid residue sites are acylated.

Whilst the protein analysis techniques described above have been useful in building an understanding of proteins and their biology, they each have their limitations. Mass spectrometry, however, is emerging as a powerful tool in the structural analysis of biomolecules, and in particular, of proteins. Mass spectrometry allows rapid and sensitive sequencing of proteins and has all but eliminated the need for Edman sequencing. In the study of modified proteins, mass spectrometry has the potential to provide a tool which can identify both the site of acylation and the acyl component.

1.3 Mass Spectrometry

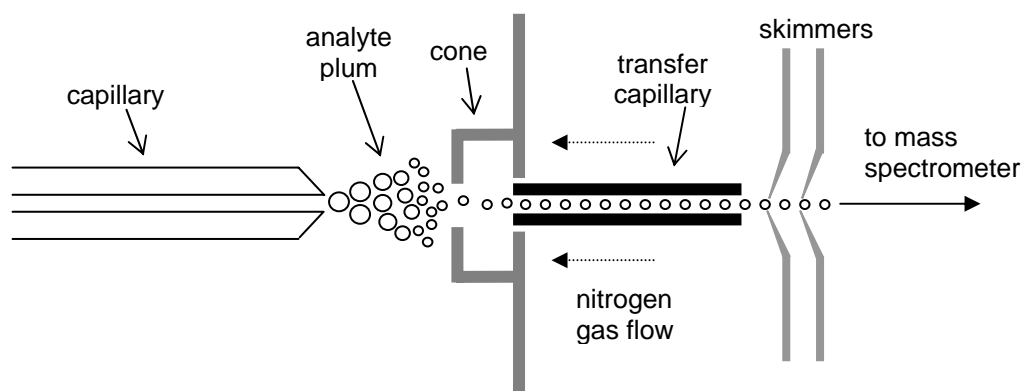
Mass spectrometry (MS) is an analytical tool by which the mass-to-charge ratio (m/z) of an ion is measured. A mass spectrometer consists of two essential elements: the ionisation source and the mass analyser. Structural information of analytes can be gained by fragmenting ions of a specific mass-to-charge ratio and measuring the m/z of

the resultant product ions. Tandem mass spectrometry (MS/MS or MSⁿ) is employed to generate structural information from a given analyte. Except in the case of trapping instruments,⁹⁰ multiple analysers are required in order to perform tandem mass spectrometry. Generally, the first analyser isolates the analyte of interest. A second analyser, or a collision cell, allows for the fragmentation of the analyte via collision with a neutral gas (collision induced/activated dissociation; referred to as CID or CAD).⁹¹ A third analyser separates the resulting fragments prior to detection.⁹² Ion-trap and Fourier transform-ion cyclotron resonance (FTICR) mass analysers are able to both isolate and fragment without the need for a separate collision cell.^{93, 94} The initial internal energy of the ionized molecule can also be harnessed to cause fragmentation either in addition to, or in place of, a tandem MS configuration.^{95, 96} The major elements of modern mass spectrometers are discussed in sections 1.4.1 – 1.4.3 below.

1.3.1 Ion Sources

The most commonly used ionisation sources for protein and peptide analysis are electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI). Electrospray ionisation (ESI) was first suggested as a means of introducing a sample to a mass spectrometer in the 1960s by Dole *et. al.*⁹⁷ and implemented by Fenn and co-workers in the 1980s.^{98, 99} ESI involves the analysis of a molecule dissolved in a polar solvent and pumped through a narrow conducting capillary (e.g. stainless steel, 75-150 µm i.d.) at a flow rate of between 1 µL/min and 1 mL/min. A nebulising gas, flowing around the outside of the capillary, aids in the evaporation of the solvent. A high voltage of 3-4 kV is applied to the tip of the capillary. As a result of the strong electric field generated at the tip, the sample emerging from the tip forms an aerosol of highly charged droplets. As the droplet size decreases due to evaporation of the solvent, the electric field increases to a point where Coulomb repulsion overcomes surface tension

(the Rayleigh Limit).¹⁰⁰ At this stage the droplet becomes unstable and divides into a number of smaller droplets. A schematic of an ESI source is shown in Scheme 1.2 below. There are two main theories as to how individual analyte ions are produced from these small droplets.¹⁰¹ The “Charged Residue Model” involves a continuation of solvent evaporation, subsequent increased electric field, followed by division of the droplet into smaller and smaller droplets until only the analyte ion remains.⁹⁷ The second model is termed the “Ion Evaporation Mechanism” and was proposed by Iribane and Thomson.¹⁰² This model proposes that when the droplets become sufficiently small, the electric field at the surface becomes sufficiently intense to cause a charged analyte ion to be expelled from the droplet’s surface into the surrounding gas. There is a general consensus that large biomolecules form multiply charged ions via the charged residue model, whilst there is still considerable debate as to the most likely mechanism for the formation of smaller analyte ions.⁹⁶ Pseudo tandem fragmentation can be induced in the source of an ESI instrument by varying the voltage within the electrospray source (between the skimmer and cone), thus imparting a greater velocity to the analyte ions. The pressure in this region is sufficient for collisions with neutral gas molecules to occur, resulting in fragmentation, referred to as “in-source CID”.⁹⁶



Scheme 1.2 Schematic of a typical ESI source (not to scale). The capillary can be replaced by a nano-spray tube, or output from an LC column. The capillary can be positioned off-axis and also at an angle to the cone.

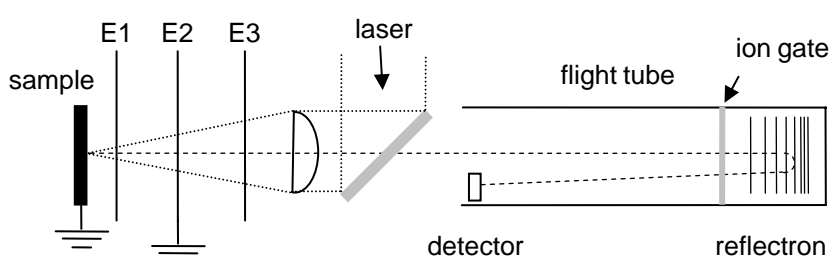
Advantages of electrospray ionisation are: good sensitivity in the high femtomole to low picomole sensitivity (approximately 1×10^{-13} moles); soft ionisation, meaning protein complexes can be maintained in the gas phase; it is useful for combining with liquid chromatography; multiple charging allows for the analysis of high mass ions (up to approximately 70,000 Da), using an analyser with a relatively low m/z range.¹⁰³ Disadvantages include a complicated spectra in mixture analysis, particular where multiple charging is present and intolerance to salts.¹⁰³

Nano-spray (or nano-ESI) is a more recent development of the ESI technique and allows for a significant improvement in sensitivity.¹⁰⁴ The sensitivity for nano-ESI is in the low femtomole range (approximately 1×10^{-15} moles).¹⁰³ The essential component in nano-ESI is that the sample is generally sprayed from a pulled silica glass tube with an orifice of approximately 1 μm . This glass tube is either coated with a conductive material or has a conducting wire inserted into the sample. The flow of sample from the nano-spray tube is mediated by the electric field applied at its tip, or by the solvent flow-rate when coupled with an initial chromatography stage. Apart from a significant improvement in sensitivity, a major advantage of nano-ESI over conventional ESI is its tolerance of salts.¹⁰⁴

Matrix-assisted laser desorption/ionisation (MALDI) was developed as a technique for the analysis of large biomolecules by Tanaka¹⁰⁵ and simultaneously by Karas and Hillenkamp.¹⁰⁶ The MALDI technique involves embedding the analyte molecules in a large excess of a matrix consisting of small organic molecules which have an absorption maximum close to that of the laser wavelength. The analytes are desorbed and ionised directly from the solid state by impact of photons generated by the laser. The analyte ions are accelerated by an electric field and their m/z measured. Delayed extraction has

been an added improvement to the MALDI process whereby the acceleration voltage is applied with a short time delay (0.1-0.5 μs) after the laser pulse.¹⁰⁷⁻¹⁰⁹ This allows for a correction in the spread of the initial velocities of the analyte ions in the initial desorption plume.

Pseudo tandem MS fragmentation information can be obtained in MALDI-TOF-MS by utilising the post-source decay (PSD) of ions created at the source.^{95, 110} PSD spectra are created by separating the ions created by metastable decay of an analyte using an ion gate prior to an ion mirror (reflectron). Product ions which had been travelling at the same velocity as the precursor ion (due to conservation of momentum) are separated at the reflectron due to the differing kinetic energies of each fragment according to their masses.⁹⁵ See Scheme 1.3 below for a depiction of a typical MALDI-TOF mass spectrometer. (The time-of-flight (TOF) mass analyser is described below in section 1.3.2.)



Scheme 1.3 Schematic of a typical MALDI-TOF mass spectrometer (not to scale). Grids E1-E3 allow the user to perform delayed extraction. The reflector at the end of the flight tube is used for both improving resolution (see text below) or when performing post-source decay analysis.

Advantages of the MALDI process are: its relatively high sensitivity (low picomole to femtomole); high mass range, up to 300,000 Da and higher with an appropriate detector;

tolerance towards mixtures and tolerance of millimolar concentrations of salt.¹⁰³ A major disadvantage of MALDI is the interference of matrix ions for masses below approximately 700 Da.¹⁰³ Due to its solid-state nature, MALDI is not as easy as ESI to couple with liquid chromatography separation.

1.3.2 Mass Analysers

Mass analysers have been used to determine the mass of ions for nearly a century. The first mass spectrometer was created by J.J. Thomson in 1913 when he passed a stream of ionised neon through a magnetic and electric field and observed the separated isotopes of neon by placing a photographic plate at the end of the flight tube.¹¹¹ This was the fore-runner of the magnetic sector mass analysers.¹¹² Today, commonly used mass analysis includes the use of time-of-flight (TOF),^{113, 114} quadrupole (Q),⁹² Fourier transform ion cyclotron resonance (FTICR),^{115, 116} and various forms of ion trap mass analysers (QIT, LIT and orbitrap).¹¹⁷⁻¹¹⁹ A separate detector is required to detect separated ions when using a quadrupole or TOF analyser but not necessarily required for use with ion-trap, orbitrap or FTICR-MS mass analysers.

Time-of-flight (TOF) mass analysers separate ions based on their differing mass-to-charge ratios after an initial acceleration under the influence of an electric field.¹¹³ An improvement was made to the basic TOF analysers with the introduction of reflectron-TOF instruments. The reflectron (or ion mirror, formed by a de-accelerating electric field) is able to focus the analyte ions such that ions of the same m/z reach the detector over the smallest possible time window. The basic principle of the reflectron is that ions with greater kinetic energy will travel further into the electric field than ions with less kinetic energy, focussing the ions at the detector, thus improving resolution.^{120, 121} The schematic of a typical MALDI-TOF instrument (Scheme 1.3) shows a flight tube containing a reflectron component. TOF mass analysers have the advantage of a

theoretically limitless mass range; however, resolution decreases significantly with higher masses. A practical m/z range is in the order of 1×10^6 , with resolution up to a m/z of 8,000. A reflectron-TOF has an increased resolution (of 15,000) at the expense of a decrease in m/z range (m/z 10,000).¹⁰³ In hybrid instruments, the TOF analyser can be separated from the ion source, and configured in an orthogonal orientation, thus reducing the influence of kinetic energy distributions created at the ion source.¹²²

A quadrupole (Q) mass analyser consists of four parallel rods which create oscillating electric fields in order to scan a range of m/z values. At any given time, ions of one m/z value are able to pass through the quadrupole to either a detector or a collision cell. The remaining ions become destabilised in the given electric field and are not transmitted through the analyser.⁹² A common configuration for quadrupole instruments is the triple quadrupole: QqQ, where Q₁ is used to filter the ion of interest, collisionally induced dissociation is allowed to occur in q₂ and Q₃ is scanned through the required mass range in order to analyse the fragment ions produced.⁹² An advantage of the quadrupole mass analyser (when used in a QqQ geometry) is the ability to perform precursor ion and neutral loss scans. Precursor ion scans involve setting Q₃ to look for a particular fragment ion (m/z), whilst Q₁ is used to scan through the required range of precursor ion m/z values, with the precursor ions being fragmented in q₂. In a neutral loss scan the second mass analyser scans at a set m/z offset lower than the first analyser, with both analysers being scanned simultaneously. These techniques are useful when analysing classes of compounds which fragment in predictable ways, for example phosphorylated peptides, which have a characteristic neutral loss of 98 u.^{123, 124} Further advantages of the quadrupole mass analyser are its ability to tolerate relatively high pressures, a mass range of up to m/z 4,000 and its relative low cost of manufacture. A disadvantage is a lower resolution compared to other analysers of approximately 4,000.¹⁰³

Fourier-transform ion cyclotron resonance (FTICR) mass analysers measure m/z by detecting an ion's movement in a magnetic field. The frequency of an ion's cyclic movement is determined by its m/z and a spectrum is produced by performing a Fourier transform on the set of image current signals for all ions present in the analyser generated over a period of time.¹¹⁶ The sensitivity of a FTICR mass analyser is greater than a quadrupole or TOF mass analyser since each ion is detected multiple times as it rotates in the magnetic field.¹¹⁶ The resolution of FTICR is also greater than that of a quadrupole due to the greater stability of a superconducting magnet when compared to the rf voltage used in these analysers.^{116, 125} A disadvantage of the FTICR instrumentation is that ultra-high vacuum conditions (approximately 10^{-10} Torr) are required to allow a coherent motion of the ions.¹⁰³ Major drawbacks of the FTICR mass analyser being used as a common instrument are the large amount space required and the relative expense of the magnets used.¹⁰³

An FTICR mass analyser is capable of being used for multi-stage mass analysis (MS^n). All ions, except for the m/z of interest, can be ejected from the FTICR and a number of different fragmentation techniques can be used within the FTICR analyser itself. Sustained-off resonance ionisation-collision activated dissociation (SORI-CAD) is one such technique where the energy required for fragmentation of the precursor ions is achieved by collisions with neutrals accompanied by the periodic excitation of the ion cyclotron radius.¹¹⁶ A second technique, electron-capture dissociation (ECD), involves the direct introduction of low energy electrons to the trapped precursor ions.^{126, 127} ECD is a particularly useful technique in that it allows for the analysis of protein modifications that are otherwise difficult to analyse by CID MS/MS, such as phosphorylation,¹²⁸ N-glycosylation¹²⁹ and O-glycosylation.¹³⁰ In addition different types of fragment ions, complementary to those observed for CID are commonly

observed.¹³¹ FTICR mass analysers can easily be coupled with other analysers, such as a quadrupole, to allow additional types of multiple-stage analysis.¹³²

The quadrupole ion trap, along with the transmission quadrupole (as described above), were first developed by Paul and co-workers in the 1950s.^{133, 134} There are three commonly used ion traps (in addition to the FTICR, which also traps ions): quadrupole ion traps (QIT), linear ion traps (LIT) and the orbitrap.^{117, 118, 135} A quadrupole ion trap analyser (also referred to as a three-dimensional Paul trap) consists of a hyperbolic ring electrode and two hyperbolic end-cap electrodes. The QIT uses an oscillating parabolic electric field applied to the ring electrode to focus ions towards the centre of the trap.¹¹⁷ QIT mass analysers were first used for chemical analysis in the late 1950s; using a power absorbance detection mechanism, without ejection of the ions from the trap.¹¹⁷ In the 1960s developments were made such that ions were ejected from the trap for external detection.¹³⁶ In the 1970s the QIT was used as a means of storing ions prior to mass analysis with a separate quadrupole being used for mass analysis.¹³⁷ In the early 1980s, Stafford and co-workers made two major technical advances. The first was the use of a mass-selective instability mode which allowed the sequential ejection of ions from the trap and simplified the use of the instrument.¹³⁸ The second advancement was made when it was discovered that the use of a damping gas (helium at approximately 1 mtorr) improved the mass resolution of the instrument significantly.¹³⁹ The 1980s and 1990s saw a continued maturation of the technique, with extension of the mass range,¹⁴⁰ the use of MSⁿ,¹⁴¹ and the coupling of the ion trap with electrospray ionisation and MALDI techniques.¹¹⁷ Fragmentation within the ion trap is induced by a resonance signal which can be adjusted to cause CID of the ions with the helium damping gas.¹⁴² In order to isolate ions of a given m/z , a technique called stored waveform inverse Fourier transform (SWIFT) is used.¹⁴³ Advantages of the ion-trap are: its compact size

(approximately the size of a tennis ball); ability to perform multiple stages of mass spectrometry, providing additional fragmentation information (MS^n ,¹⁴¹ and high sensitivity, down to the 20-100 femtomole range. The QIT has an m/z range of 4,000 which is comparable with the quadrupole mass analyser.¹⁰³ Limitations include a limited capacity, and difficulty in producing a complete set of fragmentation ions for peptide analysis.¹¹⁷

Recently, the linear ion trap (LIT, also referred to as a linear quadrupole ion trap) has gained popularity.¹¹⁸ One of the first linear ion traps was made by Church in the late 1960s, who bent linear quadrupoles into a closed ring geometry.¹⁴⁴ Modern linear ion traps are modifications of a straight quadrupole cell.¹¹⁸ A two-dimensional radio frequency (RF) field is used to confine ions radially and stopping potentials applied to end electrodes prevent the axial loss of ions. In comparison to quadrupole ion traps (three-dimensional Paul traps), linear ion traps have higher injection efficiencies and a higher capacity for the storage of ions; whilst maintaining the advantages of high sensitivity, resolution and the ability to perform MS^n .^{118, 145} Linear ion traps are generally used in hybrid instruments, rather than as a stand-alone mass spectrometer. For example, linear ion traps can be used in conjunction with quadrupole ion traps or FTICR analysers, for use with continuous ionisation sources such as ESI, to allow for the accumulation of ions in one trap, whilst the second ion trap analyses the ions.^{118, 146, 147}

In 2000 Makarov introduced a new form of ion trap, called the “orbitrap”.¹¹⁹ This instrument consists of a barrel-like outer electrode and a central axial electrode. A rapidly changing electric field is induced by applying a current to the central axial electrode. Ions are injected into the orbitrap where they are trapped at high kinetic energies around the central electrode.¹³⁵ Once a stable static electric field is established ions perform harmonic oscillations along the axis of the axial electrode. The frequency

of this oscillation is independent of the energy and spatial spread of the ions.¹¹⁹ The mass-to-charge ratio of the ions is determined by performing Fourier transform of the ion image current generated by the axial oscillation of the ions. The orbitrap mass analyser has been combined with a linear ion trap in a hybrid mass spectrometer developed by Thermo Electron (Bremen, Germany).¹³⁵ In this configuration the mass spectrometer has a sub-femtomole sensitivity,¹⁴⁸ mass range of m/z 4,000, and a standard operational resolution of 60,000.¹⁴⁹ An advantage of the orbitrap over the QIT and FTICR is its higher loading capacity.¹⁴⁸ A disadvantage of the orbitrap, when compared to the FTICR, is that the resolution decreases with the mass of the analyte, regardless of the m/z . This becomes particularly apparent in protein analysis, where multiple charging of ESI produced ions allows the formation of a relatively low m/z ion, with a high actual mass.¹⁵⁰ This decrease in resolution is thought to be due to collisions with background ions, hence a high-vacuum is critical for the orbitrap mass analyser.¹³⁵ The orbitrap mass analyser is surpassed in its analytical specifications only by the FTICR mass analyser, however, the significantly lower cost and space requirements make the orbitrap a promising analyser for the future development of mass spectrometry.

1.4 Mass Spectrometry of Peptides and Proteins

ESI and MALDI mass spectrometry are well suited to the analysis of peptides and proteins as they are considered to be “soft” ionisation techniques. This means that the analyte is able to be ionised and desorbed from its solution with minimal fragmentation. It has been suggested that in the ESI process a large amount of the internal energy of an analyte ion is dissipated during the desolvation process.⁹⁶ Peptides and proteins are usually analysed in a positive ionisation mode, whereby a small amount of acid is added to the sample to aid in the protonisation of the analyte molecules. As peptides and

proteins contain multiple protonatable sites (in particular amine groups) it is possible to analyse molecules of a high molecular weight. An advantage of both MALDI and ESI MS analysis of proteins over non-mass spectrometric methods is that only a small quantity of analyte is required for successful analysis: approximately 0.1 - 1 pmol for a 50 kDa protein.^{151, 152}

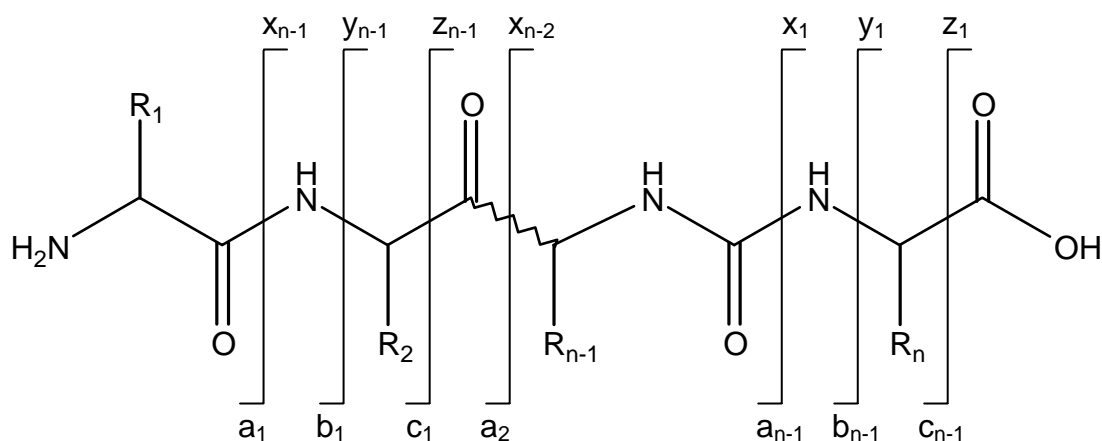
A major application of mass spectrometry in protein chemistry has been in the confirmation of sequence and purity for recombinant proteins, either as intact molecules or as a peptide mixture, resulting from enzymatic digestion.^{151, 153} A second major application is in the field of proteomics, which is the large scale analysis of proteins. Proteomic analysis often involves the analysis of proteins separated by 2D-gel electrophoresis, but is gradually being superseded by combined HPLC and mass spectrometry techniques.¹⁵² ESI mass spectrometry can also be utilised in the examination of ligand binding and protein-protein interactions.¹⁵⁴ Other applications include, but are not limited to: identification of noncovalent protein complexes; localisation of disulfide bonds and the localisation and identification of protein modifications, as discussed in more detail in section 1.5 below.¹⁵⁵

It is common for proteins to be digested into smaller fragments using a suitable enzyme prior to analysis by mass spectrometry. The resultant peptides are often pre-separated prior to MS analysis using a high performance liquid chromatography (HPLC) system.¹⁵³ HPLC involves the injection of the sample onto an appropriate column (usually C18 for small peptides and occasionally C8 for larger peptides), which is then eluted using an increasing gradient of aqueous soluble organic solvent, such as acetonitrile. The elution-point of the HPLC column is then coupled to an ESI or nano-ESI source. HPLC prior to ESI-MS has the added advantage of removing the ionic buffers used in the enzymatic digestion procedure.¹⁵³ Tandem MS of the eluted

peptides allows confirmation of the protein sequence. Although ESI is by nature more suited for the analysis of a chromatographic eluent, technical developments in solution-phase handling allowing the sequential deposition of spots on a MALDI plate for subsequent analysis.¹⁵⁶ Alternatively, nano-electrospray can be used for moderately complex samples without the need for desalting.¹⁵³

Since ESI and MALDI are inherently gentle techniques, if structural information is needed tandem or pseudo tandem mass spectrometry is required (as described above in section 1.3).¹⁵⁵ Tandem MS fragmentation of peptides, has been performed since the 1970s, prior to the invention of the ESI and MALDI ionisation techniques. Beuhler *et. al.* demonstrated the tandem MS of small peptides containing arginine at their N- or C-termini in 1973,¹⁵⁷ the MS/MS of peptides from mixtures also being reported by McLafferty in the late 1970s.^{158, 159} The MS/MS sequencing of proteins subsequent to chemical or enzymatic digestion was reported by Hunt *et. al.* in 1986.¹⁶⁰ Fragmentation of protonated peptides is predominantly along the amino acid backbone, with some side chain cleavage observed when using techniques such as high energy CID or electron capture dissociation.^{132, 161, 162} The nomenclature of peptide fragmentation, first suggested by Rorpfstorf *et. al.*,¹⁶³ is shown in Scheme 1.4 below. Biemann and Johnson have also suggested a similar nomenclature.¹⁶⁴ N-terminal ions, most commonly b_n ions when using low energy CID, are observed if the charge is retained on the N-terminal portion. C-terminal ions, most commonly y_n ions when using low energy CID are observed when the charge is retained on the C-terminal portion. Additional fragments are observed for individual amino acid residues, called immonium ions, which, apart from leucine and isoleucine, are unique for each amino acid.¹⁵⁵ Generally neither set of b -ions or y -ions is complete, and the peptide is sequenced as far as possible by

combining both sets of data.^{152, 155} Tandem mass spectra are often interpreted with computational assistance or by matching against a theoretical peak list.¹⁵²



Scheme 1.4 Nomenclature for the fragmentation of peptides by tandem mass spectrometry.

1.5 Mass Spectrometry of Modified Proteins and Peptides

Mass spectrometry has proven to be an extremely powerful tool in the detection and analysis of protein modifications. Modifications can be detected as either a mass increase or decrease from the predicted mass of the unmodified sequence. Advantages of mass spectrometry include: high sensitivity; the ability to identify the site of modification; the discovery of novel modifications; the ability to analyse complex mixtures; and the ability to quantify sub-stoichiometric, or heterogenous modification at specific sites.¹⁶⁵⁻¹⁶⁷ Some modifications, upon MS/MS, result in ions that indicate their identity, such as acetylated lysine, which results in an ion at m/z 126.¹⁶⁸ Other modifications result in a neutral loss, such as serine phosphorylation, where a neutral loss of 98 Da ($-H_3PO_4$) can be detected.¹⁶⁹

Mass spectrometry has been employed widely in the characterisation of methylated and acetylated lysines present in histone type proteins.^{170, 171} The use of tandem mass spectrometry was found to be sufficiently robust to enable the identification of multiple acetylation sites and also give a direct measurement of endogenous levels of acetylation at individual lysine residues.¹⁷⁰ The ability of mass spectrometry to allow the identification of modified lysine residues has been further utilised in probing protein surface topology.^{172, 173} Selective modification of exposed lysine residues on the surface of proteins and employing mass spectrometry to detect these lysines, gives information to complement X-ray crystallography and multidimensional NMR. It would be beneficial to biochemical analysis to extend the range of acyl type modifications capable of being studied by mass spectrometry, both in terms of the type of acylation and the amino acids at which these modifications are located.

Although mass spectrometry is proving to be a very useful tool in analysis of modified peptides and proteins, there remain analytical obstacles to overcome. Some researchers have found characterisation of peptides modified by acyl moieties a challenge to mass spectrometric analysis due to the hydrophobicity introduced by the acyl components with a subsequent decrease in ionisation efficiency.^{165, 174} When analysing large proteins, proteases are commonly utilised to cleave the protein into smaller peptides more amenable to sequencing by MS/MS. The location of a modification may cause the action of proteases to be inhibited or blocked, generating larger than expected cleavage products,¹⁷⁵ thus reducing the ability to sequence the resulting peptides by tandem mass spectrometry. Hence, this thesis seeks to overcome limitations associated with the mass spectrometric analysis of acyl modified proteins and peptides.

1.6 Aims

The aims of this thesis are to extend the utility of tandem mass spectrometry as a means of characterising acylated peptides and proteins and to do this by:

- performing tandem mass spectrometry on synthetically acylated peptides:
 - determining the conditions most suited for the successful fragmentation and sequencing of acyl modified peptides;
 - identifying distinguishing attributes of the tandem mass spectra that could be used to analyse novel acyl modified peptides and proteins;
- using tandem mass spectrometry to determine the structure of two synthetically produced HIV antagonist lipopeptides; and
- characterising a series of monodisperse polymers (synthetically modified proteins) using mass spectrometry in order to confirm sites of modification.

The first experimental chapter (Chapter 3) will demonstrate that acylated peptides are amenable to analysis by mass spectrometry. This will be shown using a number of peptides, acylated by a variety of acyl chain lengths, on a range of different amino acid side-chain and end terminal groups. A number of useful characteristic ions will be identified. The effect of acyl chain length on the collision energy required for production of these ions will be demonstrated.

The second experimental chapter (Chapter 4) involves the characterisation of two N-terminally acylated peptides. These peptides were by-products of a synthesis aimed at producing human immunodeficiency virus (HIV) antagonist peptides. The two by-product peptides were found to have an antigenic effect in an *in vitro* study performed

by collaborators.¹⁷⁶ In order to develop an understanding of the nature of the antigen effect, it was critical to ascertain the exact composition of these peptides. An acylated peptide (HIV peptide (Pam)Nef₁₋₆) which was an expected product of the synthesis procedure will be analysed by mass spectrometry for comparison with the uncharacterised peptides. A variety of mass spectrometric techniques and enzymatic digestion will be used to determine the structure of the peptides.

The third experimental chapter (Chapter 5) will extend the analysis techniques developed in Chapter 3 to a greater variety of acyl-type modifications. Building on the ability of tandem mass spectrometry to analyse proteins modified at lysines by acetylation and methylation (as described in section 1.5 above) this chapter will demonstrate the power of mass spectrometry to characterise a wide variety of acyl-type lysine modifications. A large peptide (ubiquitin) and mid-sized protein (bovine carbonic anhydrase II, BCA) acylated at their lysine residues and N-termini will be analysed by mass spectrometry either intact or subsequent to enzymatic digestion. Acylation sites and sites at which sub-stoichiometric addition has occurred will be identified. The detection, using mass spectrometry, of the degradation of BCA, via the deamidation of specific asparagine residues, will be demonstrated. It will be further shown in this chapter, that specific fragment ions, unique to lysine acylation, are produced when utilising tandem mass spectrometry.

2 Methods

2.1 Mass Spectrometry

2.1.1 Electrospray Ionisation Q-o-TOF MS/MS

In Chapters 3, 4 and 5 electrospray ionisation (ESI) mass spectra were acquired using a quadrupole orthogonal time-of-flight instrument (Q-o-TOF) (Applied Biosystems/MDS Sciex, API QSTAR Pulsar i LC/MS/MS system; Toronto, Ontario, Canada). A Protana NanoES source (Proxeon Biosystems, Odense, Denmark) was employed using uncoated glass tips (World Precision Instruments, Kwik-Fil borosilicate capillaries, Sarasota, FL), pulled in-house to *ca.* 1 μm i.d. using a micropipette puller (Sutter Instrument Co., model P-97, Novato, CA, USA). Spray was produced with a voltage of 1200-1300 V applied via a platinum wire inserted into the glass tip in contact with the sample solution. Purified peptides were dissolved in 50 % ACN: 1 % formic acid and analysed at a concentration of 0.5 – 2.0 μM . Peptide spectra were calibrated externally using the MS/MS spectrum of glu-fibrinopeptide B.

When plotting relative intensity at varied collision energy (Chapter 3) the relative intensities plotted are based on the total area for the product monoisotopic ion and associated isotopes as determined by the Analyst QSTTM software where the spectrum was obtained from an average of 10 scans. The relative intensity = area (fragment ion)/ Σ area (relevant fragment ions). Spectra were acquired between 10 eV and 75 eV, with the most appropriate collision energy range being used. Data are expressed to 2 decimal places, which reflect the general accuracy of this instrument, given the settings used.

2.1.2 MALDI Q-o-TOF MS/MS

Matrix-assisted laser desorption/ionisation (MALDI) Q-o-TOF mass spectra were obtained on an Applied Biosystems/MDS Sciex, API QSTAR Pulsar i LC/MS/MS

system (Toronto, Ontario, Canada). Samples were reconstituted or diluted in 50 % ACN: 0.5 % TFA and mixed on target with a saturated solution of α -cyano-4-hydroxycinnamic acid (α -cyano) in 50 % acetonitrile, 0.5 % TFA and allowed to air dry. Spectra were calibrated externally using the MS/MS spectrum of glu-fibrinopeptide B. Results from the MALDI Q-o-TOF mass spectrometer were used in Chapters 3 and 4 of this thesis. Data are expressed to 2 decimal places, which reflect the general accuracy of this instrument, given the settings used.

2.1.3 MALDI TOF MS and MALDI PSD MS

Matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectra were obtained on a Bruker Daltonics, Reflex IV MALDI RTOF (Billerica, MA, USA). Peptide samples were reconstituted in 50 % ACN: 0.5 % TFA. Samples were mixed on target with a saturated solution of α -cyano in 50 % acetonitrile, 0.5 % TFA and allowed to air dry. The spectra were calibrated externally using the Bruker Daltonics peptides standards mixture (Billerica, MA, USA) for intact measurements or using the spectra generated for ACTH (adrenocorticotrophic hormone, clip 18–39 monoisotopic m/z 2465.1983; Bruker Daltonics peptides standards mixture; Billerica, MA, USA) when analysing by post source decay (PSD).

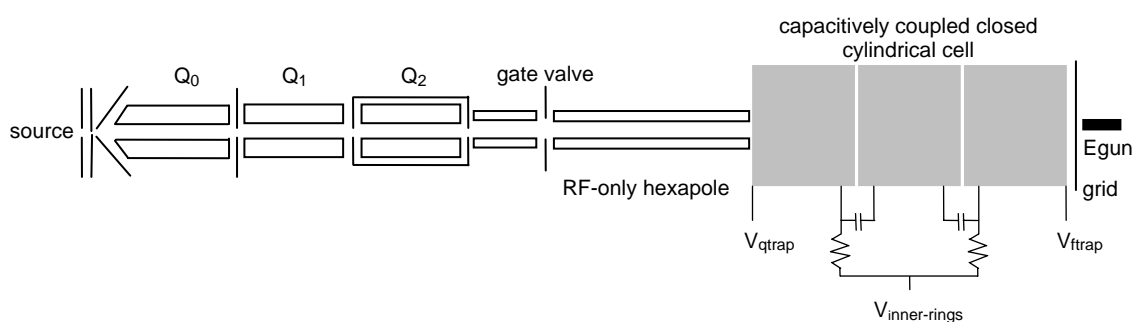
Monodisperse polymers (synthetic derivatives of ubiquitin) were diluted to an appropriate concentration in 50 % ACN: 0.5 % TFA and were mixed on target with matrix composed of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid in a 9:1 ratio in 50 % acetonitrile, 0.5 % TFA and allowed to air dry.

The MALDI-TOF mass spectrometer was used in Chapters 3 and 5 of this thesis; MALDI-PSD was used in Chapter 3. MALDI-TOF MS data are expressed to 2 decimal

places and MALDI-PSD MS data to 1 decimal place, which reflect the general accuracy of this instrument, given the settings used.

2.1.4 Electrospray Ionisation-Fourier Transform Ion Cyclotron-Resonance MS/MS (ESI-FTICR MS/MS)

Additional MS/MS analysis of Nex1 and Nex2 (Chapter 4) was conducted on an electrospray ionisation (ESI) Fourier transform ion cyclotron-resonance mass spectrometer (qQq FTICR MS), equipped with a 7 T actively-shielded magnet (Cryomagnetics, Oak Ridge, TN, USA), built in-house in collaboration with MDS Sciex (Toronto, ON, Canada; see Scheme 2.1).^{132, 177} The front end of the instrument consists of a set of quadrupoles in qQq configuration which have the ability to select, fragment and accumulate ions which are subsequently transmitted into the FTMS for detection. The design and operation of this instrument are discussed in detail elsewhere.¹⁷⁸ The front-end quadrupoles were controlled using the program LC2Tune 1.5 (MDS Sciex) and the program IonSpec99 (Irvine, CA, USA) controlled data acquisition in the ion cyclotron-resonance cell. Spectra were analysed using the Boston University Data Analysis (BUDA) software, developed in-house.¹⁷⁹



Scheme 2.1 Schematic of the QqQ FTICR mass spectrometer used in this thesis.

The synthetic peptide sample analysed in Chapter 4 was diluted 1:40 from its original concentration using 1:1 (vol/vol) methanol to water, 1 % formic acid and was sprayed

using a home-built nanospray source. The solution was loaded into a pulled-glass Kwik-Fil borosilicate capillary tip (1 μm orifice diameter) pulled in-house with a micropipette puller (Model P-97 Flaming/Brown, Sutter Instruments Co. Novato, CA, USA).

The synthetic peptides were isolated and fragmented using external collisionally activated dissociation (CAD). The precursor ions were isolated by Q1 and accelerated at 10 eV into Q2 for collision with N_2 gas. Fragment ions were accumulated in Q3 for 1000 ms and transmitted to the ICR cell for detection. All data were analysed without apodisation and with two zero-fills. Spectra were internally calibrated based on ions the m/z of the precursor ion and other significant product ions. Peaks above 1 % of the base peak were automatically picked using BUDA software. Peaks below 1 % of the base peak were manually identified. Data are expressed to 3-5 decimal places as appropriate, which reflects the general accuracy of this instrument, given the settings used.

2.1.5 Materials Required for MS Analysis

Unless otherwise stated, chemicals, other than solvents were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC grade water, isopropanol and acetonitrile were obtained from Burdick & Jackson (Morristown, NJ, USA).

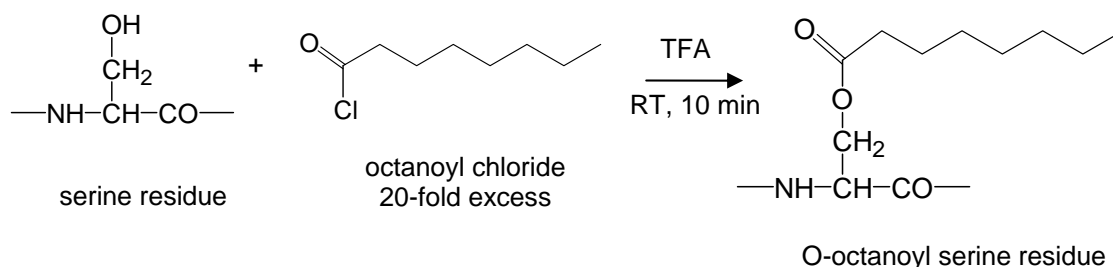
For FTMS analysis Kwik-Fil borosilicate glass capillaries, HPLC-grade methanol and formic acid were obtained from VWR International (Bristol, CT, USA). All water was filtered using a Milli-Q Gradient A10 filtration system from Millipore (Billerica, MA, USA).

2.2 Preparation of Acylated Peptides

Acylated peptides were synthesised for mass spectrometric analysis in Chapter 3 of this thesis.

2.2.1 Synthesis

Synthetically acylated peptides were prepared using a method modified from Yousefi-Salakdeh *et. al.* (see Scheme 2.2).¹⁸⁰ Peptides were dissolved in 100 % TFA (for *O*- and *S*- acylation) or 50 % TFA, 50 % chloroform (for *N*-terminal *N*-acylation). Acyl chlorides (octanoyl, C8; myristoyl, C14 or palmitoyl, C16) were added in a 20 fold excess. After 15 minutes reaction mixtures were dried down under N₂ or quenched with 80 % ethanol and dried under vacuum.



Scheme 2.2 Reaction scheme for the acylation of side-chain residues. (The example given is for the serine amino acid residue. The same reaction conditions were used for cysteine (-SH) and tyrosine (-OH) side-chain residue reactions.)

The naturally occurring acylated peptide, ghrelin, required de-acylation in order to compare the mass spectra of the acylated and de-acylated peptides. 20 μL of 1 M ammonium hydroxide was added to 1 nmol of ghrelin and allowed to react for 1-2 hours at room temperature.

2.2.2 High Performance Liquid Chromatography (HPLC) Purification

Reaction mixtures were purified by reversed-phase HPLC using a C4 column (Vydac C-4 column, 300Å particle size, Deerfield, IL, USA) on a Beckman-Coulter HPLC System Gold (Fullerton, CA, USA). Buffer A: 95 % H₂O, 5 % ACN (acetonitrile), 0.1 % TFA

(tri-fluoroacetic acid); Buffer B: 85 % ACN, 10 % IPA (isopropanol), 5 % H₂O, 0.1 % TFA. The reaction mixture was reconstituted in Buffer A immediately prior to purification. The gradient was 0-100 % B over 20 min. The solvents were modified for the purification of palmitoylated glutathione, where Buffer A: 50 % H₂O, 50 % ACN, 0.1 % TFA and Buffer B was as previously described, the gradient used remained the same. The detector was set at 214 nm. The resulting eluent was dried under vacuum prior to being re-constituted for mass analysis (using a Thermo Electron Corporation Speed Vac, Waltham, MA, USA).

2.2.3 Materials for Peptide Acylation

Peptides and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and solvents were purchased from Burdick & Jackson (Morristown, NJ, USA).

2.3 Preparation of Monodisperse Polymers

The monodisperse polymers studied in this work were synthesised by collaborators and detailed methods are published elsewhere.¹⁷ The synthesis involved the initial denaturation of commercially available proteins, bovine ubiquitin and bovine carbonic anhydrase II (BCA), in the presence of 10 mM SDS, in 0.1 M HEPES buffer (pH 8.2). This was followed by nucleophilic reaction of the exposed lysine side chains with activated carboxylic acid groups (at a ratio of 20 moles per amine group for ubiquitin and 30 moles per amine group for BCA). To remove potential modifications at the hydroxyl groups of tyrosine or serine, the perfunctionalised proteins (100 µM) were incubated in 0.1 M LiOH for 1 hour at 4°C, followed by dialysis against water. The resulting mixtures were analysed by capillary electrophoresis (by our collaborators) and mass spectrometry. Capillary electrophoresis indicated that the completely modified protein, at amine residues, was the sole product, with 93 % yield for ubiquitin and 90 % yield for BCA. Reagents were purchased from Sigma (St. Louis, MO, USA) and Pierce (Rockford, IL, USA).

2.4 Enzymatic Digestions

2.4.1 Procedure

Samples were diluted in 50 mM ammonium bicarbonate, pH 8.5. The relevant enzymes were added to separate aliquots at an enzyme to substrate ratio of 1:50-1:500 and incubated for several hours or overnight at 37°C. Following digestion the samples were either immediately prepared for analysis or stored in the freezer or dried down under vacuum (using a Thermo Electron Corporation Speed Vac, Waltham, MA, USA) to prevent further digestion.

2.4.2 Materials for Enzymatic Digestions

The enzymes LysC, AspN, GluC and trypsin were obtained from Roche (Mannheim, Germany). Ammonium bicarbonate was obtained from Sigma-Aldrich (St. Louis, MO, USA) and water was filtered using a Milli-Q Gradient A10 filtration system from Millipore (Billerica, MA, USA).

2.5 Liquid Chromatography / Mass Spectrometry (LC/MS)

HPLC-ESI-Q-o-TOF MS/MS was utilised in obtaining the results presented in Chapter 5 of this thesis.

2.5.1 Sample Preparation and Instrumentation

A Waters CapLC® (Milford, MA, USA) with an auto-sampler was used to perform the liquid chromatographic separation (HPLC) of peptide mixtures resulting from the enzymatic digestion of proteins (see Section 2.3). Enzymatic digestion mixtures were diluted to a concentration of approximately 1 pmol/μL with buffer A (0.1 % aqueous formic acid) and 1 μL was loaded via a trapping column (Waters Optipak™, Symmetry300™ C18, 5 μM). The sample was washed for 4 minutes at a flow rate of 10 μL/min with 3 % acetonitrile, 0.1 % formic acid using the auxiliary pump. The gradient

used was 3 % buffer B (97 % acetonitrile, 3 % water, 0.1 % formic acid) to 60 % buffer B over 60 min, 60 %-80 % buffer B over 2 min, 80 %-90 % buffer B over 1 min, 90 %-3 % buffer B over 2 min followed by 17 min column re-equilibration between samples. The flow rate was approximately 400 nL/min, split down from an instrument flow rate of 5 μ L/min. The HPLC column used for the separation of peptides was a Waters Symmetry™ C18 3.5 μ m particle size, 75 μ m internal diameter, 100mm column. Eluent was sprayed at 200 V from a distally coated nano-spray tip (New Objective, Woburn, MA, USA) into an ESI-Q-o-TOF as described above using intelligent data acquisition (IDA™). Ions that exceeded 10 counts in MS were subjected to MS/MS. Pre-set collision energies were used, based on the charge state and m/z of the selected ion (approx. 18-60V). The spectra were calibrated externally using the MS/MS spectrum of glu-fibrinopeptide B.

2.5.2 Materials for LC/MS

HPLC grade water and acetonitrile were obtained from Burdick & Jackson (Morristown, NJ, USA). Formic acid was obtained from VWR International (Bristol, CT, USA).

2.6 Theoretical Mass Calculations

The online computer program Protein Prospector v.5.3.2 and previous iterations was used for theoretical mass calculations.¹⁸¹ MS-Product was used to determine the theoretical m/z for the product ions in MS/MS experiments. MS-Digest was used in order to determine the products of enzymatic digestions. User defined sequences and user defined amino acids were used to input the various modifications examined in this thesis. Table 2.1 below lists the sequences used as input in the Protein Prospector programs.

Table 2.1 User in-input sequences and user defined amino acids.

Sequences		User defined amino acids [†]
Chapter 3		
Neurokinin A	HKTDuFVGLM-NH ₂	u = S(Oco): C11 H19 N1 O3 u = S(My): C17 H31 N1 O3
Eledoisin	pyro-EPuKDAFIGLM-NH ₂	u = S(Oco): C11 H19 N1 O3
Ghrelin	GSuFLSPEHQRVQQRKESKKPPAKLQPR	u = S(Oco): C11 H19 N1 O3
Renin Substrate Tetradecapeptide	DRVuIHPFHLLVYS or DRVYIHPFHLLVuS	u = Y(Oco): C17 H23 N1 O3 u = Y(My): C23 H35 N1 O3 u = Y(Pam): C25 H39 N1 O3
Glutathione	EuG	u = C(Oco): C11 H19 N1 O2 S1 u = C(My): C17 H31 N1 O2 S1 u = C(Pam): C19 H35 N1 O2 S1
Substance P	(Oco)RPKPQQFFGLM-NH ₂	(Oco)R: C14 H26 N4 O2
Chapter 4		
Nef ₁₋₆	uGKWSK	u = (Pam)G: C18 H33 N1 O2
Nex1 / Nex 2	uGKWSKvSKWSK	u = (Ole)G: C20 H35 N1 O2 u = (Ste)G: C20 H37 N1 O2 v = ester-linked kynurenine: C10 H9 N1 O3
Chapter 5		
Ubiquitin	(Acet)MQIFVuTLTG DTIENVuAuI QRLIFAGuQL IQuESTLHLV	uTITLEVEPS QDuEGIPPDQ EDGRTLSDYN LRLRGG (Acet)M: C7 H11 N O2 S (Acet)S: C3 H5 N O2
BCA	(Acet)SHHWGYGuHN IANGERQSPV PALuPLALVY NGHSFNVEYD GPLTGTYRLV DQGSEHTVDR HWNTuYGDFG VVGvFLuVGD ALDSIuTuGu SLLPNVLDYW LLESVTWIVL MLuFRTLNFN ANWRPAQPLu	GPEHWHuDFP DIDTuAVVQD GEATSRRMVN DSQDuAVLuD QFHFHWGSSD uuYAAELHLV TAAQQPDGLA ANPALQuVLD STDFPNFDPG TYPGSLTTPP uEPISVSSQQ AEGEPELLML NRQVRGFPu u = K(Acet): C5 H7 N O3 u = K(Iodoacate): C8 H13 N2 O2 I u = K(trifluoropropionate): C9 H13 N2 O2 F3 u = K(PEG): C10 H18 N2 O4 u = K(Benzoate): C13 H16 N2 O2 u = K(Glutamate): C11 H18 N2 O4

[†]acyl listed prior to amino acid indicates an N-terminal modification; acyl listed after amino acid indicates a side-chain addition.

3 Analysis of Acylated Peptides by Mass Spectrometry

3.1 Introduction

Acylation occurs naturally in a limited number of forms, including palmitoylation and myristoylation of proteins and octanoylation and decanoylation of peptides.^{39, 182} Additional, similar lipid modifications include S-prenylation, employing farnesyl and geranylgeranyl groups, cholesterol and glycosylphosphatidyl inositol (GPI) anchors.³⁹ Acylation of peptides and proteins plays numerous important biological functions; however, this class of modification has not been studied to the extent of other classes of modification, such as phosphorylation¹⁸³ and glycosylation.¹⁸⁴

S-acylation is used as a general term applying to the reversible post-translational modification of cysteine residues, by any long-chain fatty acids via thioester linkages. The protein myelin proteolipid protein, for example, which plays an important role in stabilising the multilamellar structure of myelin contains palmitoyl, palmitoleoyl ($C_{16}H_{28}O$), stearoyl ($C_{18}H_{34}O$) and oleoyl ($C_{18}H_{32}O$) as its major acyl groups.^{185, 186} *S*-acylation of cysteine residues can occur enzymatically¹⁸⁷ or non-enzymically via autoacylation, as is the case for at least some G proteins.¹⁸⁸ *S*-acylation is a dynamic process and the *S*-acyl groups often have a much shorter half-life than that of the protein.¹⁸⁹ There is no well-defined consensus sequence which determines the location of an *S*-acyl modification,¹⁹⁰ although, *in-vitro* studies of short peptides showed that *S*-acylation is favoured in the presence of basic and hydrophobic residues and is inhibited by the presence of acidic residues.¹⁹¹ Consequently, *S*-acylation must be identified at the level of individual proteins, and mass spectrometry plays an important role in facilitating this.¹⁹²

Palmitoylation refers to permanent and reversible modifications by the acyl group containing $C_{16}H_{31}O$ (see Table 1.1). In addition to *S*-acylation, palmitoylation often occurs on lysine residues, such as those found in the toxin protein of *Bordetella pertussis*. The *Bordetella pertussis* RTX adenylate cyclase toxin-hemolysin, for example, acquires biological activity upon a single amide-linked palmitoylation of the ϵ -amino group of lysine-983 by the accessory fatty-acyltransferase CyaC.¹⁹³

Palmitoylation, and more generally, *S*-acylation, are associated with numerous biological functions, including: neurotransmitter release;^{194, 195} signal transduction;¹⁹⁶ compaction of the myelin sheath;¹⁸⁶ enzyme regulation;¹⁹⁷ photoreceptor activity;¹⁹⁸ membrane binding;¹⁹⁹ reversible membrane association;^{200, 201} vacuole fusion;²⁰² and bacterial surface adhesion.²⁰³ In the scientific literature, the term palmitoylation is often used ambiguously in place of the term *S*-acylation.

Myristoylation ($C_{14}H_{27}O$; see Table 1.1) is generally considered a permanent modification of which there are two different naturally occurring types: the more common N-terminal myristoylation of glycine and the rare myristoylation of lysine.^{204, 205} N-terminal myristoylation has been found to occur co-translationally.²⁰⁶ Protein myristoylation has a similar range of biological functions to that of palmitoylation²⁰⁷ and often plays a role in conjunction with cysteine palmitoylation, such as in subcellular interactions and signalling.²⁰⁸ Myristoylation has been associated with cancer progression, with the enzyme myristoyl-CoA:protein N-myristoyltransferase found at elevated levels in carcinomas when compared to normal tissue.^{209, 210}

3.1.1 Ghrelin: An Appetite Controlling Acylated Peptide

In the late 1990s a novel acylated peptide, called ghrelin, was discovered.²¹¹ Ghrelin contains an unusual acyl modification as the acyl moiety is found on a serine residue in

the rat and human forms and on a threonine residue in bull-frog ghrelin.²¹² The acyl modification is reversible and plays an important role in appetite control and the regulation of insulin and growth hormone release.^{213, 214} Human ghrelin, isolated from the stomach, has been found in four different forms: non-acylated, octanoylated (C8:0), decanoylated (C10:0) and possibly decenoylated (C10:1).¹⁸² The sequence of octanoylated human ghrelin is shown in Scheme 3.1.



Scheme 3.1 Sequence of human ghrelin.

Ghrelin is secreted in the stomach and found in the gut and brain.^{215, 216} Ghrelin acts along with leptin, insulin, protein YY and cholecystokinin to control hunger. Ghrelin, which is an appetite stimulant, and protein YY and cholecystokinin, which are appetite suppressants, have been linked to short term feeding habits.^{217, 218} In contrast, leptin and to a lesser extent insulin, which are both appetite suppressors, have been linked with weight maintenance over the longer term.^{219, 220} The mis-regulation, intolerance or non-recognition of any of these stimulants or suppressors has the potential of causing obesity, an increasingly problematic disease.²²¹ Ghrelin levels in the blood rise before meals and fall after eating.²²² In obesity low levels of ghrelin are found, however, after dieting, ghrelin levels increase.²²³ Contrastingly, following gastric bypass surgery, ghrelin levels go down, independent of weight loss.²²² Humans suffering from the rare and inherited Prader-Willi syndrome overproduce ghrelin and become obese, often dying before the age of 30 due to obesity related diseases.^{224, 225} Blocking ghrelin receptors may act as a weight loss aid; consequently, a number of synthetic pharmaceuticals have recently been proposed as potential ghrelin antagonists.²²⁶⁻²²⁸

The discovery of ghrelin suggests the likely presence of other yet undiscovered acylated peptides in the human body. Hence there is a need to develop techniques suitable for identifying acylated peptides in biological tissues. The structural divergence of the acyl moiety on ghrelin suggests the need for a robust method of identifying acylated peptides and proteins that can be adapted to analyse multiple acyl chain lengths. Additionally, a method needs to be able to adapt to the different amino acids at which acyl modification occurs. Ghrelin was used as a model compound in this work, as an example of a naturally occurring acylated peptide, and analysed along side synthetically acylated peptides.

3.1.2 Mass Spectrometry of Acylated Peptides and Proteins

Mass spectrometry has been used to successfully identify a number of acylations found on peptides and proteins. For example, mass spectrometry was used to identify the heterogeneous nature of *S*-acylation in GAP-43¹⁹⁹ and N-terminal and *S*-acylation in the Src family kinases.²²⁹ Mass spectrometry has also been used to observe a reversible acylation-deacylation mechanism for the inactivation of human sputum elastase with the inhibitor, ONO-5046.²³⁰ However, of the *S*-palmitoylation sites reported in proteins, very few have actually been confirmed using mass spectrometry or other analytical techniques. Usually a tentative identification of palmitoylation sites has been deduced by site-directed mutagenesis of candidate cysteine residues.¹⁹⁰

The acylated ghrelin peptide has previously been analysed by ESI-electron capture dissociation (ECD)-FTICR mass spectrometry.²³¹ The acyl moiety was found to be retained on peptide backbone cleavage fragments and the acylation was successfully identified as being attached to Ser-3. The ECD fragmentation also resulted in ester bond cleavage to cause a neutral loss of the octanoic acid from the ghrelin molecular ion. The same research group also analysed ghrelin using SORI-CAD, and found the resultant

fragmentation was not useful in localising the acyl moiety on the ghrelin peptide. The group did not specifically identify any marker type ions useful for discovering acyl modifications on peptides.

When analysing peptides or proteins of known sequence using mass spectrometry, intact mass analysis can give an indication of the number and type of acylations present. Treatment of the peptide or protein with proteolytic enzymes followed by mass spectrometry can further localize any modifications.²³² Subsequent MS/MS of the enzymatic peptides can then be used to identify the acylated amino acid residue(s) and any heterogeneity in the acyl chain length. Identification of acyl-modified amino acids within proteins by mass spectrometry has at times proved challenging. Standard analytical procedures, such as enzymatic digestion of peptides prior to mass spectrometry, have resulted in the apparent disappearance of acyl moieties that had been previously observed in the intact mass of the protein. For example, Serebryakova *et. al.* found when analysing hemagglutinin, a transmembrane glycoprotein from *Influenza A* virus, that treatment with thiol reagents as part of the enzymatic digestion procedure resulted in a decrease in the number of observed sites of acylation. They concluded that cysteine bound acylations are sensitive to the use of thiol reagents.²³³ In another study, Sachon and co-workers had been unable to identify palmitoylation on synthetically *N*-palmitoylated human growth hormone. In order to identify the palmitoylation sites successfully, enzymatic digestion followed by liquid-liquid extraction was performed directly on a MALDI sample plate.⁵⁶ The more polar palmitoylated peptides were extracted from the aqueous phase using ethyl acetate. Spotting the non-aqueous phase at a different position on the MALDI target enabled the sub-stoichiometric, multi-site palmitoylation to be identified.

Very little research has been performed to identify the most useful mass spectrometric conditions for the discovery and analysis of acylated peptides. The most systematic study to date, using MALDI and ESI tandem mass spectrometry to identify useful marker ions for acylated peptides, has been by Hoffman *et. al.* This group analysed myristoylation on N-terminal glycine and palmitoylation on cysteine residues, finding a small number of useful neutral losses and potential marker ions.²³⁴ The characteristic fragments observed were a_1 , b_1 and neutral loss ions for peptides containing *N*-terminally myristoylated glycine. For peptides containing palmitoylated cysteine, two neutral loss species were observed, being the loss of the palmitoyl moiety and the loss of a thioacid.

A thorough investigation of the mass spectrometry of acylated peptides encompassing both a variety of modified amino acids and a variety of acyl chain lengths has not been conducted to date. The purpose of this chapter is to determine the most useful tandem mass spectrometry conditions for the analysis of a variety of acylated peptides, including N-terminal-linked, *O*-linked and *S*-linked acylations.

3.2 Results and Discussion

Synthetically acylated peptides and the naturally occurring acylated peptide ghrelin were analysed using three different tandem mass spectrometry techniques: MALDI PSD MS; MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS. The influence of the collision energy on the fragmentation observed in ESI-Q-o-TOF MS/MS was also studied for a number of acylated peptides. The data processing used to create the collision energy profiles graphs is described in the methods chapter (Chapter 2.1.1).

The synthesis of acylated peptides is described in Chapter 2.4 as previously published by Yousefi-Salakdeh *et. al.*¹⁸⁰ The synthesis gave a yield of 10-50% as determined by

HPLC. Figure 3.1 shows the HPLC trace for *O*-octinoylated Eledoisin. Similar HPLC traces were observed for the purification of the remaining acylated peptides (data not shown). Due to the sensitivity of mass spectrometry, sufficient acylated peptide was produced in the initial synthesis to complete the mass spectral analysis required, hence it was not necessary to optimise reaction conditions further.

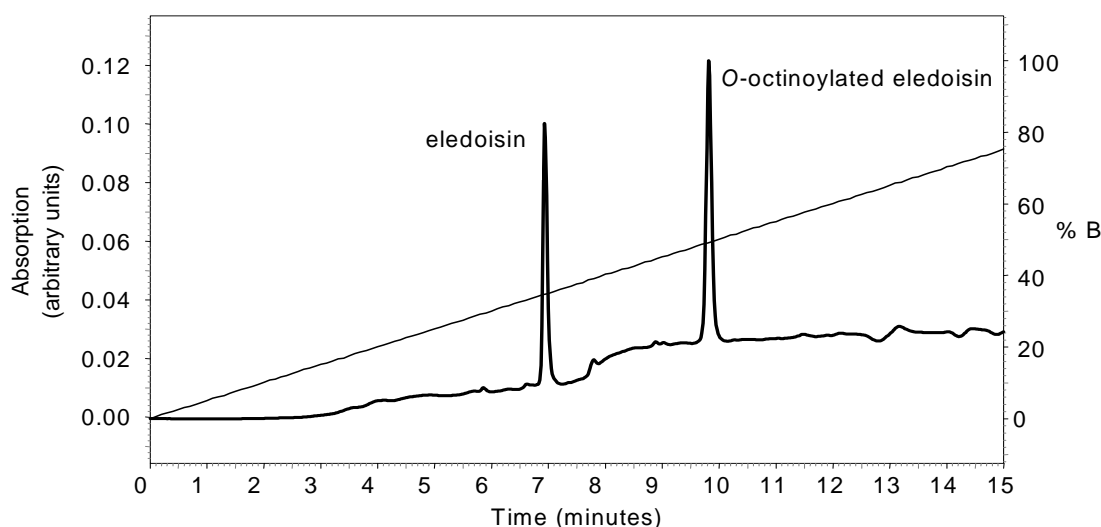


Figure 3.1 HPLC trace for the purification of *O*-octanoylated eledoisin. Buffer A: 95% H₂O, 5% ACN (acetonitrile), 0.1% TFA (trifluoroacetic acid); Buffer B: 85% ACN, 10% IPA (isopropanol), 5% H₂O, 0.1% TFA. The gradient was 0-100% B over 20 min. The detector was set at 214 nm.

3.2.1 Tandem Mass Spectrometry of *O*-Acylated Neurokinin A

Neurokinin A (also known as substance K) was selected for the study of an *O*-acylated peptide as a synthetic analogue to ghrelin. Neurokinin A contains only one serine residue compared with four in ghrelin, and is a shorter peptide (ten residues compared with twenty eight). Neurokinin A was modified with octanoyl and myristoyl acyl moieties and examined by MALDI PSD MS, MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS techniques. The mass spectra resulting from the use of these three mass spectrometry sequencing techniques for *O*-acylated and native neurokinin A are shown in Figures 3.2 – 3.4.

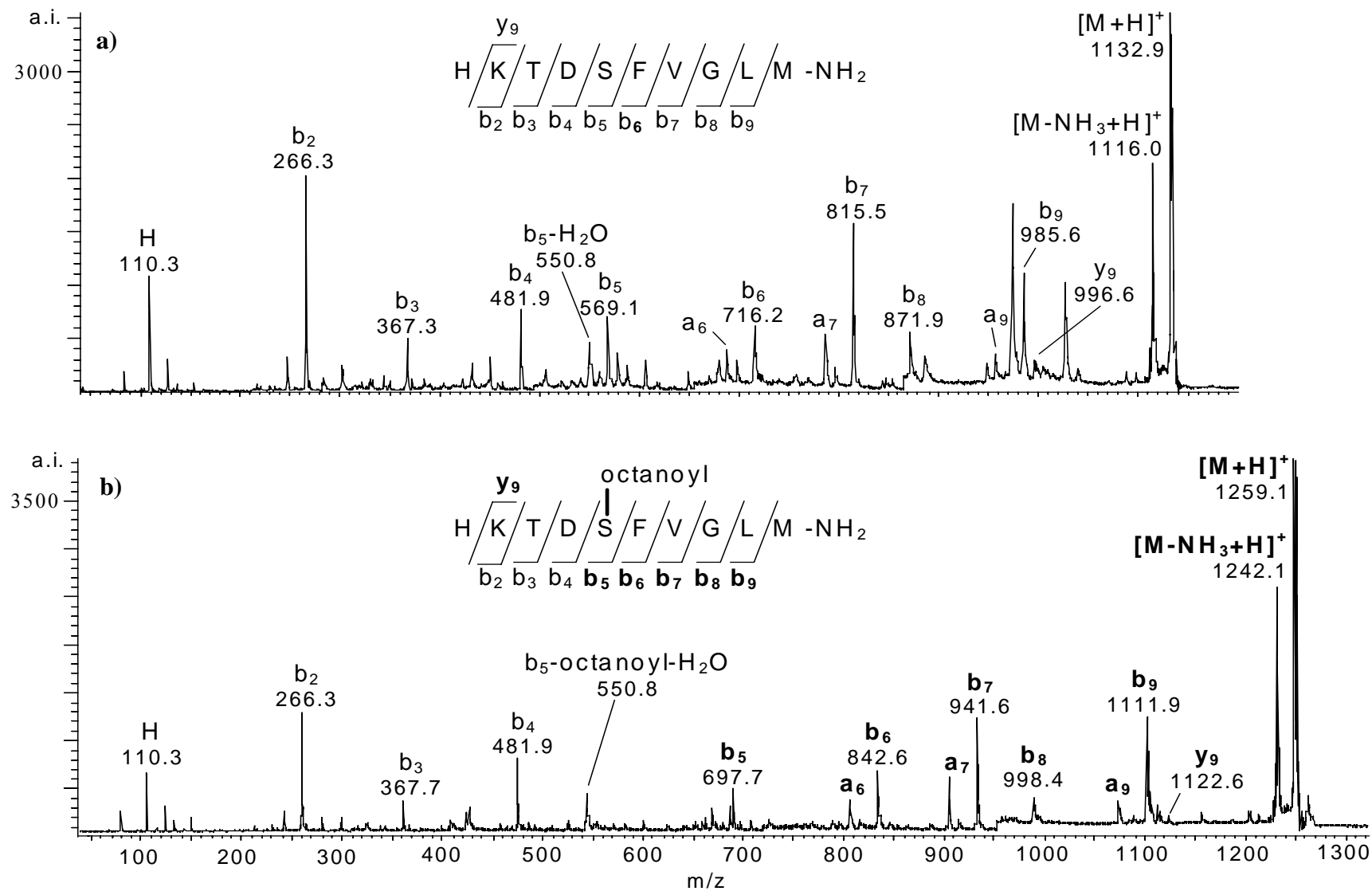


Figure 3.2 MALDI PSD mass spectra of neurokinin A (a) and *O*-octanoyl-neurokinin A (b), showing the stability of the acylated serine modification. Bold-face type indicates ions containing the acyl moiety.

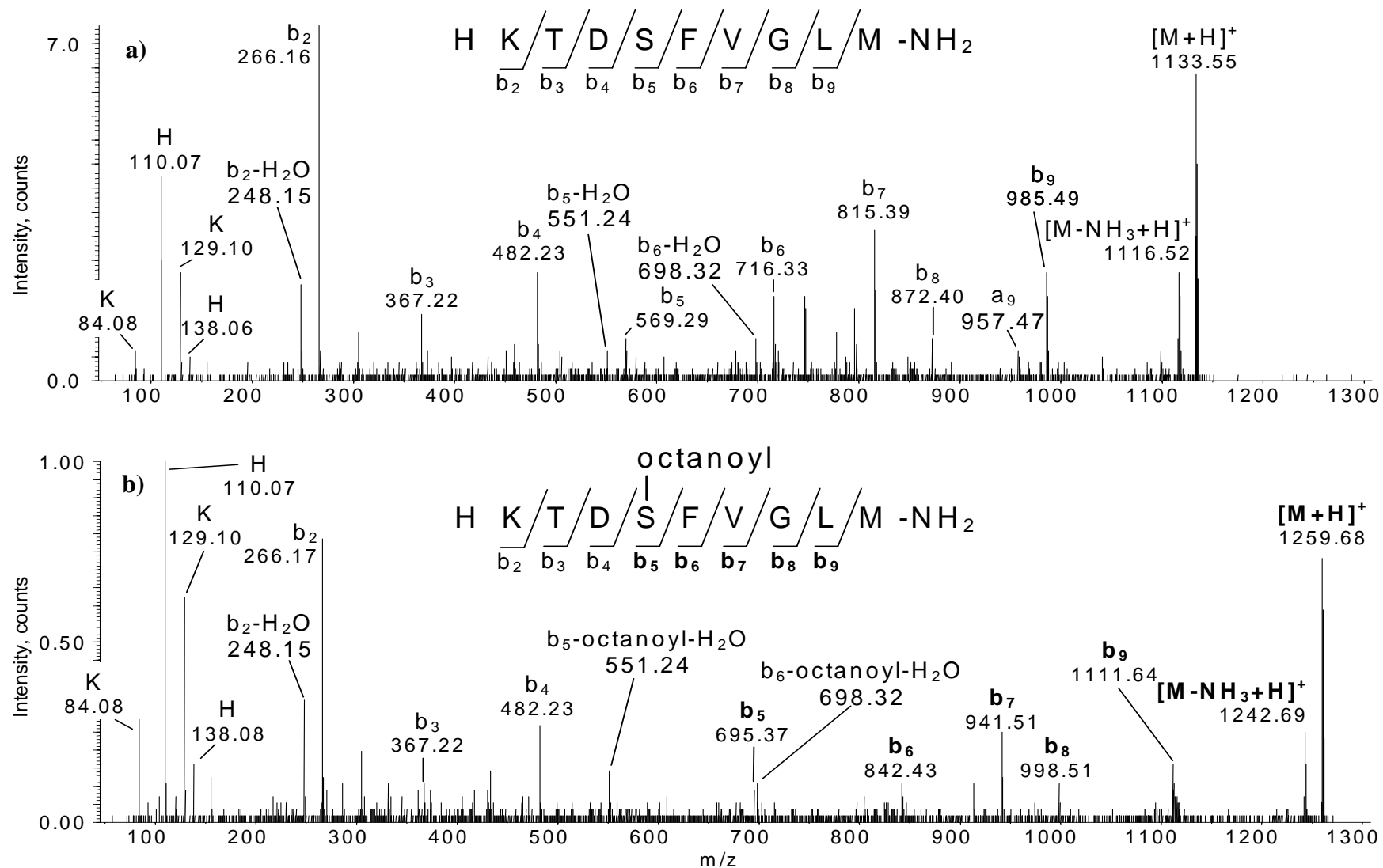


Figure 3.3 MALDI-Q-o-TOF CID MS/MS analysis of neurokinin A (C.E 110 eV) (a) and *O*-octanoyl-neurokinin A (C.E 110 eV) (b). Bold-face type indicates ions containing the acyl moiety.

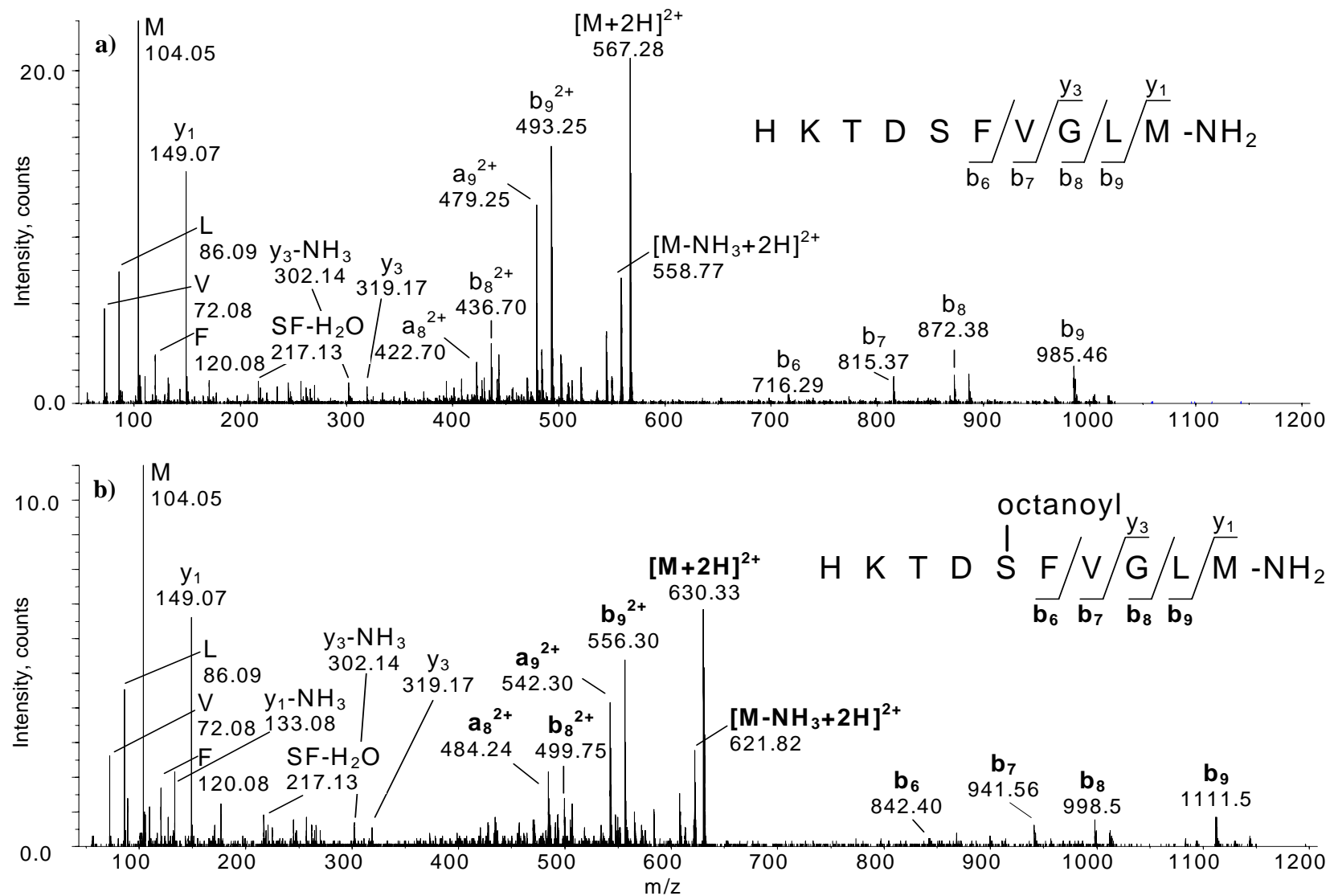


Figure 3.4 ESI-Q-o-TOF CID MS/MS analysis of neurokinin A (C.E. 25 eV; 2⁺ precursor ion) (a) and *O*-octanoyl-neurokinin A (C.E. 30 eV) (b). Bold-face type indicates ions containing the acyl moiety. (Continued on next page.)

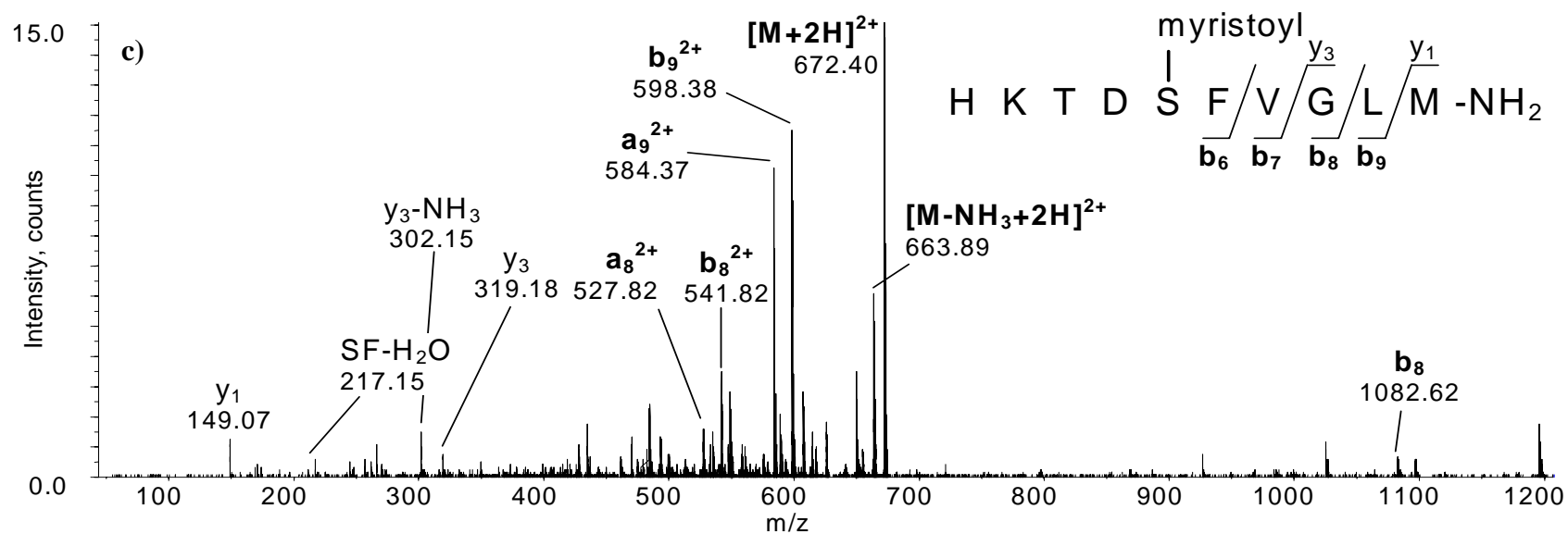


Figure 3.4 cont. ESI-Q-o-TOF CID MS/MS analysis of *O*-myristoyl-neurokinin A (C.E. 30 eV) (c). Bold-face type indicates ions containing the acyl moiety.

The MALDI PSD MS and MALDI-Q-o-TOF MS/MS (Figures 3.2 and 3.3 respectively) resulted in very similar mass spectra and clearly show the stability of the acyl modification on serine in the neurokinin A peptide. The presence of the acyl modification has no obvious influence on the fragmentation pattern observed for the neurokinin A peptide, with both the location of fragmentation along the peptide backbone and the observed intensity of the resultant ions being preserved in the acylated peptide when compared to the native peptide. The b-ion peptide backbone series predominates due to the presence of a readily protonated lysine residue toward the N-terminus. This series easily allows for the identification of the acylated residue, ser-5, by a mass difference between fragment ions b_4 and b_5 of 213 u, which is the differential mass of the serine residue (87 u) plus octanoyl moiety (126 u). A loss of the octanoyl moiety in the form of octanoic acid is observed from the b_5 ion in both MALDI mass spectra and also from the b_6 ion in the MALDI-Q-o-TOF spectrum ($b_{5/6}$ -octanoyl- H_2O). A similar loss was observed by Hoffman *et. al.* when studying S-acylation, where a loss of a palmitic thioacid was observed in MALDI mass spectra.²³⁴ There were no readily observable peaks in either MALDI mass spectra that corresponded to potential marker ions for O-acylated peptides (for example, an octanoylated serine immonium ion or an octanoyl carbenium ion²³⁵).²³⁴

The ESI-Q-o-TOF MS/MS spectra of native, O-octanoylated and O-myristoylated Neurokinin A (Figure 3.4) are noticeably different to the corresponding MALDI mass spectra. A significant difference, for Neurokinin A in particular, is the reduced number of sequence ions observed in the ESI-Q-o-TOF MS/MS spectrum. At the collision energies shown, the ESI-Q-o-TOF MS/MS spectra for the O-octanoylated and O-myristoylated neurokinin are very similar to that of the native peptide, allowing for the mass shift corresponding to the additional mass of the respective acyl moieties. (Note: a

collision energy of 30 eV was required to fragment the acylated neurokinin A peptides to the same degree as the native neurokinin a at collision energy of 25 eV.) The ESI-Q-o-TOF mass spectra are dominated by $b_6 - b_9$ ions in the singly and doubly charged states. Similar to the MALDI mass spectra, the acyl moiety remains intact on the observed sequence ions. At the collision energies used for the ESI-Q-o-TOF MS/MS mass spectra shown in Figure 3.4, fragmentation due to neutral loss of either the acyl moiety or of an acyl carboxylic acid from sequence ions was significantly less pronounced compared to the MALDI spectra. The peptide sequence ions observed did not alter significantly with a change in the collision energy (data not shown), however, the ion intensity of the larger sequence ions decreased with an increase in collision energy and the ion intensity of smaller fragment ions increased. At higher collision energies the neutral loss of both the acyl moiety and acyl carboxylic acid (independently) from the precursor ion is observed. This is in contrast to Hoffman *et. al.* who observed only the neutral loss, without accompanying SH_2 loss in their ESI-mass spectrometry study of *S*-acylation.²³⁴ In addition, acylated serine immonium ions and acyl carbenium ions are observed at higher collision energies. A summary of the characteristic marker ions and useful sequence ions that were observed in the MALDI PSD MS, MADLI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS spectra of *O*-acylated neurokinin is given in Table 3.1.

Table 3.1 Characteristic and useful sequence ions identified for *O*-acylated neurokinin A.

Peptide	Characteristic and useful sequence ions	Optimum collision energy ^a (eV)
<i>O</i> -octanoyl neurokinin A HKTDS(oco)FVGLM-NH ₂	<p><i>ESI-Q-o-TOF</i> (2+ charge state):</p> <p>neutral loss of octanoyl (-126 u)</p> <p>neutral loss of octanoic acid (-144 u)</p> <p>immonium ion <i>O</i>-octanoyl-S (<i>m/z</i> 186)</p> <p>octanoyl carbenium ion (<i>m/z</i> 127)</p> <p><i>MALDI PSD</i>:</p> <p>b₅-neutral loss of octanoic acid (-144 u)</p> <p>b₄, b₅ sequence ions with additional mass shift of <i>m/z</i> 126</p> <p><i>MALDI-Q-o-TOF</i>:</p> <p>b₅-neutral loss of octanoic acid (-144 u)</p> <p>b₆-neutral loss of octanoic acid (-144 u)</p> <p>b₄, b₅ sequence ions with additional mass shift of <i>m/z</i> 126</p>	<p>35</p> <p>35</p> <p>75+</p> <p>45+</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>
<i>O</i> -myristoyl neurokinin A HKTDS(myrist)FVGLM-NH ₂	<p><i>ESI-Q-o-TOF</i> (2+ charge state):</p> <p>neutral loss of myristanoyl -210 u) [C.E. 35-45 eV]</p> <p>neutral loss of myristanoic acid (-228 u) [C.E. 35 eV]</p> <p>immonium ion <i>O</i>-myristoyl-S (<i>m/z</i> 270) [C.E. 45+ eV]</p> <p>myristoyl carbenium ion (<i>m/z</i> 211) [C.E. 45+ eV]</p>	<p>35-45</p> <p>35</p> <p>45+</p> <p>45+</p>

^aApproximate optimum collision energy as determined by ESI-Q-o-TOF MS/MS, and estimated from plots in Figure 3.5.

To investigate the influence of collision energy on the fragmentation and production of acylated marker ions, collision energies ranging from 10 to 75 eV were employed to fragment native, *O*-octinoylated and *O*-myristoylated neurokinin. For acylated neurokinin A, only a small degree of fragmentation was observed below 30 eV, hence the plots are shown over the range 30-75 eV for *O*-acylated neurokinin A (Figure 3.5).

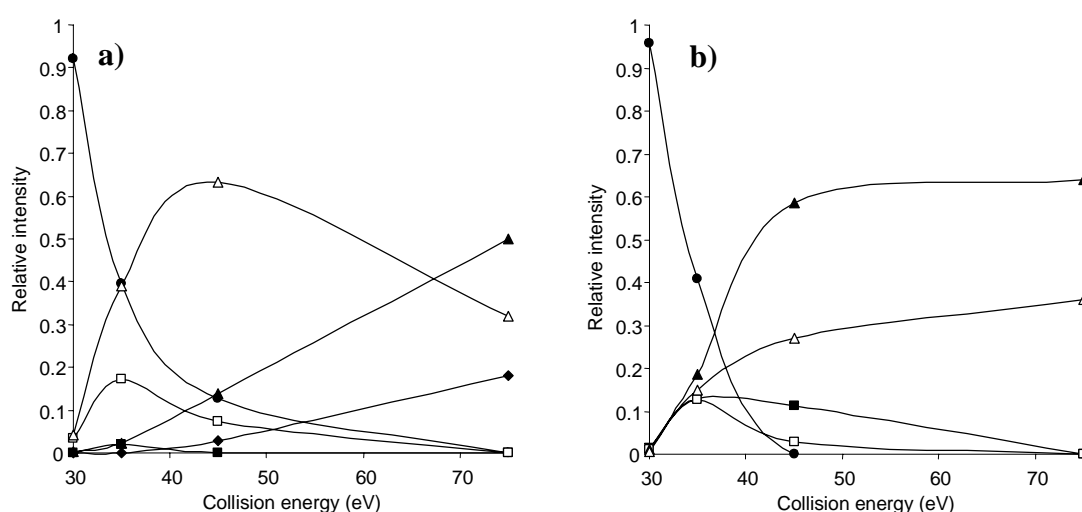


Figure 3.5 Collision energy profiles for the precursor ion and product ions of: *O*-octanoyl neurokinin A (a) and *O*-myristoyl neurokinin A (b) (2⁺ precursor ions). Legend: precursor ion, [M+2H]²⁺, (●); neutral loss, [M-acyl+2H]²⁺, (■); neutral loss, [M-acyl carboxylic acid+2H]²⁺, (□); acylated serine immonium ion, (▲); acyl carbenium ion, CH₃(CH₂)_(n-2)CO⁺, (Δ); des-acyl serine immonium ion, (◆).

The initial fragmentation of *O*-octanoyl neurokinin A shows the loss of the acyl moiety as a neutral species, either as a loss of 126 u (C₈H₁₄O) or as a loss of the carboxylic acid form, delta mass 144 u (C₈H₁₆O₂). The more abundant ion of these two species is formed as a result of the loss of a carboxylic acid molecule. These neutral loss species reach a maximum at approximately 35 eV. There was no obvious loss of H₂O from the native neurokinin A precursor ion, although this is a common neutral loss from serine containing peptides.¹⁶¹ This suggests that the additional loss of H₂O from the acylated

precursor peptide is directly related to the presence of the acyl moiety, with the acyl group being lost from the precursor ion as octanoic acid.

The predominant marker ion indicating the presence of an octanoylated serine residue, is the octanoylated serine immonium ion (m/z 186). A des-acyl serine immonium ion is also present in the spectra of *O*-octanoyl neurokinin A; this product ion may be the result of a further fragmentation of the acylated serine immonium ion, or result from a further fragmentation of the $[M\text{-octanoyl}+2H]^{2+}$ ion. The relative abundance of both the octanoylated serine immonium ion and the des-acyl serine immonium ion increase with increasing collision energy. The corresponding fragments in the MS/MS of native neurokinin A show a simple relationship between the decrease in the intensity of the precursor ion and the increase in the intensity of the serine immonium ion at increasing collision energies.

The predominant acyl marker ion, which is not also indicative of an octanoylated serine residue, is an octanoyl carbenium ion ($CH_3(CH_2)_6CO^+$, m/z 211). The relative intensity of this ion peaks at a collision energy of approximately 45 eV. At collision energies up to at least 45 eV, the octanoyl carbenium ion is significantly more abundant than the octanoylated serine immonium ion. It is possible that, for *O*-octanoylated neurokinin A, this ion is more intense than the ion corresponding to the acylated serine immonium ion as only one bond must be broken to form the octanoyl carbenium ion whereas two peptide backbone bonds are required to be broken to form the octanoylated serine immonium ion.

The initial fragmentation of *O*-myristoyl neurokinin A is similar to that of *O*-octanoyl neurokinin A with the loss of the acyl moiety as a neutral species, either as a loss of 210 u ($C_8H_{14}O$) or as a loss of myristic acid, 228 u ($C_8H_{16}O_2$). The relative abundance

of the myristoyl neutral loss peaks between 35 and 45 eV, and at approximately 35 eV for the myristic acid. However, fragmentation is different, as the ion resulting from myristoyl neutral loss is more similar in abundance to the ion resulting from the myristic acid loss, whereas the octanoic acid loss is more prevalent in the case of *O*-octanoyl neurokinin A. This may result from the increased stability of the myristoyl molecule due to its increased chain length, relative to the energy required to produce the corresponding myristic acid fragment. The fragmentation of *O*-myristoylated neurokinin A also resulted in the formation of a myristoylated serine immonium ion (m/z 270) and a myristoyl carbenium ion (m/z 211). In contrast to the case of octanoyl the myristoylated serine immonium ion is in greater abundance than the myristoyl carbenium ion. It is possible the difference in stability between the *O*-octanoylated and *O*-myristoylated neurokinin A acyl bonds is due to the increased chain length of the myristoyl moiety. The fragmentation of *O*-myristoyl neurokinin A is also marked by the absence of the formation of a des-acyl serine immonium ion (m/z 60). This is a further indication of the increased stability of the myristoyl modified neurokinin A over the octanoyl form.

3.2.2 Tandem Mass Spectrometry of *O*-acylated Eleodoisin

Eleodoisin is a peptide with a similar sequence to Neurokinin A. They consist of identical C-termini (Gly-Leu-Met-NH₂) and contain one serine residue: neurokinin A at residue 5 and eleodoisin at residue 3. In this work, eleodoisin was modified with the octanoyl acyl moiety and examined by MALDI PSD MS, MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS techniques. The resultant spectra for the fragmentation of *O*-octanoyl serine are shown in

Figure 3.6 – 3.8 below. As was the case for *O*-acylated neurokinin A, the spectra produced for MALDI PSD MS and MALDI-Q-o-TOF MS/MS are very similar to each

other. One noticeable difference between the spectra produced for *O*-octanoyl eledoisin and *O*-acylated neurokinin A, is the reduced sequence coverage for *O*-octanoyl eledoisin. The three mass spectrometry techniques used resulted in tandem or pseudo tandem mass spectra with incomplete backbone sequence coverage. The two MALDI MS techniques resulted in spectra containing only b-ions. The ESI-Q-o-TOF MS/MS spectrum resulted in sequence ions predominantly from cleavages at the C-terminal end of the peptide. The spectra for octanoylated eledoisin were essentially the same as that of native eledoisin. Only the MALDI-Q-o-TOF MS/MS gave sufficient sequence coverage to determine the position of the octanoyl addition at Ser-3, by a mass difference between fragment ions b_3 and b_4 of 213 u, which is the differential mass of the serine residue (87 u) plus octanoyl (126 u).

At the collision energies used in the ESI-Q-o-TOF MS/MS analysis of native and acylated eledoisin (over a range of 15 – 70 eV, with the spectra shown at 25 eV and 30 eV respectively) there were no distinctive marker ions, nor neutral losses (either as octanoyl or as octanoic acid) observed for the octanoyl moiety on the Ser-3 residue of *O*-octanoyl eledoisin. In addition, no neutral losses were observed in either of the MALDI mass spectrometry techniques employed, where neutral loss of the octanoyl moiety had been observed using both MALDI MS techniques for *O*-octanoyl neurokinin A. The lack of useful ions produced from eledoisin may be due to the presence of a proline residue to the immediate N-terminal side of the acylated serine residue. This

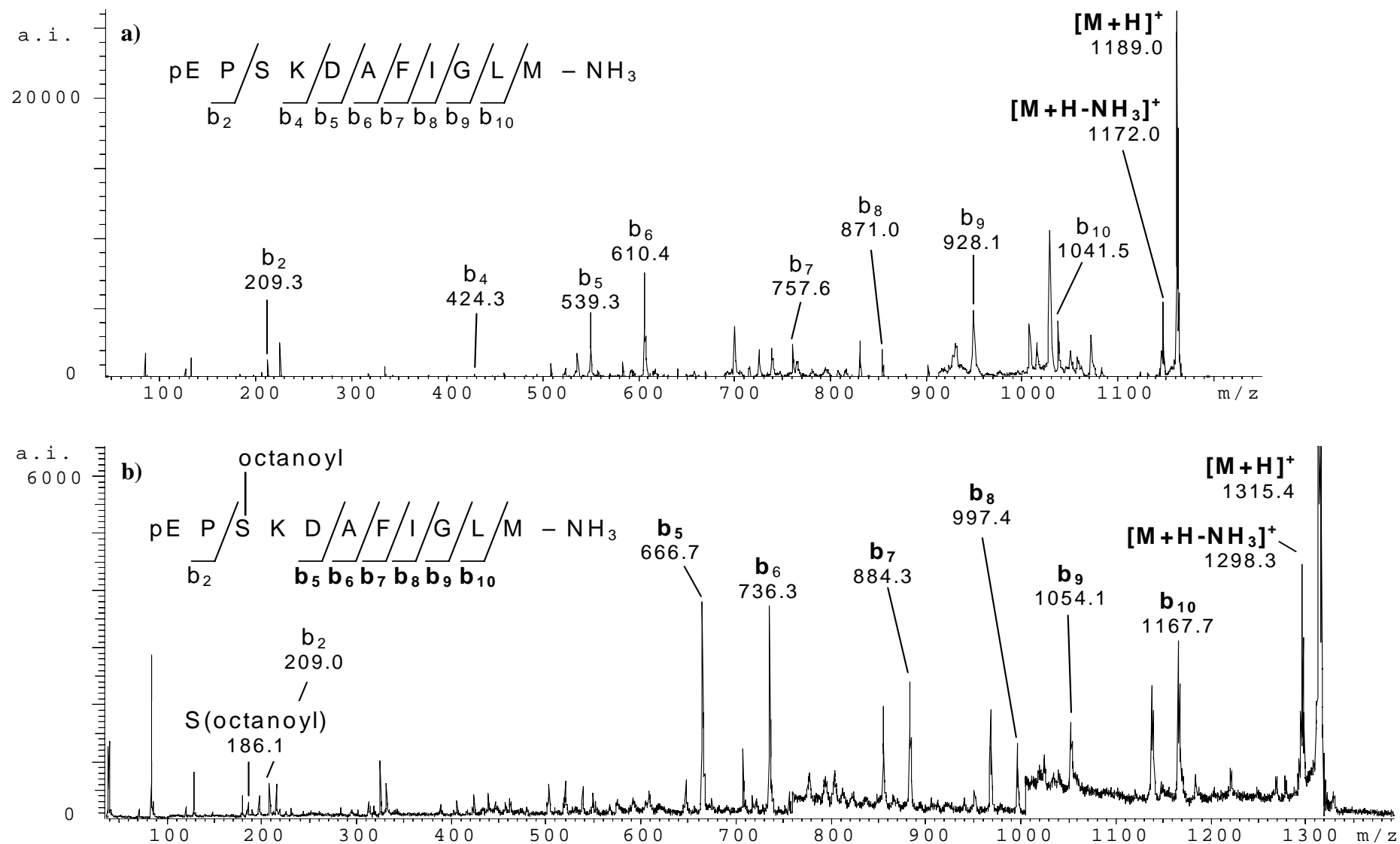


Figure 3.6 MALDI PSD mass spectra of eldoisin (a) and *O*-octanoyl eldoisin (b). Bold-face type indicates ions containing the acyl moiety.

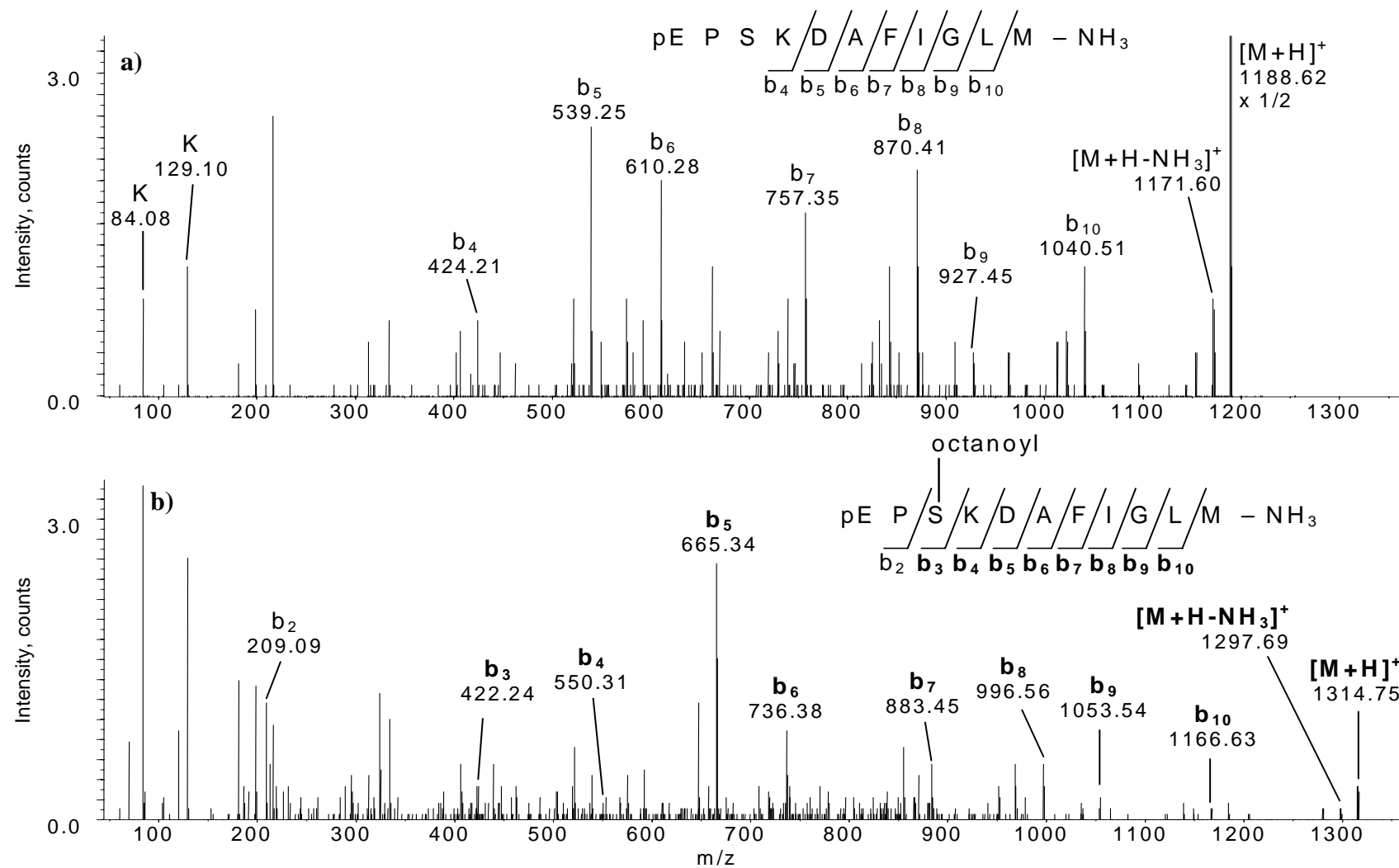


Figure 3.7 MALDI-Q-o-TOF CID MS/MS of eldoisin (C.E. 110 eV) (a) and *O*-octanoyl eldoisin (C.E. 110 eV) (b). Bold-face type indicates ions containing the acyl moiety.

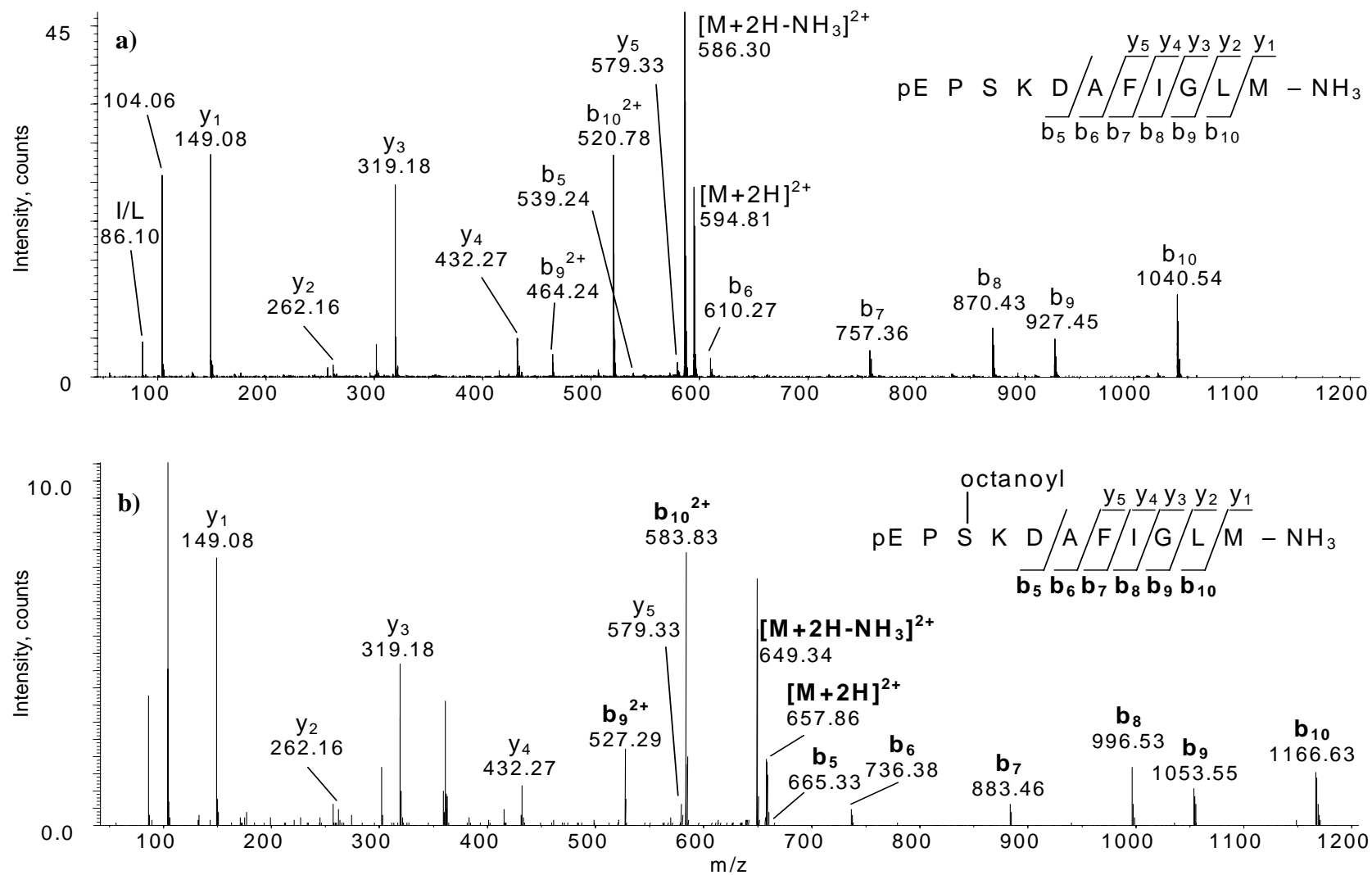


Figure 3.8 ESI-Q-o-TOF CID MS/MS of eledoisin (C.E. 25 eV; 2⁺ precursor ion) (a) and *O*-octanoyl eledoisin (C.E. 30 eV) (b). Bold-face type indicates ions containing the acyl moiety.

residue may provide a stabilising influence on the acyl moiety, or otherwise reduce interactions or collisions at the acyl site due to steric considerations.

3.2.3 Tandem Mass Spectrometry of Ghrelin

Ghrelin (an *O*-octanoylated peptide) contains four serine residues and is octanoylated at Ser-3 (in the human and rat forms). Human ghrelin was analysed by MALDI PSD MS and ESI-Q-o-TOF MS/MS. The MALDI PSD mass spectra (Figure 3.9) shows a number of y sequence ions, which are expected for a peptide with an arginine at the C-terminus. MALDI PSD MS is rarely used for peptides which are as large as ghrelin (protonated average mass: 3372 u) as the MALDI process does not impart sufficient energy to easily fragment peptides above *ca.*1500 u.^{236, 237} In addition, mass accuracy and resolution can be problematic when the instrument is not calibrated over the required mass range, as evident from the molecular weight shown in Figure 3.9 of *m/z* 3366. The PSD function use in this analysis was calibrated using ACTH fragment 18-39, with an average mass of 2465 u. The MALDI-Q-o-TOF could not be used for analysis of the ghrelin sample due to the restricted mass range of the instrument at the time of analysis. (High-mass rods are available for a MALDI-QSTAR mass spectrometer; however, they were not installed on the mass spectrometer used in this analysis.) A potential advantage of the MALDI-Q-o-TOF mass spectrometer is the ease of calibrating a required mass range compared with a MALDI PSD mass spectrometer, where calibration can be an arduous task. To overcome the mass range limitation of the MALDI-Q-o-TOF instrument, the octanoyl-containing tryptic peptide of ghrelin was analysed (see Section 3.2.4 below). Consequently, MALDI MS analysis under the usual conditions found in a mass spectrometry laboratory was not found to be useful in the analysis of large intact acylated peptides such as ghrelin.

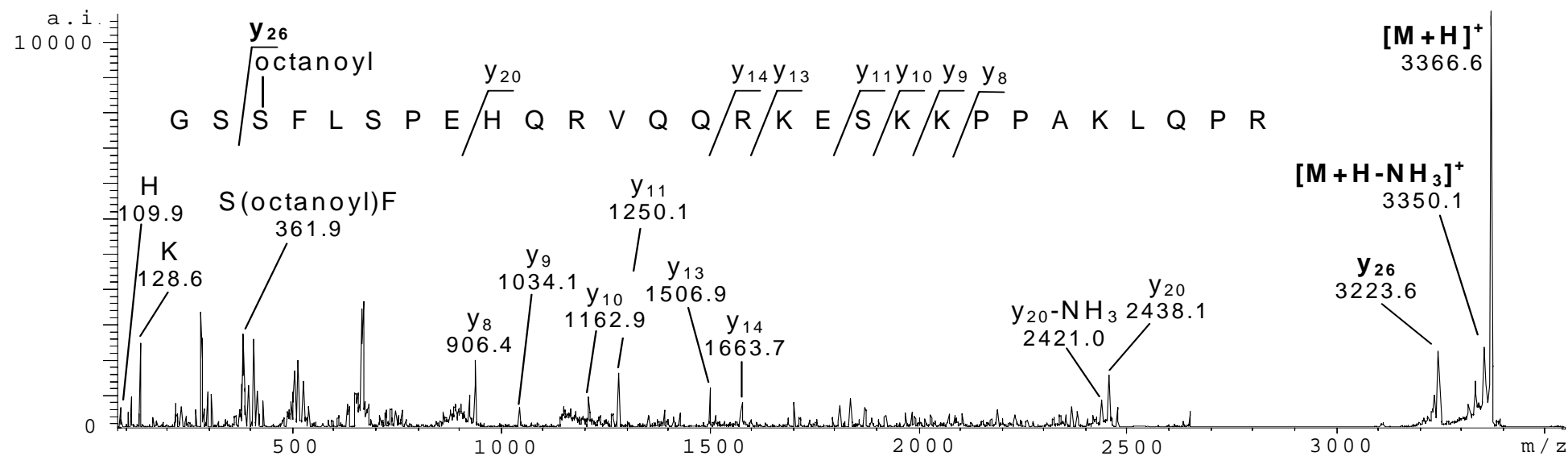


Figure 3.9 MALDI PSD mass spectrum of ghrelin. Bold-face type indicates ions containing the acyl moiety.

ESI-Q-o-TOF MS/MS spectra were obtained for both des-acyl ghrelin and ghrelin (see Figure 3.10 and Figure 3.11 respectively; each spectrum is shown over two lines for clarity). Des-acyl ghrelin was prepared by treating ghrelin with ammonium hydroxide and purified by HPLC, as described previously in Chapter 2.2. The spectra are very similar for the acylated and des-acyl forms of ghrelin, as found for the other acylated and native peptides analysed in this chapter. A neutral loss of H₂O is readily observed from the intact des-acyl ghrelin molecule, and also from the (acylated) ghrelin molecule. An ion at m/z 438.49 in the ESI-Q-o-TOF MS/MS spectrum of ghrelin can be attributed to an apparent loss of octanoic acid (144 u).

In contrast to neurokinin A, the potential marker ions octanoylated serine immonium ion (m/z 186) and octanoyl carbenium ion (m/z 127), are readily observed at the collision energy used to produce useful sequence fragmentation of the ghrelin peptide (see Figure 3.11). This is most likely due to the overall size of the ghrelin molecule relative to the size of the octanoyl moiety. The most abundant marker ions observed for MS/MS fragmentation of ghrelin were the ion resulting from a neutral loss of octanoic acid (-144 u), the *O*-octanoylated serine immonium ion (m/z 186) and the octanoyl carbenium ion (m/z 127). The ion resulting from the neutral loss of the octanoyl moiety (-126 u) was of relatively low abundance. The spectra shown in Figures 3.10 and 3.11 are for the 6+ charge-state of ghrelin and des-acyl ghrelin. ESI-Q-o-TOF MS/MS spectra were obtained for 4+ to 7+ charge states, with the spectrum resulting from the fragmentation of the 6+ ion resulting in the greatest number of useful sequence ions. A summary of the characteristic marker ions and useful sequence ions that were observed in the ESI-Q-o-TOF MS/MS spectra of ghrelin can be found in Table 3.2.

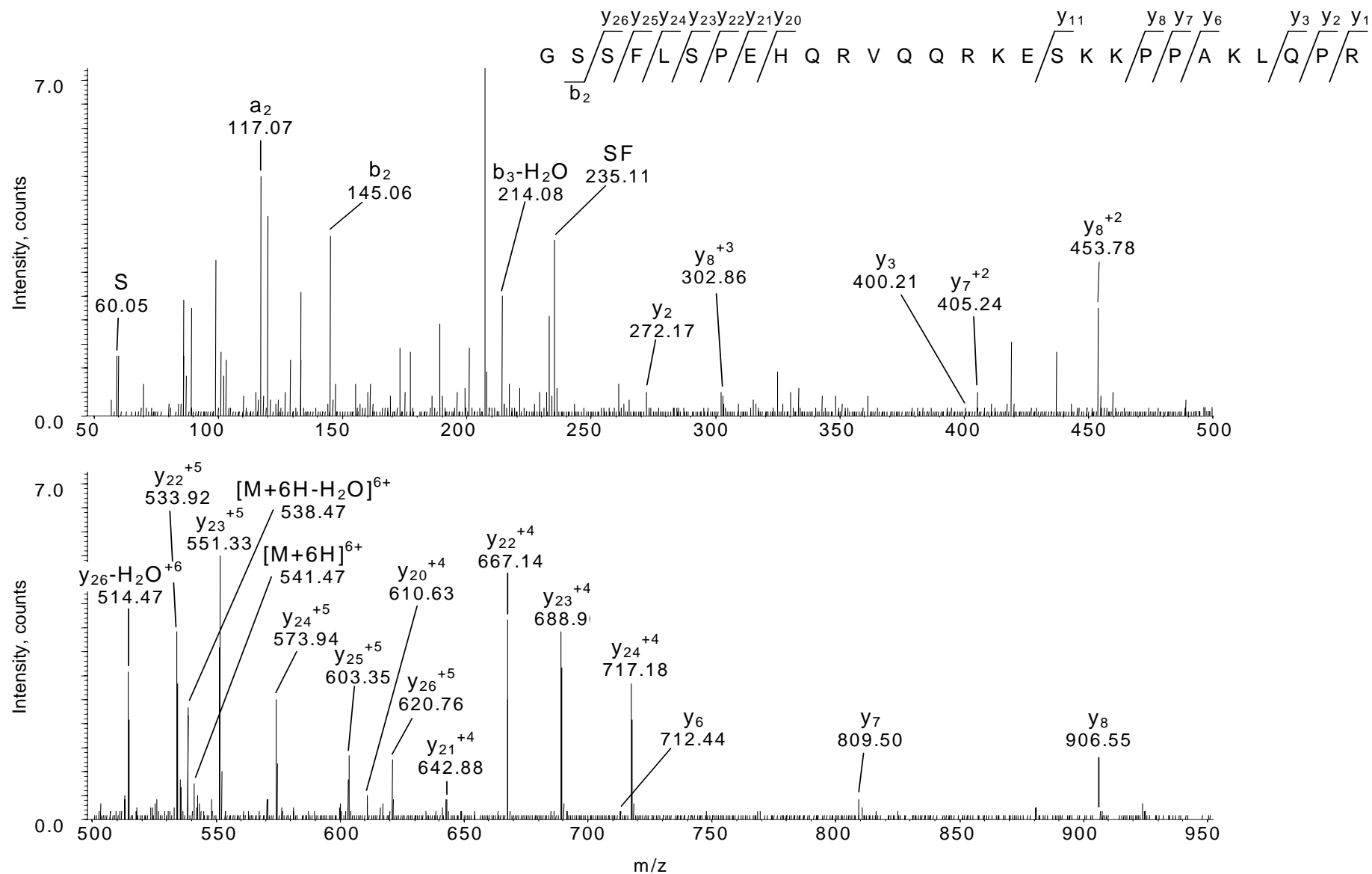


Figure 3.10 ESI-Q-o-TOF CID MS/MS of des-acyl ghrelin. Bold-face type indicates ions containing the acyl moiety; internal fragment ion labels are omitted for clarity. (C.E. 25 eV; 6+ precursor ion.)

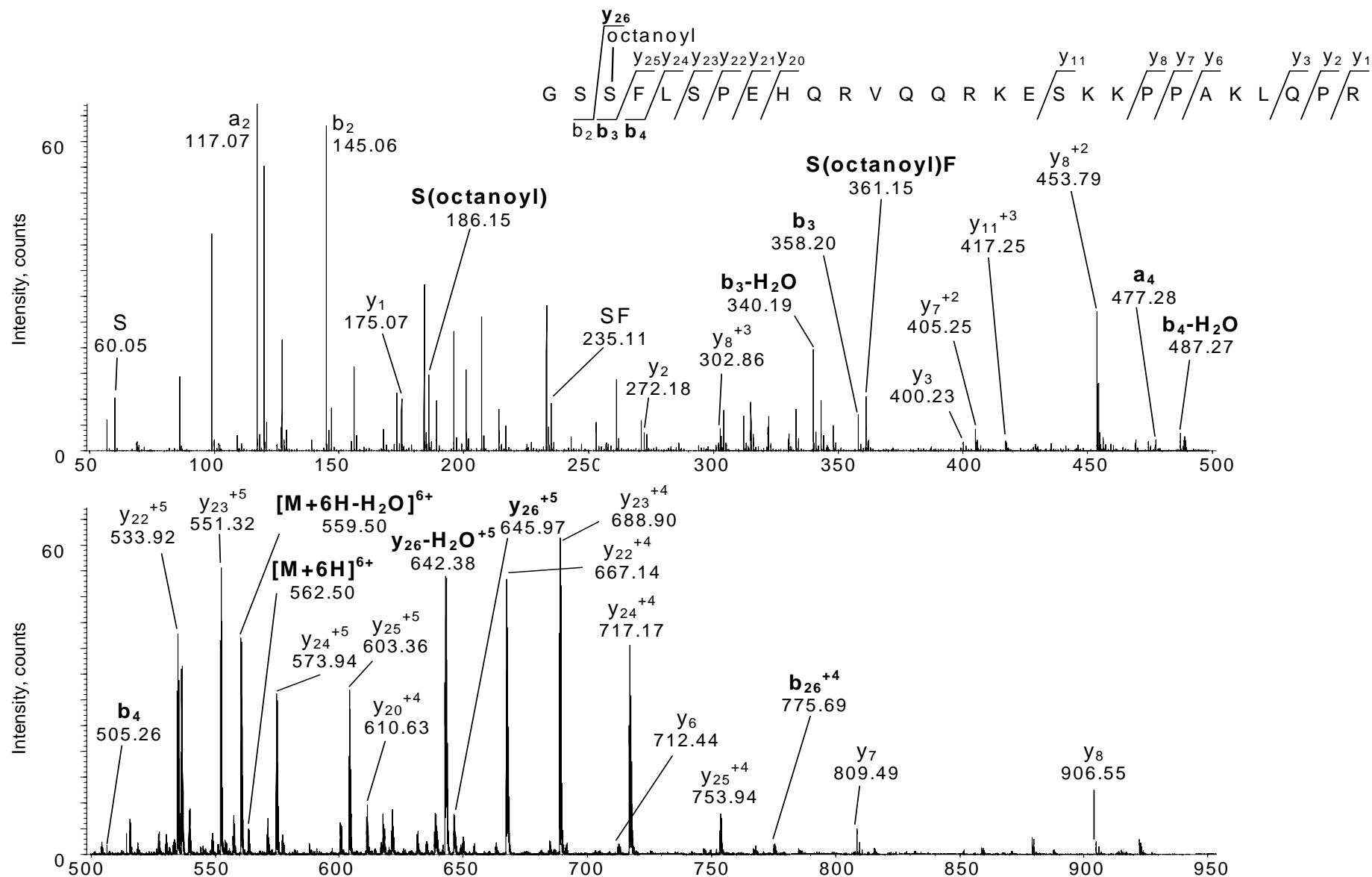


Figure 3.11 ESI-Q-o-TOF CID MS/MS of ghrelin. Bold-face type indicates ions containing the acyl moiety; internal fragment ion labels are omitted for clarity; C8:0 = octanoyl carbenium ion. (C.E. 25 eV; 6+ precursor ion.)

Table 3.2 Characteristic and useful sequence ions identified for ghrelin.

Peptide	Characteristic and useful sequence ions	Optimum collision energy ^a (eV)
GSS(oco)FLSPEHQRVQ QRKESKKPPAKLQPR	<i>ESI-Q-o-TOF</i> : neutral loss of octanoyl (-126 u) ^b neutral loss of octanoic acid (-144 u) immonium ion <i>O</i> -octanoyl-S (m/z 186) octanoyl carbenium ion (m/z 127)	25 25 35 35+

^a Approximate optimum collision energy as determined by ESI-Q-o-TOF MS/MS, and estimated from plots in Figure 3.12.

^b Low abundance ion.

A study of the influence of collision energy on the formation of marker ions for peptide acylation was conducted for ghrelin. The data for the graph shown in Figure 3.12 was computed as described previously in Chapter 2.1.1. The fragmentation and subsequent production of marker ions in ghrelin is slightly more complicated than that for neurokinin A as ghrelin contains four serine residues, compared to the single serine residue of neurokinin A. The majority of the serine immonium ion formation can be accounted for by the unacylated serine residues and some as a secondary fragmentation from the octanoylated serine immonium ion.

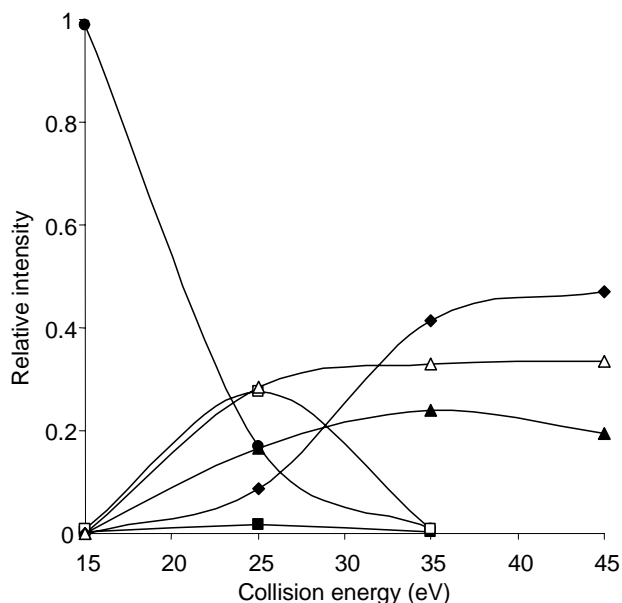


Figure 3.12 Collision energy profile for the precursor ion and product ions of ghrelin ($6+$ precursor ion). Legend: precursor ion, $[M+6H]^{6+}$ (●); neutral loss, $[M\text{-octanoyl}+6H]^{6+}$ (■); neutral loss, $[M\text{-octanoic acid}+6H]^{6+}$ (◻); *O*-octanoylated serine immonium ion (▲); octanoyl carbenium ion, $\text{CH}_3(\text{CH}_2)_6\text{CO}^+$ (Δ); native/des-acyl immonium ion (◆).

3.2.4 Tandem Mass Spectrometry of Ghrelin Tryptic Peptide T1-11

The ghrelin peptide was found to be too long for facile fragmentation in a MALDI mass spectrometer. Therefore, ghrelin was digested using trypsin to enable MALDI MS analysis of this naturally occurring acylated peptide. This also enabled investigation of the influence of peptide chain length on the resultant fragmentation spectra.

The ghrelin tryptic peptide T1-11 was successfully analysed by MALDI PSD, MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS (See Figure 3.13 - Figure 3.15). The spectra are dominated by *y* sequence ions as expected for a tryptic peptide, in this case containing a basic arginine residue at the C-terminus. Both the MALDI PSD MS and ESI-Q-o-TOF MS/MS spectra allow for the confirmation of the octanoyl containing residue by mass difference between the y_8 and y_9 sequence ions (213 u, being the sum of a serine residue, 87 u, and the mass difference of an octanoyl moiety, 126 u).

The MALDI PSD mass spectrum (Figure 3.13) shows nine of the possible ten y sequence ions, and one b ion. This spectrum is similar to the full-length ghrelin MALDI PSD mass spectrum (Figure 3.9) as it is dominated by y ions. This is explained by the fact that both sequences contain an arginine residue at their C-termini. Although the MALDI spectra of the full-length ghrelin peptide and the tryptic peptide T1-11 are similar in terms of the ions produced, the spectrum for the shorter tryptic peptide has significantly better sequence coverage. These results indicate that MALDI PSD MS is suitable for the analysis of naturally occurring acylated peptides given that longer peptides can be treated enzymatically to produce peptides of a more suitable length.

In comparing the ESI-Q-o-TOF MS/MS spectra of the full-length (Figure 3.11) and tryptic peptide T1-11 (Figure 3.15), it is found that the fragmentation is very similar for the portion of the full-length ghrelin that is represented in the tryptic peptide. The ESI-Q-o-TOF MS/MS spectrum of ghrelin contains the sequence ions b_2 - b_4 and y_{20} - y_{26} , compared to the ions b_2 , b_3 and y_1 - y_9 for the ESI-Q-o-TOF MS/MS of the tryptic peptide (y_9 in the tryptic peptide being equivalent to the y_{26} ion in the full-length ghrelin peptide). Table 3.3 summarises the characteristic and useful ions found in the ESI-Q-o-TOF MS/MS spectrum of ghrelin tryptic peptide T1-11.

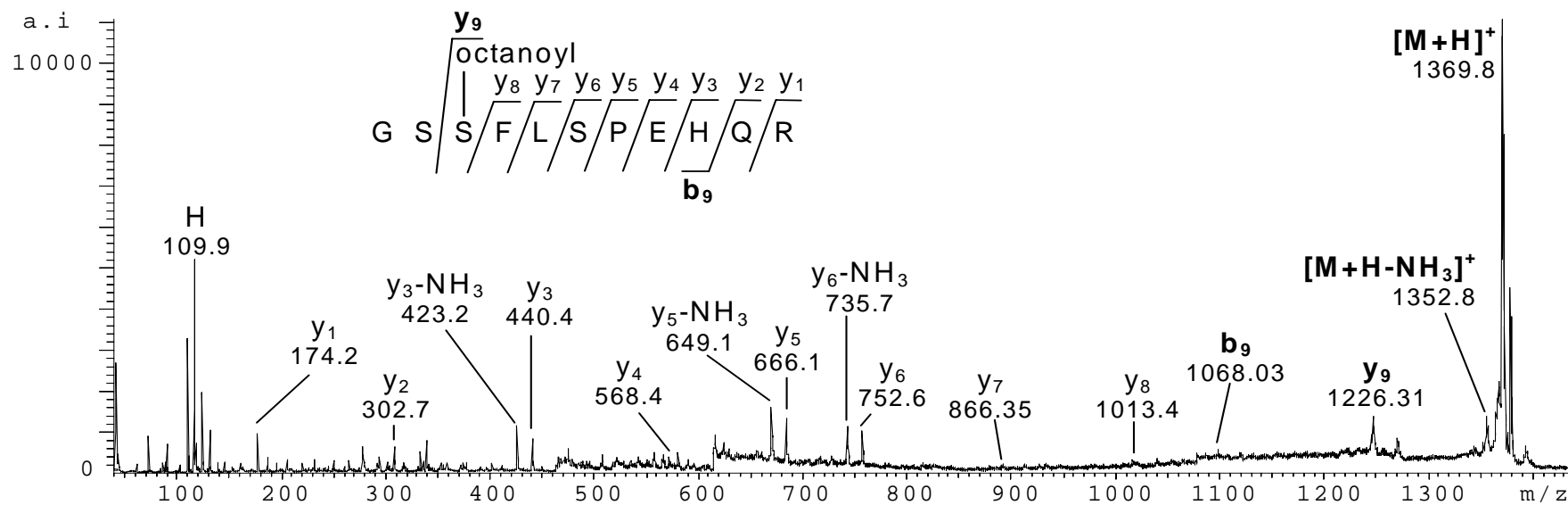


Figure 3.13 MALDI PSD mass spectrum of the ghrelin tryptic peptide T1-11. Bold-face type indicates ions containing the acyl moiety.

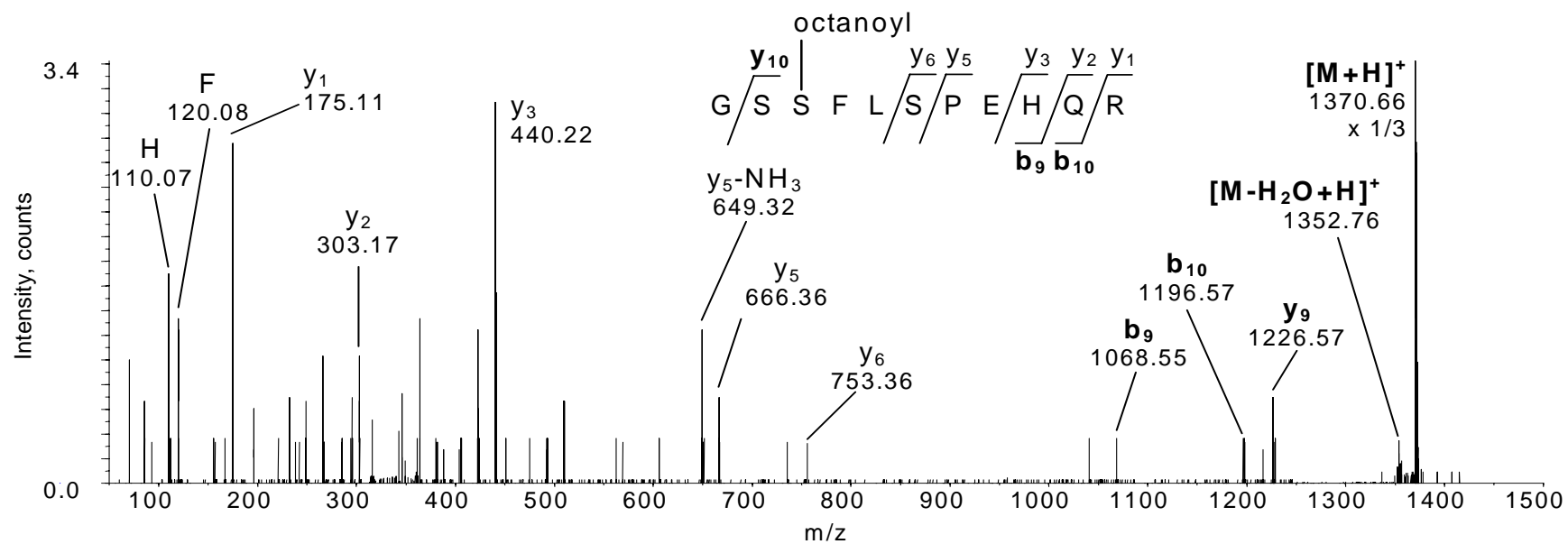


Figure 3.14 MALDI-Q-o-TOF CID MS/MS of the ghrelin tryptic peptide T1-11. Bold-face type indicates ions containing the acyl moiety. (C.E.110 eV)

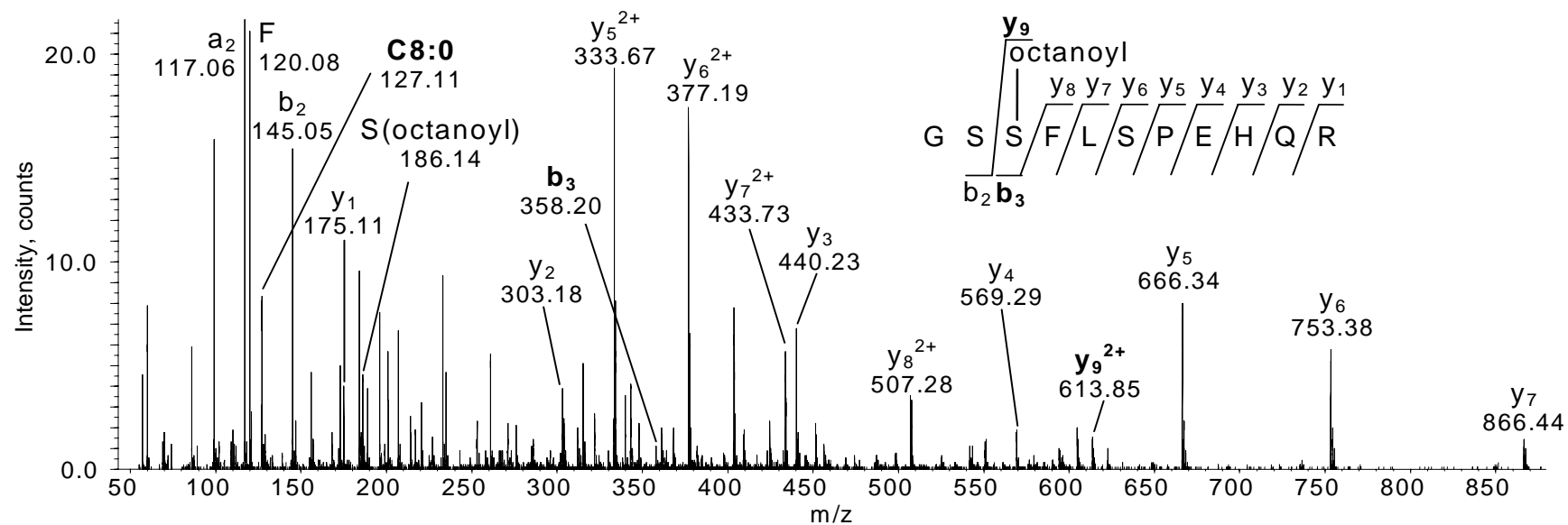


Figure 3.15 ESI-Q-o-TOF CID MS/MS of the ghrelin tryptic peptide T1-11 (C.E. 25 eV; 3+ precursor ion). Bold-face type indicates ions containing the acyl moiety; C8:O = octanoyl carbenium ion.

Table 3.3 Characteristic and useful sequence ions identified for ghrelin tryptic peptide T1-11.

Peptide	Characteristic and useful sequence ions	Optimum collision energy ^a (eV)
GSS(oco)FLSPEHQR	<i>ESI-Q-o-TOF:</i> neutral loss of octanoyl (-126 u) ^b neutral loss of octanoic acid (-144 u) immonium ion <i>O</i> -octanoyl-S (m/z 186) octanoyl carbenium ion (m/z 127)	35 25-35 25-35 25-35

^aApproximate optimum collision energy as determined by ESI-Q-o-TOF MS/MS, and estimated from plots in Figure 3.16.

^b Low abundance ion.

Figure 3.16 shows the effect of collision energy on the formation of marker ions for the ESI-Q-o-TOF MS/MS of the ghrelin tryptic peptide T1-11. The formation of marker ions was found to be very similar to that observed for the full-length ghrelin peptide. The more abundant marker ions being the ion resulting from the neutral loss of octanoic acid, an acylated serine immonium ion (m/z 186), and an acyl carbenium ion, $\text{CH}_3(\text{CH}_2)_6\text{CO}^+$ (m/z 127). This was also observed for the fragmentation of the full-length ghrelin peptide (Figure 3.12). The low abundance of the ion resulting from the neutral loss of octanoic acid (m/z 415.6) is also similar to that observed for the full-length ghrelin. In contrast to the fragmentation for the full-length ghrelin peptide, the abundance of the serine immonium ions increases at higher collision energies in the MS/MS of the ghrelin tryptic peptide T1-11. This is expected, as the octanoyl-containing tryptic peptide contains two serine residues (one of which is octanoylated) in comparison to four in the intact peptide.

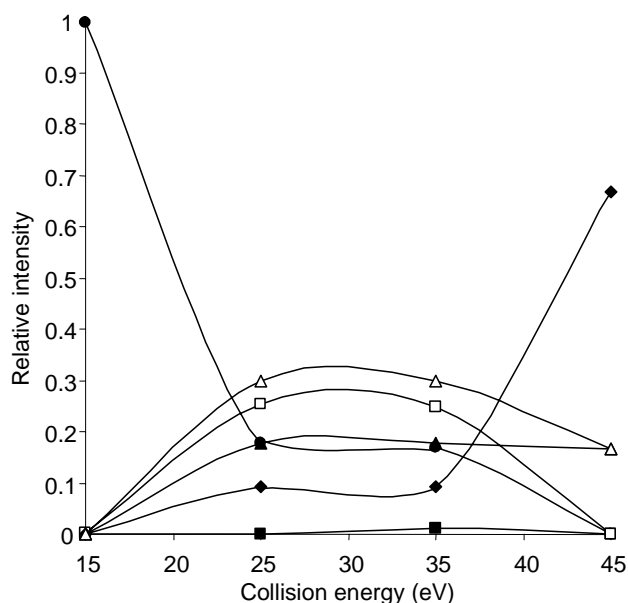


Figure 3.16 Collision energy profile for the precursor ion and product ions of: ghrelin tryptic peptide T1-11. (3+ precursor ion.) Legend: precursor ion, $[M+3H]^{3+}$ (●); neutral loss, $[M\text{-octanoyl}+3H]^{3+}$ (■); neutral loss, $[M\text{-octanoic acid}+3H]^{3+}$ (□); acylated serine immonium ion (▲); acyl carbenium ion, $\text{CH}_3(\text{CH}_2)_6\text{CO}^+$, (Δ); native/des-acyl serine immonium ion (◆).

3.2.5 Tandem Mass Spectrometry of Tyrosine(4) and Tyrosine(13) *O*-acylated Renin Substrate Tetradecapeptide

Renin substrate tetradecapeptide (renin substrate; also referred to as angiotensinogen 1-14) was initially selected for study as a peptide containing a serine residue at the C-termini. However, using the synthesis conditions described in Chapter 2.2, renin substrate was found to be preferentially acylated at either or both tyrosine residues in its sequence (tyr-4 and/or tyr-13). This novel synthetic acylation is of interest for study, as an example of an acylation not yet detected in biological samples. Ghrelin was discovered relatively recently, and is acylated at serine in the mammalian form and at threonine in the amphibian form, of the peptide with neither of these amino acids being previously identified as sites of acyl modification. ESI-Q-o-TOF MS/MS spectra were obtained for native, tyr(4) *O*-octanoyl renin substrate and tyr(13) *O*-octanoyl, myristoyl and palmitoyl renin substrate (Figure 3.17). A striking feature of the Tyr(13) *O*-acylated

renin substrate MS/MS spectra (Figure 3.17 parts b - d) is the dominance of an acylated tyrosine immonium ion (at m/z 262.18, m/z 346.26 and m/z 374.29 for the Tyr(13) *O*-octanoylated, myristoylated and palmitoylated peptides respectively). This ion is not observable at the collision energy shown for the MS/MS spectrum of the Tyr(4) *O*-octanoylated renin substrate (Figure 3.17 part e) and only forms to a very small degree at higher collision energies (see Figure 3.18 and related discussions below). The difference in the abundance of the acylated tyrosine immonium ion suggests that either proximity to the peptide end or neighbouring amino acids strongly influences the formation of these ions.

Apart from the differences in the spectra due to the presence and/or absence of the acylated tyrosine immonium ion, the MS/MS spectra for the native, Tyr(4) *O*-octanoylated and Tyr(13) *O*-acylated peptides are very similar. This is consistent with the observations of peptides acylated at serine residues as discussed in the previous sections. There is a noticeable gap in the fragmentation ion sequence of the native and *O*-acylated renin substrate peptides. This can be explained by the presence of a proline residue. Using low-energy collision-induced fragmentation it is usual to see a predominant cleavage to the N-terminal side of proline. Fragmentation at this position is induced by the presence of the cyclic structure which inhibits C-terminal cleavage due to steric factors.²³⁸

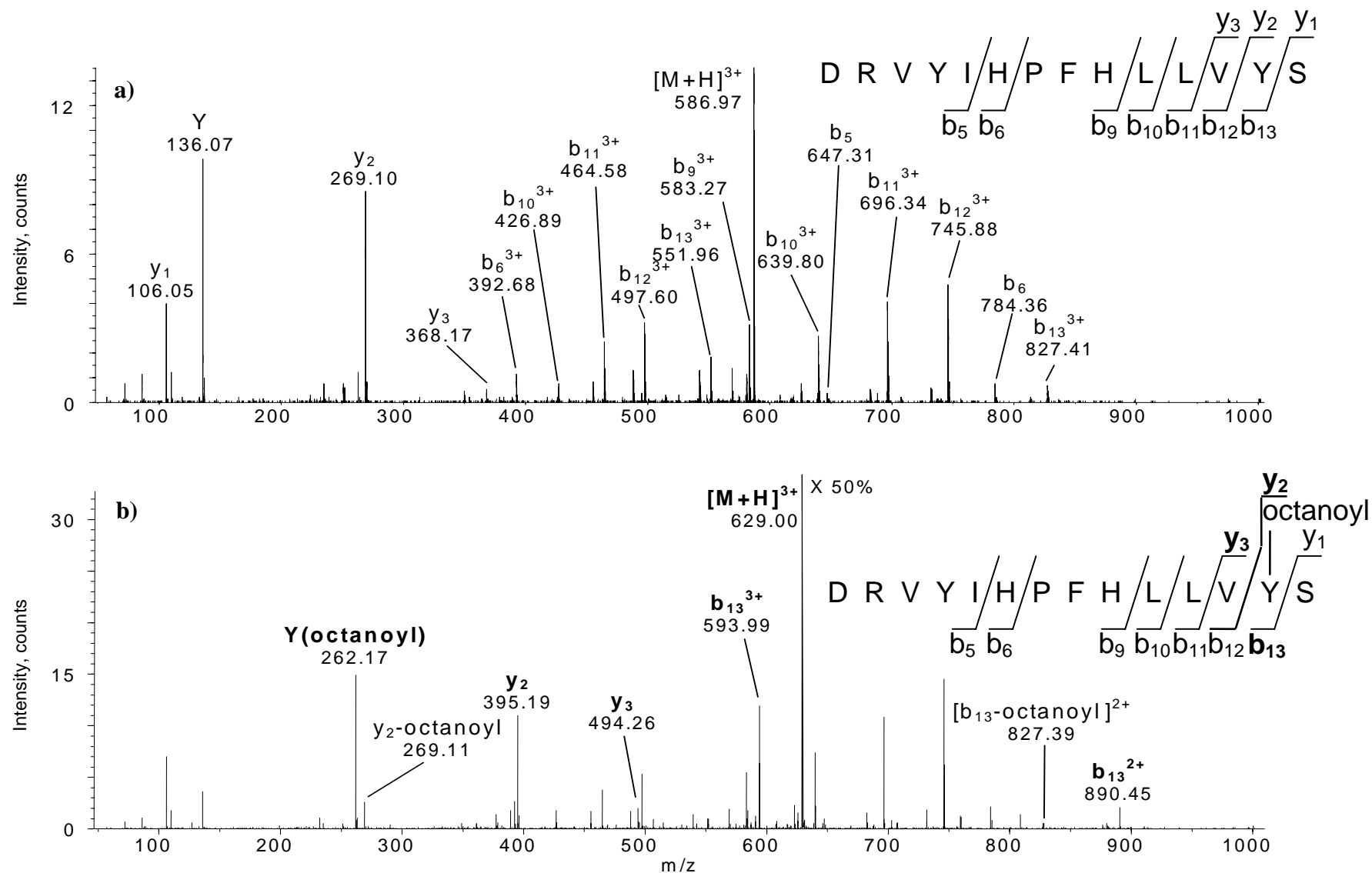


Figure 3.17 ESI-Q-o-TOF MS/MS of native and *O*-acyl-renin substrate tetradecapeptide (3⁺ charge state). For clarity only fragment ions of interest are labelled, bold-face type indicates ions containing the acyl moiety. Native peptide (C.E. = 25 eV) (a); Tyr(13) *O*-octanoyl-renin substrate tetradecapeptide (C.E. = 25 eV) (b). (Figure continued on next two pages.)

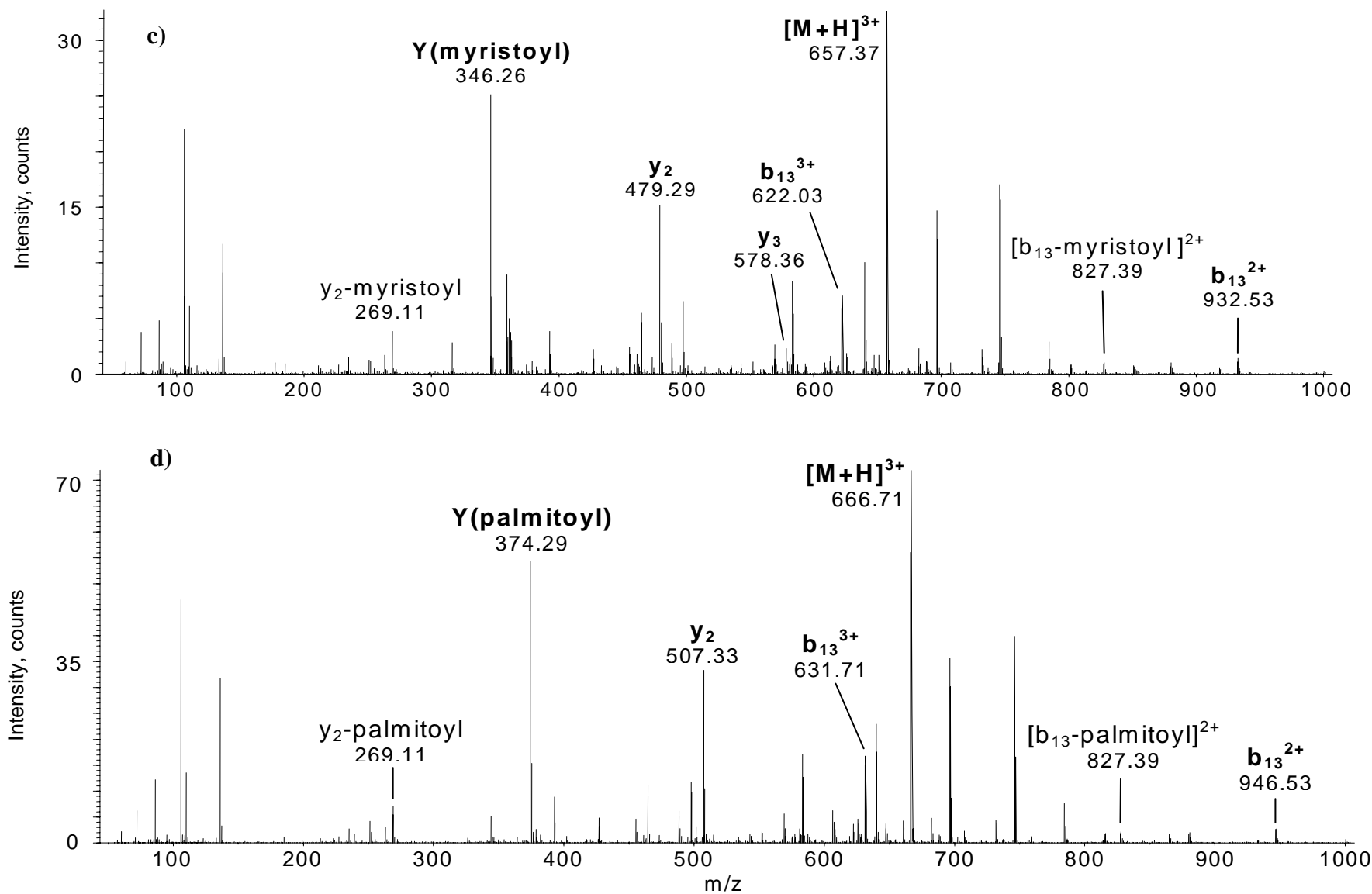


Figure 3.17 cont. Tyr(13) *O*-myristoyl-renin substrate tetradecapeptide (C.E. = 30 eV) (c), Tyr(13) *O*-palmitoyl-renin substrate tetradecapeptide (C.E. = 30 eV) (d). For clarity only ions containing the acyl moiety are labelled in (c) and (d). (Figure continued on next page.)

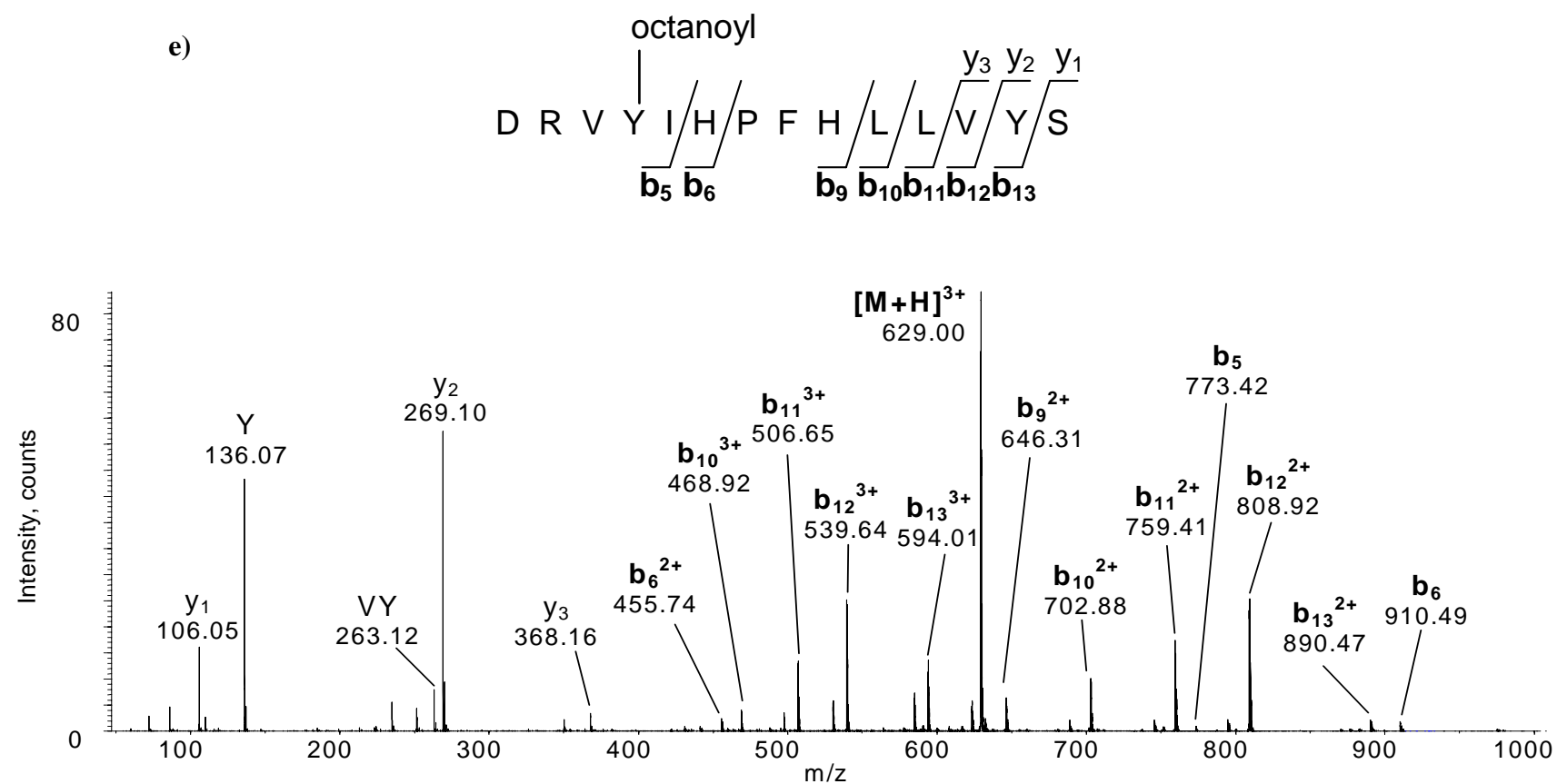


Figure 3.17 cont. Tyr(4) *O*-octanoyl-renin substrate tetradecapeptide (C.E. = 25 eV) (e).

A mass shift, attributed to the loss of the acyl moiety on Tyr(13), is observed from the y_2 and b_{13} product ions. The mass shift allows for the localisation and assignment of the acyl modification to the tyrosine (13) residue. The spectra shown in Figure 3.17 are for the triply charged precursor ions; similar spectra were observed for the fragmentation of the quadruply charged precursor ions. In contrast to the observations made for acylation at serine residues, there were no abundant ions at m/z values attributed to the neutral loss of either the acyl moiety or carboxylic acid group from the acylated precursor ions. Ions of low abundance were observed for these neutral losses from the Tyr(13) *O*-acyl renin substrate, however, no ions were observed for the neutral loss of the acyl moiety from the Tyr(4) *O*-acyl renin substrate. In the fragmentation spectrum of tyr(4) *O*-octanoyl renin substrate, the octanoylated tyrosine immonium ion and octanoyl carbenium ion were both of low abundance and only seen at higher collision energies (above a collision energy of 35 eV). This result illustrates that the position of the acyl moiety within a peptide sequence plays a crucial role in the types of ions formed in CID-MS/MS fragmentation. A summary of the marker ions observed, and the approximate optimum collision energy, is shown in Table 3.4.

Figure 3.18 shows the fragmentation profiles of a number of markers ions present in the MS/MS of the various *O*-acylated renin substrate peptides. The dominant marker ion observed in the MS/MS spectra of tyr(13) *O*-acylated renin substrate was the acylated tyrosine immonium ion, as shown in the spectra in Figure 3.17. Marker ions, resulting from the neutral loss of either the acyl component or acyl carboxylic acid and an acyl carbenium ion, were also present in the spectra of the fragmented tyr(13) *O*-acyl renin substrate peptides. These ions, however, were at significantly lower abundance compared with the acylated tyrosine immonium ions (shown in Figure 3.18 part (ii)).

Table 3.4 Characteristic and useful sequence ions identified for *O*-acylated renin substrate tetradecapeptide.

Peptide	Characteristic and useful sequence ions	Optimum collision energy ^a (eV)
tyr(13) <i>O</i> -octanoyl renin substrate DRVYIHPFHLLVY(oco)S	<i>ESI-Q-o-TOF</i> : neutral loss of octanoyl (-126 u) ^b neutral loss of octanoic acid (-144 u) ^b immonium ion <i>O</i> -octanoyl-Y (m/z 262) octanoyl carbenium ion (m/z 127) ^b	30 30 35 35
tyr(13) <i>O</i> -myristoyl renin substrate DRVYIHPFHLLVY(myrist)S	<i>ESI-Q-o-TOF</i> : neutral loss of myristoyl (-210 u) ^b neutral loss of myristic acid (-228 u) ^b immonium ion <i>O</i> -myristoyl-Y (m/z 346) myristoyl carbenium ion (m/z 211) ^b	30 30 35 30-35
tyr(13) <i>O</i> -palmitoyl renin substrate DRVYIHPFHLLVY(pam)S	<i>ESI-Q-o-TOF</i> : neutral loss of palmitoyl (-238 u) ^b neutral loss of palmitic acid (-256 u) ^b immonium ion <i>O</i> -palmitoyl-Y (m/z 374) palmitoyl carbenium ion (m/z 239) ^b	30 30 35 30-35
tyr(4) <i>O</i> -octanoyl renin substrate DRVY(oco)IHPFHLLVYS	<i>ESI-Q-o-TOF</i> : immonium ion <i>O</i> -octanoyl-Y (m/z 262) ^b octanoyl carbenium ion (m/z 127) ^b	>45 45+

^aApproximate optimum collision energy as determined by ESI-Q-o-TOF MS/MS, and estimated from plots in Figure 3.18.^bLow abundance ion.

Amongst the lower abundance marker ions, an acyl carbenium ion was preferentially formed for the octanoylated peptide (denoted by an open triangle). In the neutral loss species, the loss of the carboxylic acid (denoted by squares), were favoured in the fragmentation of the myristoylated and palmitoylated peptides. A similar trend was observed for the formation of acyl carbenium ions as a result of the fragmentation of *O*-acylated neurokinin A (see Figure 3.5, above); where the octanoyl carbenium ion was relatively more abundant than the myristoyl carbenium ion in the ESI-Q-o-TOF MS/MS spectra of their respective peptides.

A comparison of the collision energies required to fragment native and Tyr(13) *O*-acylated renin substrate precursor ions is shown in Figure 3.19. To obtain this graph the precursor ion area was determined at collision energies between 15 and 35 eV and normalised to the area at 15 eV (at which little fragmentation of the precursor ion was observed; as calculated by the Analyst QST[™] software). These values were plotted against collision energy (eV). Both the 3+ and 4+ precursor ions were analysed in this way. The resulting data in Figure 3.19 indicates that a higher collision energy is required to fragment the Tyr(13) *O*-acylated renin substrate peptide than to fragment the native peptide, and that the collision energy required to fragment the acylated peptide increases with increasing acyl chain length.

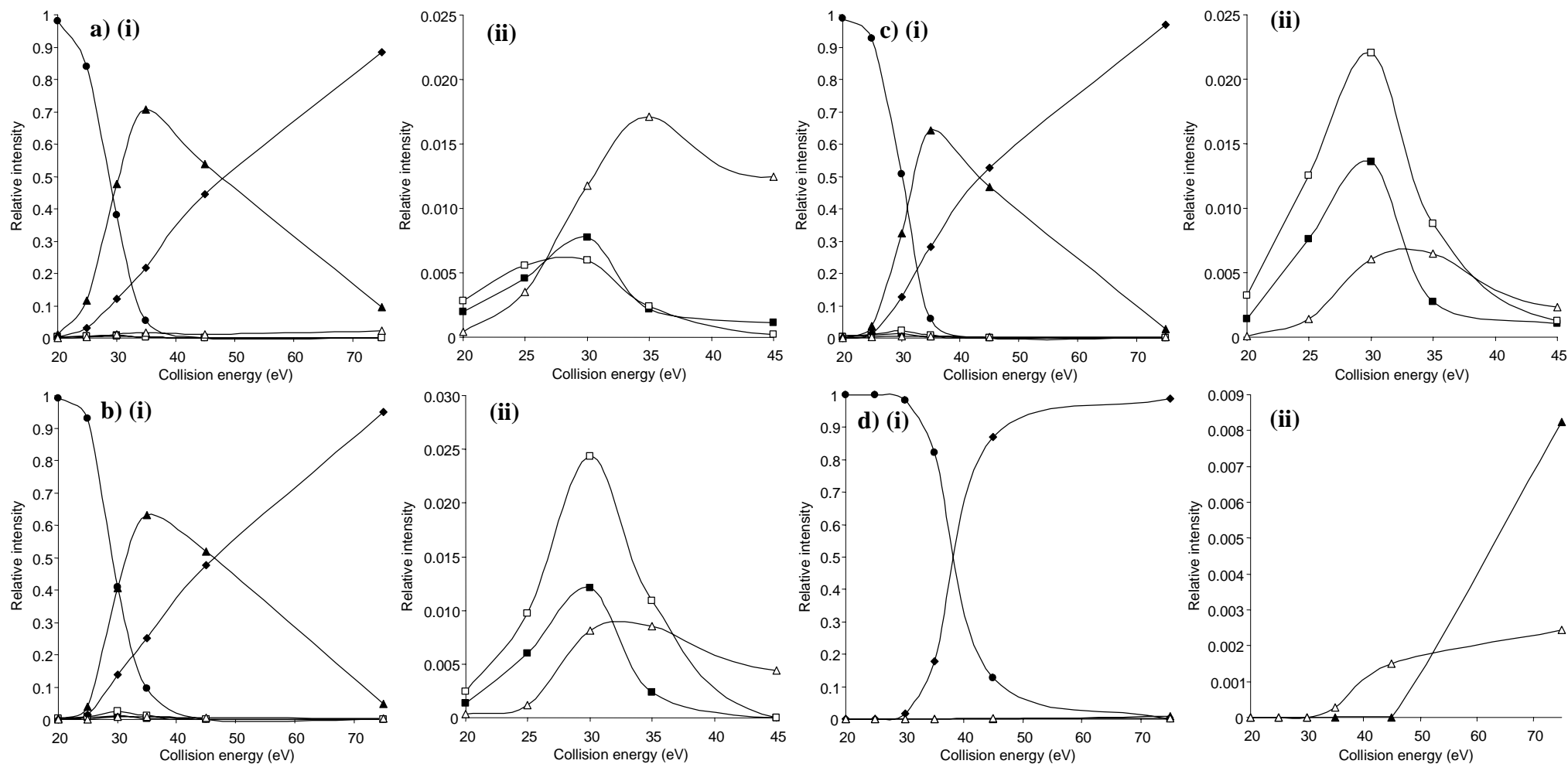


Figure 3.18 Collision energy profiles for the precursor ion and product ions of: tyr(13) *O*-octanoyl renin substrate (a); tyr(13) *O*-myristoyl renin substrate (b); tyr(13) *O*-palmitoyl renin substrate (c); tyr(4) *O*-octanoyl renin substrate (d) (3+ precursor ions). Part (i) graphs show all ions; part (ii) graphs are an expansion showing the low abundance ions. Legend: precursor ion, $[M+3H]^{3+}$, (●); neutral loss, $[M-acyl+3H]^{3+}$, (■); neutral loss, $[M-acyl\ carboxylic\ acid+3H]^{3+}$, (□); acylated tyrosine immonium ion, (▲); acyl carbenium ion, $CH_3(CH_2)_{(n-2)}CO^+$, (△); des-acyl tyrosine immonium ion, (◆).

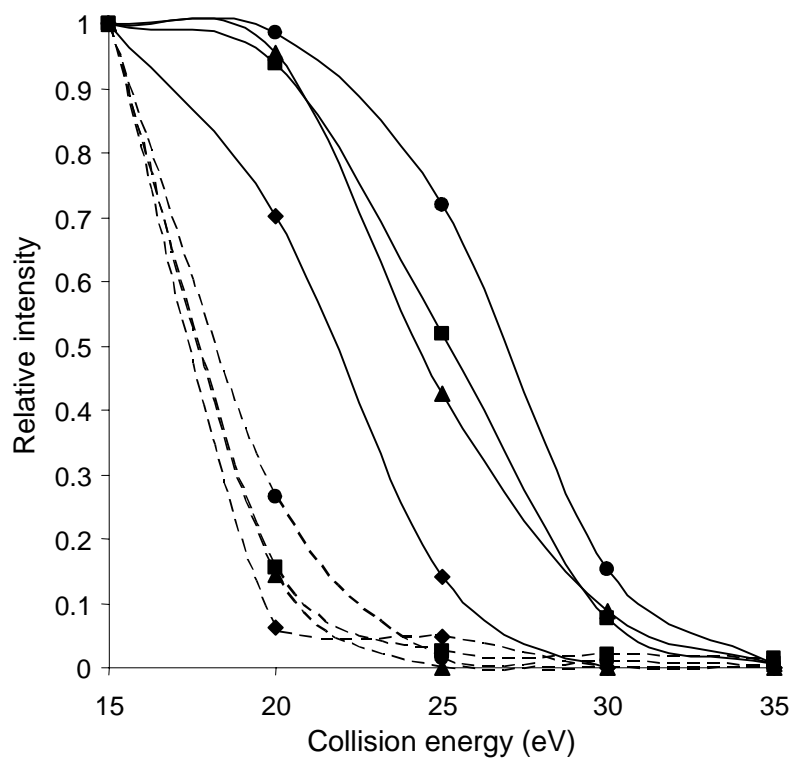


Figure 3.19 Collision energy profile for the fragmentation of native and Tyr(13) *O*-acylated renin substrate tetradecapeptide (3+ (solid line) and 4+ (dashed line) charge states). Legend: native peptide(♦); octanoylated peptide (▲); myristoylated peptide (■); and palmitoylated peptide (●).

3.2.6 Tandem Mass Spectrometry of *S*-acylated Glutathione

Glutathione was selected as a peptide suitable for *S*-acylation as it contains one cysteine residue and no other residues susceptible to acylation using the acylation procedure used in this work (Chapter 2.2). *S*-octanoyl, myristoyl and palmitoyl glutathione were analysed by ESI-Q-o-TOF MS/MS (Figure 3.20). As observed with the other peptides studied in this chapter, the MS/MS spectra of the acylated glutathione peptides are similar to the MS/MS spectra of the native peptide.

Marker ions associated with the presence of the acyl moiety on the peptide included an acylated cysteine immonium ion and an acylated cysteine acylium ion (b_2y_2 ion). The presence of acylated cysteine immonium ions is similar to that observed for peptides containing acylated serine and tyrosine residues. Acylated acylium ions (single amino acid internal cleavage ions) were not observed for the other peptides studied. The presence of the acylated acylium ion is possibly due to the stabilising influence of the acyl moiety on the cysteine residue in this short (3-mer) peptide. There was no observable acylium ion present in the MS/MS spectrum of native glutathione, and acylated cysteine acylium ions have not been reported in the literature. An additional ion found in all three *S*-acylated glutathione MS/MS spectra but not present in the spectrum of native glutathione, is the y_2-H_2O ion. The presence of this ion may result from the stabilising influence of the acyl moiety. Alternatively, the oxygen atom on the acyl group may be eliminated via a reaction with the N-terminal γ -glutamic acid, in which case the fragmentation would be a sequence specific reaction, rather than a general marker for acylation. There is evidence for the loss of the acyl moiety as a neutral forming protonated glutathione (m/z 308). There was no evidence for an ion formed as a result of a neutral loss of an acyl thioacid (m/z 274). In contrast, Hoffman

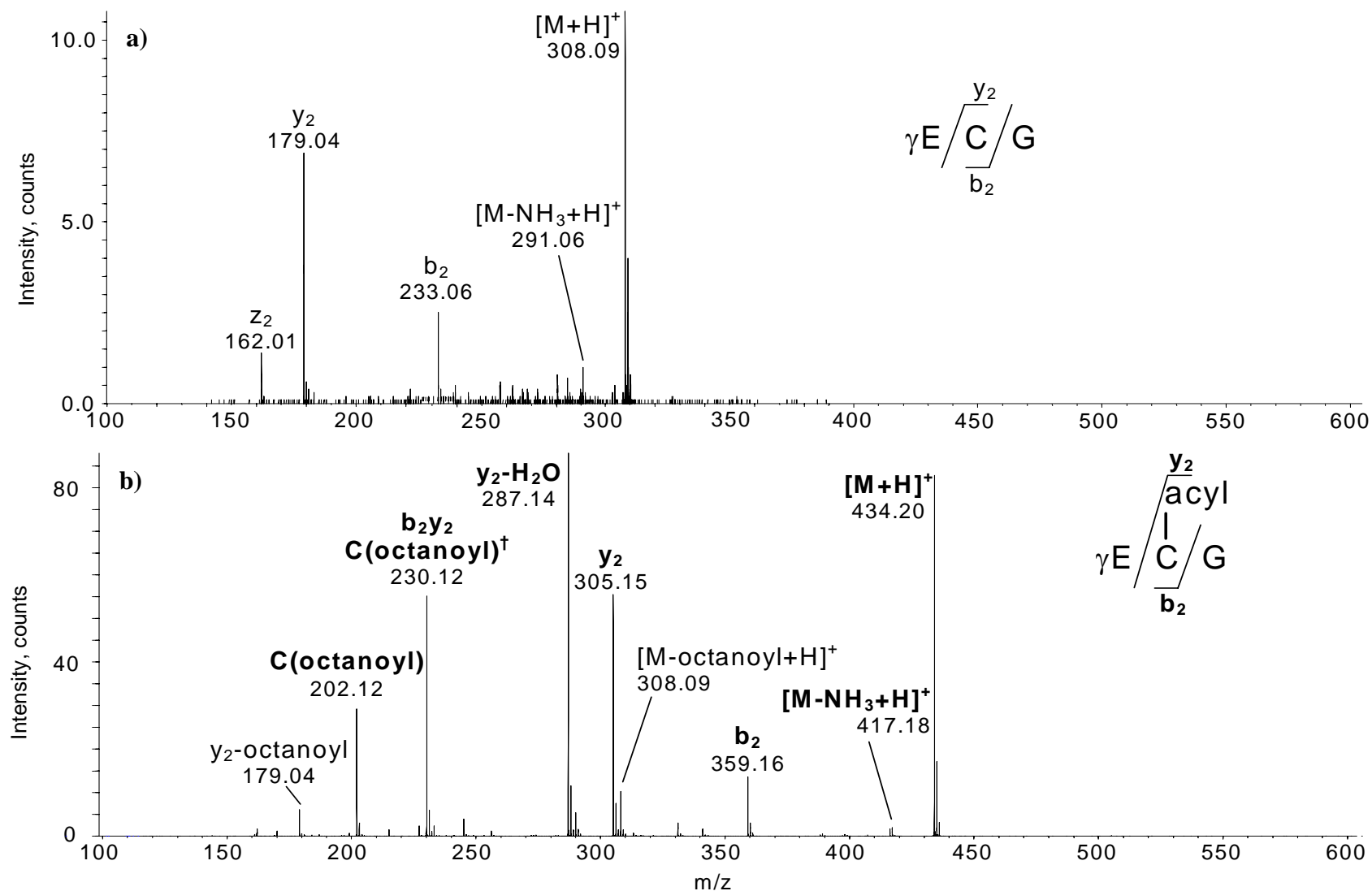


Figure 3.20 ESI-Q-o-TOF MS/MS of native and *S*-acyl-glutathione (1+ charge state). For clarity only fragment ions of interest are labelled, Bold face type indicates ions containing the acyl moiety. Native peptide (C.E. = 15 eV) (a); *S*-octanoyl glutathione (C.E. = 20 eV) (b). † indicates acylated cysteine acylium ion. (Figure continued on next page.)

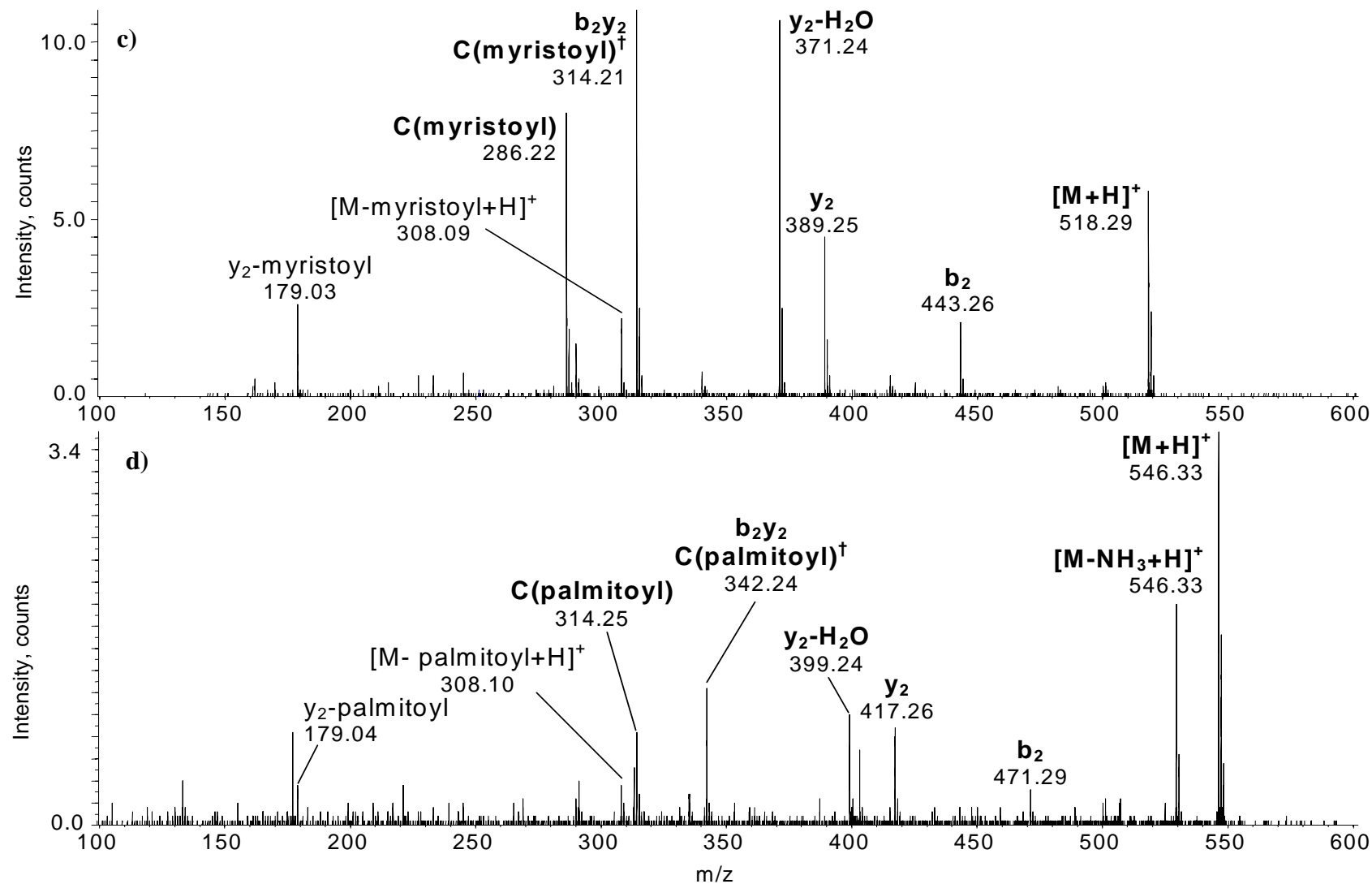


Figure 3.20 cont. ESI-Q-o-TOF MS/MS of native and *S*-acyl-glutathione (1+ charge state). For clarity only fragment ions of interest are labelled, Bold face type indicates ions containing the acyl moiety. *S*-myristoyl glutathione (C.E. = 25 eV) (c), *S*-palmitoyl glutathione (C.E. = 25 eV) (d). † indicates acylated cysteine acylium ion.

Table 3.5 Characteristic and useful sequence ions identified for *S*-acylated glutathione.

Peptide	Characteristic and useful sequence ions	Optimum collision energy ^a (eV)
<i>S</i> -octanoyl glutathione γ -EC(oco)G	<i>ESI-Q-o-TOF</i> : neutral loss of octanoyl (-126 u) octinoylated cysteine acylium ion (m/z 230.1) immonium ion <i>O</i> -octanoyl-C (m/z 202.1)	25 25 35+
<i>S</i> -myristoyl glutathione γ -EC(myrist)G	<i>ESI-Q-o-TOF</i> : neutral loss of myristoyl (-210 u) myristoylated cysteine acylium ion (m/z 314.2) immonium ion <i>O</i> -myristoyl-C (m/z 286.2)	25+ 25-30 35
<i>S</i> -palmitoyl glutathione γ -EC(pam)G	<i>ESI-Q-o-TOF</i> : neutral loss of palmitoyl (-238 u) palmitoylated cysteine acylium ion (m/z 342.24) immonium ion <i>O</i> -palmitoyl-C (m/z 374) palmitoyl carbenium ion (m/z 239)	25-35 30 30 35

^a Approximate optimum collision energy as determined by ESI-Q-o-TOF MS/MS, as estimated from plots in Figure 3.21.

et. al. saw a neutral loss of a thioacid in the MALDI of a palmitoylated peptide in their research.²³⁴ This discrepancy may be a result of the different peptide sequences studied here and in the work of Hoffman *et. al.* or as a result of the different ionisation method used. A small ion attributed to the palmitoyl carbenium ion was observed in the ESI-Q-o-TOF MS/MS spectrum of *S*-palmitoyl glutathione, however no acyl carbenium ions were observed in the ESI-Q-o-TOF MS/MS spectra of either the octanoyl or myristoyl modified peptides. A summary of the useful marker ions observed in the ESI-Q-o-TOF MS/MS spectra of acylated glutathione peptides is given in Table 3.5.

The fragmentation of *S*-acylated glutathione was studied over a range of collision energies. The results are shown in Figure 3.21. The relative abundance of the acylated cysteine acylium ion increased steadily until a collision energy of 25-30 eV, after which it steadily declined. It is likely that at high collision energies this ion fragments to form an acylated immonium ion, as there was no evidence for the formation of an unacylated cysteine acylium ion. A palmitoylated carbenium ion is present in the ESI-Q-o-TOF MS/MS spectrum of *S*-palmitoylated glutathione, whilst equivalent ions are not observed in the spectra of *S*-octanoylated or *S*-myristoylated glutathione. This may be the result of the increased stability of the palmitoyl moiety as a leaving group due to its increased carbon-chain length. The *S*-palmitoylated glutathione spectrum also shows an increased proportion of a des-acyl cysteine immonium ion relative to the *S*-octanoylated and *S*-myristoylated glutathione peptides.

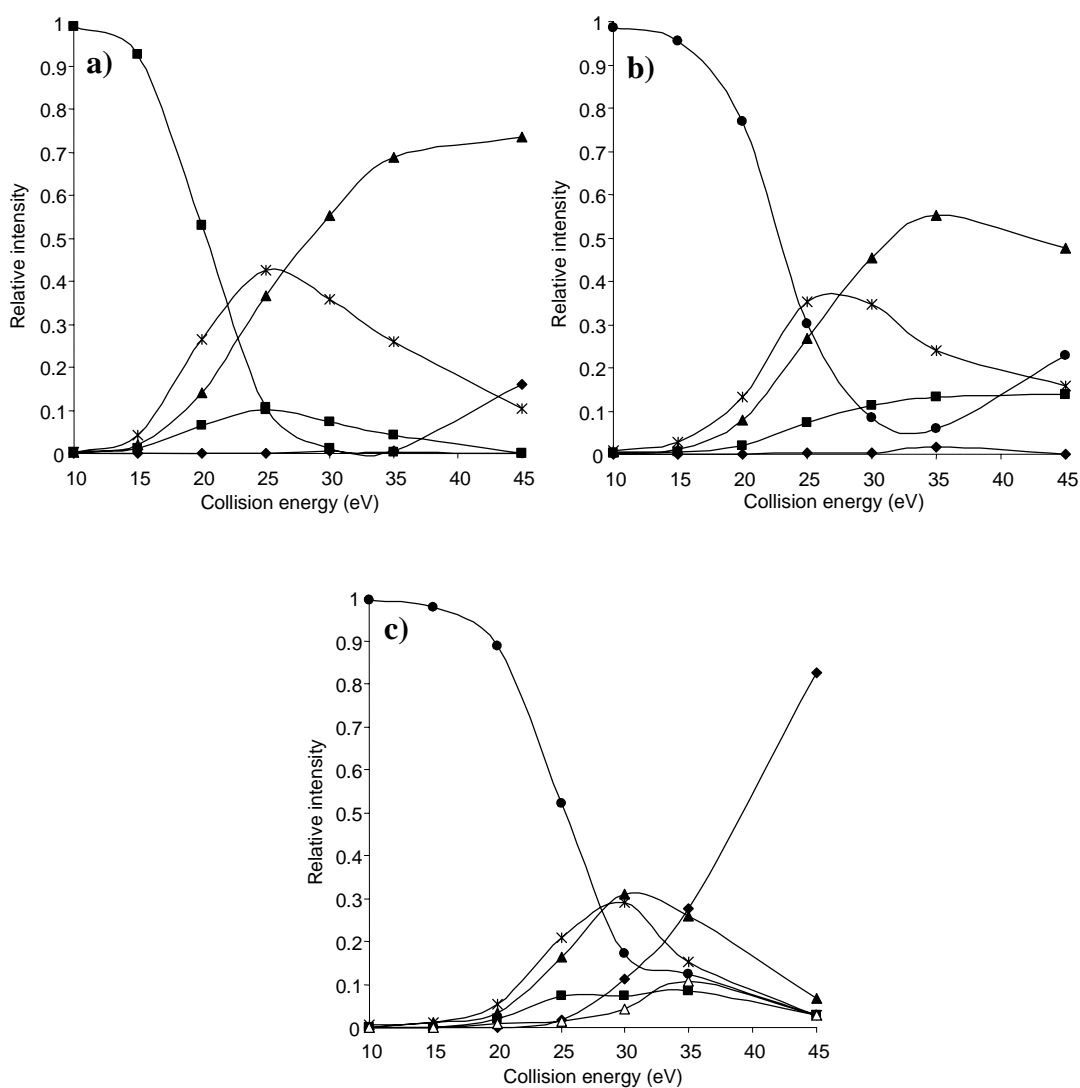


Figure 3.21 Collision energy profiles for the precursor ion and product ions of: *S*-octanoyl glutathione (a); *S*-myristoyl glutathione (b); *S*-palmitoyl glutathione (c) ($1+$ precursor ions). Legend: precursor ion, $[M+2H]^{2+}$, (●); neutral loss, $[M-acyl+2H]^{2+}$, (■); acylated cysteine acylium ion (*); acylated cysteine immonium ion, (▲); acyl carbenium ion, $CH_3(CH_2)_{(n-2)}CO^+$, (Δ); des-acyl cysteine immonium ion, (◆).

A graph showing the effect of collision energy on the fragmentation of the acylated glutathione precursor ions is shown in Figure 3.22. As was found for the *O*-acylated renin substrate peptide (section 3.2.5 above), increasing collision energies were required to fragment *S*-acylated glutathione containing longer acyl chain moieties.

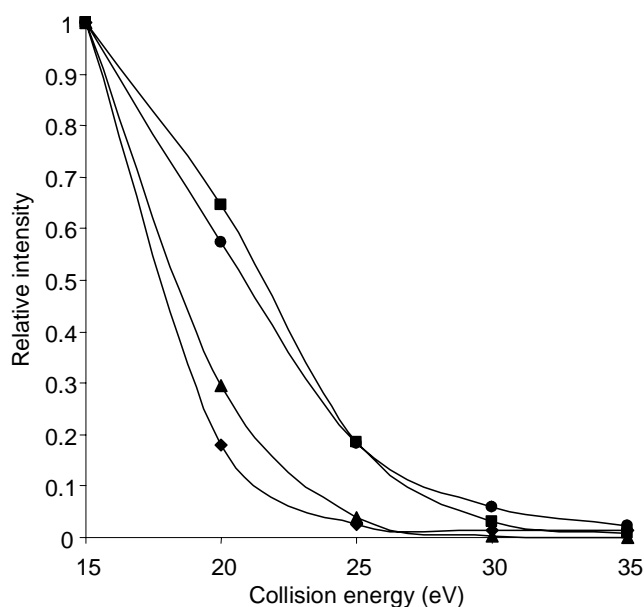


Figure 3.22 Collision energy profile for the fragmentation of native and *S*-acylated glutathione (2+ charge state). Legend: native peptide(♦); octanoylated peptide (▲); myristoylated peptide (■); and palmitoylated peptide (●).

3.2.7 Tandem Mass Spectrometry of *N*-acylated Substance P

Substance P was chosen as a peptide suitable for *N*-acylation, due to the lack of amino acid residue side chain sites susceptible to *O*- or *S*- acylation, using the acylation procedure employed (see Chapter 2.2). Native and *N*-octanoylated substance P were analysed by MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS. The resultant spectra are shown in Figure 3.23 and Figure 3.24. The MS/MS spectra produced for the native and octanoylated substance P, using both MS techniques, are relatively similar. The spectra are dominated by N-terminal sequence fragments, which can be explained by the presence of the basic residue, arginine at the N-terminus of the peptide.

An important difference between the spectra of the native and octanoylated substance P is the presence of a b_1 ion in the spectra produced by fragmentation of the octanoylated peptide (m/z 283.2). This is absent in the spectra of the native peptide. The MS/MS fragmentation of peptides does not usually result in the formation of b_1 ions due to the absence of a carbonyl group at the N-terminal side of an unmodified N-terminal amino acid.²³⁸ In the case of an N-terminally acylated peptide, a b_1 ion is able to form via a five-membered ring attack.²³⁹ This b_1 formation has been observed by others, including Schlosser and Lehmann, who analysed a short N-myristoylated peptide,²³⁸ and Hoffman and co-workers, who also observed an a_1 ion.²³⁴ Jedrzejewski and Lehmann observed a_1 and b_1 ions when analysing similarly N-terminally acylated peptides.²⁴⁰ In the present study, there was no a_1 ion detected in the MALDI or ESI-Q-o-TOF MS/MS of octanoylated substance P.

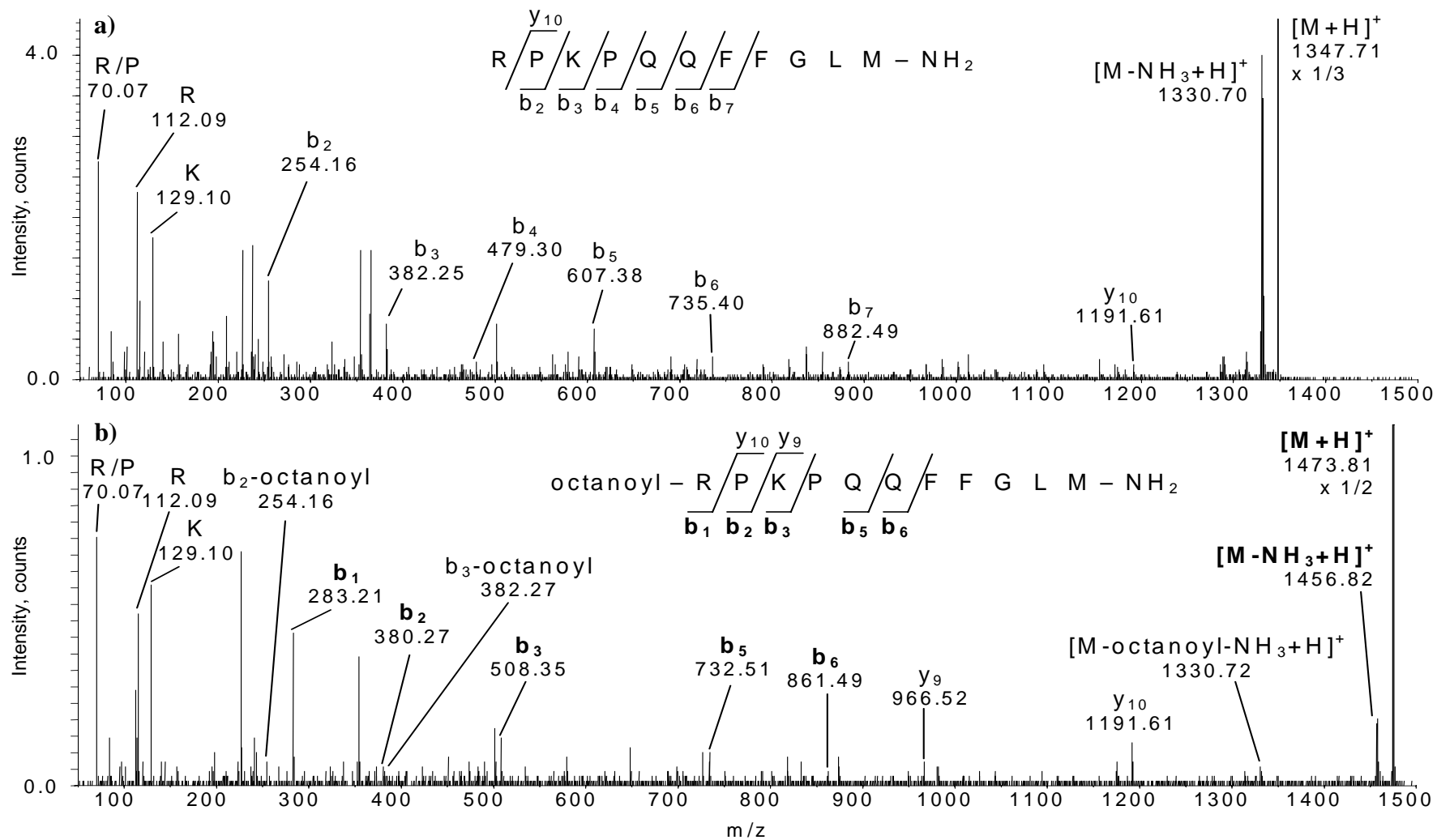


Figure 3.23 MALDI-Q-o-TOF MS/MS of substance P (C.E.110 eV) (a) and *N*-octanoylated substance P (C.E. 110 eV) (b). Bold-face type indicates ions containing the acyl moiety.

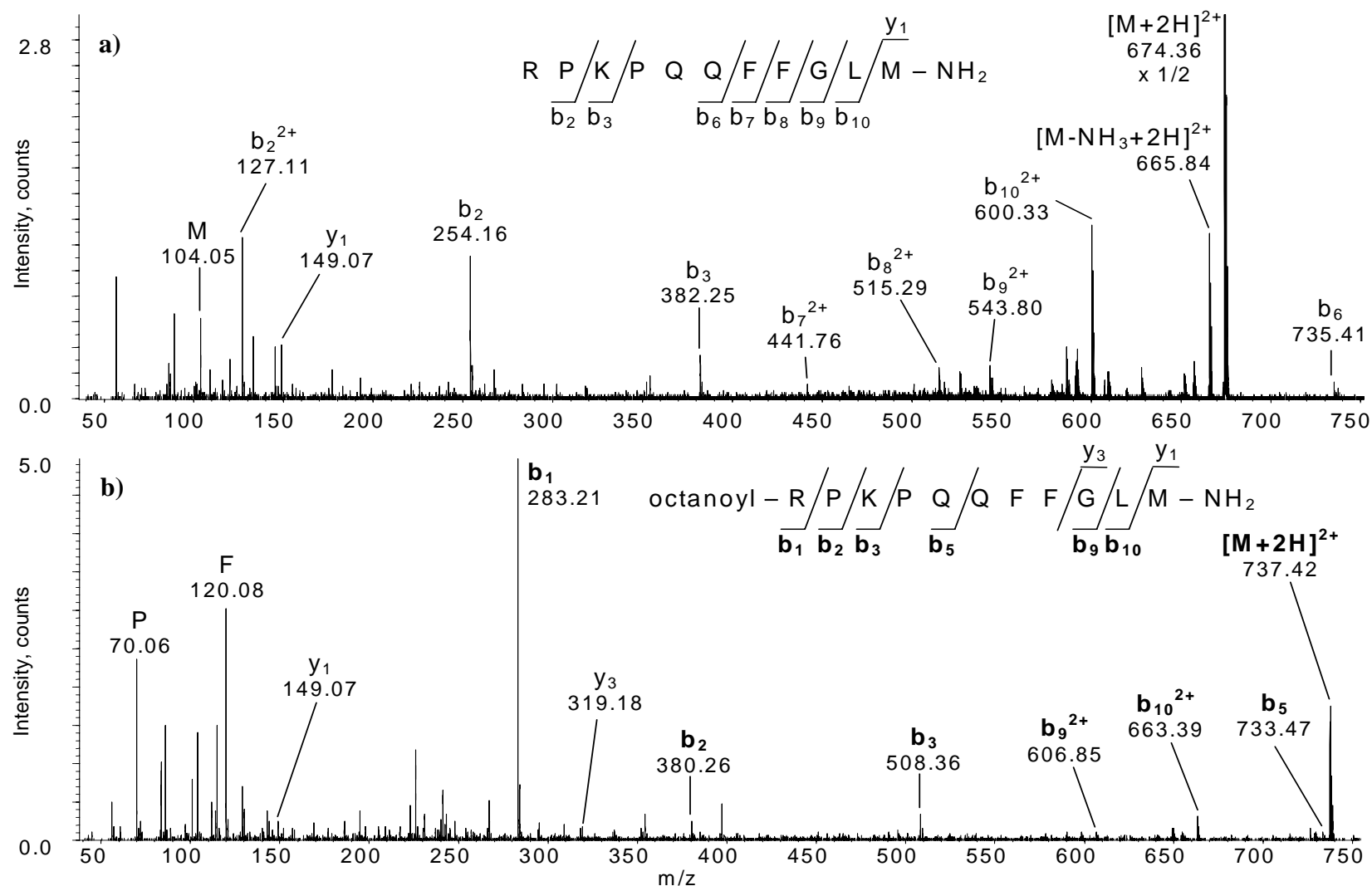


Figure 3.24 ESI-Q-o-TOF CID MS/MS of substance P (C.E.25 eV) (a) and *N*-octanoylated substance P (C.E. 25 eV) (b). Bold-face type indicates ions containing the acyl moiety. (No significant ions were observed above m/z 750.)

No octanoyl product ion was observed in the fragmentation of *N*-octanoylated substance P. However, as was noted for acylated glutathione, a palmitoyl ion observed although neither the octanoyl nor myristoyl ions were seen. This suggests that if a longer chain acyl moiety were in place of the octanoyl, an acyl fragment ion may be observed. Indeed, Jedrzejewski and Lehmann observed myristoyl ions when analysing a *N*-myristoylated protein.²⁴⁰ As noted for other acylated peptides above, the presence of particular product ions (for example, ions resulting from neutral losses and protonated acyl ions) may be sequence dependant. In the case of *N*-terminal acylation, the identity of the *N*-terminal amino acid would also be critical in determining the likelihood of various ions forming in the CID or PSD process. Additionally, in contrast to the majority of peptides studied in this thesis, there was no ion observed resulting from the neutral loss of the octanoyl moiety (with a calculated m/z value of 674.4). The lack of marker ions may be attributed to the presence of a proline residue immediately adjacent to the *C*-terminal side of the acylated arginine residue, which may stabilise the acyl moiety, as observed in the case of synthetically *O*-acylated eledoisin (section 3.2.2 above).

3.3 Conclusions

This work shows that a variety of synthetically acylated peptides, and the naturally occurring acylated peptide ghrelin, are amenable to analysis by tandem mass spectrometry (including the pseudo tandem technique of MALDI PSD MS). The ionisation techniques of MALDI and ESI were found to be complementary. In general, MALDI produced a spectrum which was more easily interpreted and generally gave better sequence coverage when compared with ESI. This was true for both PSD and CID produced fragment ions. However ESI, when coupled with CID, allowed for the

production of a greater range of acylation marker ions. This was due to the ability to adjust the collision energy allowing for greater control of fragmentation when using ESI-Q-o-TOF MS/MS. ESI has the additional advantage of allowing analysis of peptides of a higher mass range, due to the formation of multiply-charge ions. This was demonstrated in the analysis of ghrelin, with a monoisotopic mass of 3369.9 u, which was analysed using ESI with a charge state of 6+ (m/z 559.5). This peptide was found to be difficult to analyse by MALDI PSD MS, with only a small amount of fragmentation being observed. When analysing acylated peptides by ESI-Q-o-TOF MS/MS it was found that a higher collision energy was required to fragment acylated peptides than their native, or des-acyl, equivalents. In addition, increasing collision energies were required with increasing acyl chain length in order to obtain the same degree of fragmentation.

Acyl modifications were found to be relatively stable to the fragmentation processes involved in mass spectrometry, with the MS/MS spectra of acylated peptides being very similar in appearance to that of the native, or des-acyl, peptide. This allows for the site of acylation to be determined directly by mass difference or by comparison with the native or des-acyl form of the peptide. Although relatively stable to the tandem MS process, the acyl modifications were generally found to be sufficiently labile to allow the formation of a variety of useful marker ions in the fragmentation spectra of acylated peptides. The marker ions observed in the ESI-Q-o-TOF MS/MS spectra of acylated peptides are described in Table 3.6 below.

Table 3.6 Marker ions observed in the ESI-Q-o-TOF MS/MS spectra of acylated peptides.

Marker	Ion observed in MS/MS spectrum ^a	Indicates
Neutral loss of acyl	$[M-\text{acyl}+nH]^{n+}$	acyl chain length and presence of acyl moiety in peptide
Neutral loss of carboxylic acid	$[M-(\text{acyl} + H_2O)+nH]^{n+}$	acyl chain length and ester bond to amino acid residue
acyl carbenium ion	$CH_3(CH_2)_{(n-2)}CO^+$	acyl chain length
acylated immonium ion	$[I+\text{acyl}]^+$	acyl chain length and identity of modified amino acid
b ₁ ion	b ₁	N-terminally modified peptide

^a acyl indicates delta mass of the acyl moiety; I indicates mass of immonium ion.

Three of the marker ions observed give a direct indication of the identity of the acylated amino acid. These are the ions resulting from a neutral loss of an acylated immonium ion and a b₁ ion in the case of an N-terminally acylated peptide. Acylated immonium ions, where present, identify an acylated amino acid. The identity of which can be determined by the addition of the masses of the amino acid immonium ion mass and the delta mass of the acyl moiety (if known or predicted). One of the most useful features of the MS/MS spectrum resulting from the fragmentation of an N-terminally acylated peptide, is the presence of a b₁ ion not normally observed in the spectrum of peptide, not otherwise derivatised at its N-terminus.

Ions resulting from a neutral acyl loss or the presence of an acyl carbenium ion both give an indication that an acylation is present on the peptide being analysed. If there are sufficient sequence ions present in the MS/MS spectrum, the location of the acyl moiety on the peptide can be determined by *de novo* sequencing. The neutral loss of a carboxylic acid indicates that the acyl moiety was attached to a hydroxyl group, which

could originate from a number of amino acids, such as serine, tyrosine or threonine. If more than one marker ion is observed (for example, an acyl carbenium ion and an acylated immonium ion) then the identity of the acyl moiety and the modified amino acid can be given with greater certainty.

The marker ions observed for a particular acylated peptide were found to be dependent on both the peptide sequence and the chain-length of the acyl moiety. Not all peptides studied produced the complete set of marker ions as described in Table 3.6. A greater number of marker ions were observed for acylated peptides containing longer-chained acyl moieties compared to shorter-chained acyl moieties. No significant influence of peptide length was observed on the production of marker ions in the comparison of the fragmentation of ghrelin and the ghrelin tryptic peptide T1-11. Tandem mass spectrometry CID fragmentation of one of the acylated peptides studied, eledoisin, failed to produce any marker ions. It is postulated that this is as a result of the proline residue situated to the immediate N-terminal side of the acylated serine residue in this synthetically acylated peptide. Tandem mass spectrometry of the synthetically N-acylated *N*-octanoyl substance P resulted in the formation of only a b_1 marker ion, without the formation of other acylated ions. This result could also be attributed to the presence of a proline to the immediate C-terminal side of the N-acylated arginine residue.

It was found that formation of the various marker ions depends on the collision energy used. The collision energy required to produce useful marker ions was generally found to be similar to or slightly higher, than that required to produce ions suitable for *de novo* sequencing. Consequently it may be necessary, when analysing acylated peptides, to acquire MS/MS spectra at two different collision energies; one suitable for *de novo*

sequencing and another at a slightly elevated collision energy to aid in confirming the identity of the modified amino acid.

In order to analyse peptides of mass greater than *ca.* 2000 u, longer peptides can be enzymatically digested to produce shorter peptides more amenable to analysis by mass spectrometry. This was demonstrated when analysing ghrelin (monoisotopic mass 3369.9 u), which was digested with trypsin, producing an acylated peptide with a monoisotopic mass 1370.71. This shorter peptide produced easily interpreted MALDI PSD MS and MADLI Q-o-TOF MS/MS spectra.

This chapter of work demonstrates that a number of useful ions may be produced when analysing acylated peptides by tandem (or pseudo tandem) mass spectrometry. The ions produced were found to be influenced by the amino acid sequence in the locality of the acyl moiety, the identity of the acyl moiety and the mass spectrometry conditions used. This understanding of the types of acyl marker ions produced will be applied in the characterisation of novel acylated peptides and proteins in the chapters that follow.

4 Sequence Determination for a Potential HIV Immune Response Antagonist

4.1 Introduction

4.1.1 Peptide and Lipopeptide Vaccines

The technique of vaccination for viral disease prevention has been utilised for over 200 years.²⁴¹ Conventional viral vaccines contain either attenuated or inactivated varieties and, more recently, whole proteins, whole DNA gene and live recombinant vectors.^{242,}
²⁴³ At a cellular level the immune response involves the presentation of an antigen to T cells by antigen binding protein found on the surface of antigen presenting cells.^{244, 245} With the current understanding of molecular biology, it is now possible to define sites on viral particles that are responsible for eliciting an immune response and to use synthesised peptides representing these sites (epitopes) as potential vaccines. The peptide-based approach to vaccines gives the ability to direct the immune response against highly conserved epitopes that might be crucial for the pathogen, or to avoid epitopes that could induce undesirable immune responses.^{242, 243} This new targeted approach not only overcomes some of the safety and efficacy concerns of conventional vaccines, but also allows for the development of vaccines for diseases that have proved elusive using the conventional vaccine approach.²⁴¹

Peptide vaccines are also under development to treat non-communicable diseases, such as cancer and allergenic responses.²⁴⁶⁻²⁴⁸ Peptide vaccines are being trialled in the treatment of melanoma patients, among other cancer diseases, with some success in reducing the incidence of repeat metastases.²⁴⁹ In the treatment of allergies, peptide vaccines have the potential to inhibit T-cell function without inducing potentially life-threatening anaphylaxis.²⁴⁸

Peptide molecules in vaccines often require the use of potentially toxic adjuvants to provide effective immunogenicity.⁵⁷ (Adjuvants are agents which aid in stimulating the immune system, thereby making a vaccine more effective, without having any direct antigenic effect.²⁵⁰) Lipopeptides, however, have been found to illicit comprehensive immune responses, without the need for an adjuvant,⁵⁷ and have been found to have enhanced immune responses over equivalent unacylated peptides^{242, 251} In some cases lipopeptides have induced an immune response where the corresponding peptide without a lipidic moiety did not.²⁵² This observation could be partly explained by the lipid moiety allowing the endocytosis of lipopeptides into dendritic cells.^{243, 253} Lipopeptides were first trialled as potential vaccines in the 1980s, when Hopp and co-workers found a significant improvement in the antibody response to hepatitis, where viral epitopes were conjugated to a dipalmitoyl-lysine moiety.²⁵¹ Lipopeptide vaccines have been trialled in the therapeutic treatment of Hepatitis B virus, one of the most prevalent viral pathogens of humans with around 350 million chronically infected patients,²⁵⁴ as AIDS vaccines in healthy individuals, and as a therapeutic vaccine in chronically HIV-infected patients.²⁵⁵⁻²⁵⁷

Peptide and lipopeptide vaccines offer many advantages over conventional vaccines. In addition to purity and specificity, peptide vaccines can be modified or replaced by pseudopeptides that contain changes in the amide bond, resulting in more stable and immunogenic molecules.²⁵⁸ Lipopeptide vaccines, in particular, are amenable to non-invasive application methods, such as through the skin or mucosal membranes, which allows for the efficient uptake of the antigen by lymphoid tissues.^{57, 258} Non-invasive application methods have economic and health benefits, particularly for widespread inoculation required in developing nations. The underlying cellular mechanisms by which lipopeptides elicit immune responses are poorly understood and the exact

processing and presentation pathways involved in the binding of antigens by CD proteins requires further examination.^{57, 176}

4.1.2 Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS)

4.1.2.1 History and Mechanism

The current acquired immune deficiency syndrome (AIDS) pandemic was first identified in 1981, when an unusually high number of requests for the drug to treat a particular strain of pneumonia (*Pneumocystis carinii*) was noticed by a technician at the Centers for Disease Control in the United States. The requests came from both California and New York.²⁵⁹ Around the same time, an increase in the occurrence of Kaposi's sarcoma amongst young gay men was seen in New York.^{260, 261} By July of 1982, cases of AIDS had been reported in injecting drug users, Haitians and haemophiliacs.²⁵⁹ By the end of 1982, AIDS had been reported in children and a similar disease had been identified in Africa.^{259, 262, 263}

The human immunodeficiency virus (HIV), which leads to AIDS, was first isolated by researchers in France in 1983.²⁶⁴ There have been two major strains of HIV found in humans: HIV-1 and HIV-2. HIV-1 appears to have been passed from the chimpanzee to humans on at least three separate occasions; this strain is prevalent in North America and Europe.²⁶⁵ HIV-2 was passed to humans from the sooty mangabey and/or rhesus macaques and is prevalent in West Africa.²⁶⁶ Within the HIV-1 strain there are currently five identified subtypes and two circulating recombinant forms.²⁶⁷ It has been suggested that the current pandemic started in the mid-to late 1970s, with sporadic cases of HIV occurring prior to this.²⁶⁸ It is estimated that approximately 36 million adults and children are currently living with HIV, with around 3 million new HIV infections per year, and up to 2.5 million AIDS related deaths per year. The rate of new infections,

however, has dropped by 30% from its peak in 1996.²⁶⁹ In the western world, HIV and AIDS are becoming prevalent in the aging population. This is partially due to the effectiveness of anti-viral drugs allowing HIV infected persons to live longer, and partially due to the lack of anti-HIV campaigning amongst the older members of the community.²⁷⁰

HIV is an enveloped retrovirus (a virus whose genetic component is composed of RNA, ribose nucleic acid, and is surrounded by a layer of lipopeptide). HIV infects CD4+ T cells, with the protein CD4 being a glycoprotein receptor allowing the HIV virus to enter T cells (also referred to as T lymphocytes).²⁷¹ HIV replicates after the activation of infected CD4+ T cells.²⁷² Primary HIV infection is often asymptomatic, causing flu-like symptoms in approximately 50% of cases.²⁶⁵ The first immune response is an increase in levels of CD8+ and CD4+ T cells.²⁷³ Two to 6 months after the initial infection, HIV may enter a prolonged clinically latent period. The CD8+ T cell response is believed to maintain this apparently dormant period.^{274, 275} During this time, however, the virus remains active in lymphoid tissue.^{276, 277} Most important for the long-term course of HIV within an individual is the pool of latently infected CD4+ T cells which is established very early during primary infection.²⁷⁸ The final stages of HIV and the onset of AIDS occur when the virus escapes immune control.²⁷³ The gradual decrease in functional CD4+ T cells leads to immunodeficiency and eventual death in the untreated patient.²⁷⁹ Wasting is often pronounced in the later stages of the disease and the ultimate cause of death is usually infection by an opportunistic pathogen.²⁶⁵

HIV accumulates many mutations during the course of an infection in a single individual: approximately $10^9 - 10^{10}$ new virions are produced per day, with many mutations per nucleotide per replication cycle.^{280, 281} The high degree of mutations is due to the lack of DNA proof reading procedures in the HIV viral replication

mechanism.²⁸² The effectiveness of the antibody response is subsequently thwarted by rapid genetic changes in the envelope protein that allow the virus to escape recognition by antibodies in circulation at that time. The anti-body response is itself responsible for applying evolutionary pressure on the virus.²⁷⁵ The constant evolution of the HIV virus leads to difficulties in vaccine development.

Some people are immune to HIV: those with an inherited mutation in a chemokine receptor which is essential as a co-receptor for HIV entry, and a group of Gambian and Kenyan prostitutes, who's source of natural immunity is unknown but is speculated to be related to cytotoxic T cell activity.^{265, 283, 284}

4.1.2.2 Treatment Options

Current treatments for HIV and AIDS include combinations of viral protease inhibitors or nucleoside analogues.^{285, 286} A decrease in mortality due to AIDS has been observed in Europe and North America where antiretroviral agents have been utilised. Several strains of HIV have developed resistance to anti-viral drugs.²⁸⁷ However, there have been continued efforts in producing new therapies for HIV drug resistant strains, such as the introduction of Etravirine, a second-generation nonnucleoside reverse transcriptase inhibitor (NNRTI) active against NNRTI-resistant strains of HIV.²⁸⁸ An HIV vaccine, however, offers the best long-term option to control the AIDS pandemic, since the vast majority of individuals do not have adequate access to treatment, and new infections continue to occur.^{242, 243}

Exposed, uninfected people, such as sex workers have been found to have cytotoxic T lymphocytes against HIV. However, in HIV infected people, it has been found that many of these cells are non-functional, possibly due to impaired maturation as a result of infection.^{243, 289-292} A vaccine that could elicit the appropriate T lymphocyte response

prior to infection could have great potential for controlling the HIV virus.²⁴³ It has also been suggested that a vaccine that does not prevent infection, but reduces HIV levels and preserves uninfected memory CD4+ T cells, would be beneficial.²⁷³

Lipopeptides are promising vectors for an AIDS vaccine and are currently being evaluated as a therapeutic vaccine in chronically HIV-infected patients.²⁵⁵ Lipopeptide vaccines are presented to T cells by various pathways. Andrieu *et. al.*²⁵⁵ found that a vaccine consisting of the peptide Pol₄₆₁₋₄₈₄, that had been modified with palmitoyl at the N-terminal end, depended on the proteasome for processing. Conversely the peptide Nef₆₆₋₉₇, which had been modified at the C-terminal end, depended on endosomal acidification for processing. (Pol refers to the HIV Gag-Pol polyprotein; Nef is described in section 4.1.3.) Andrieu suggests that the position of the acyl moiety may direct the lipopeptide to different processing sites.²⁵⁵ Sequences, modified by palmitoylation and derived from HIV proteins Gag and Env, have also been tested in combination.²⁴² In addition, work by Gahery-Segard *et. al.* suggests that T cell epitopes induced in volunteers who have been vaccinated can be different from those induced in HIV infected patients.²⁴³

Using a combination of peptides may be advantageous, particularly in the case of HIV, to avoid the emergence of virus escape mutants. These were found in studies involving preclinical HIV related simian immunodeficiency virus (SIV) infected macaques, where a single lipopeptide was used, and the resultant immune response was monoepitopic.²⁹³ Gahery-Segard *et. al.* found that by using a combination of lipopeptides they were able to induce and maintain a multiepitopic CD8+ T cell response. This response was not necessarily random, but may depend on precise molecular peptide processing.²⁴³ Lipopeptides have the potential to be generated to cover many different epitopes that can react with different major histocompatibility complexes (MHCs). The identification

of highly conserved and widely recognised epitopes is a critical step in the development of a lipopeptide vaccine. The use of a mixed lipopeptide vaccine may also work in a greater number of individual patients where immune responses vary from individual to individual.²⁴³ Lipopeptides are currently being tested as AIDS vaccines in healthy individuals and as a therapeutic vaccine in chronically HIV infected patients.²⁵⁵⁻²⁵⁷

4.1.3 HIV type-1 Nef Protein

4.1.3.1 Background

Nef is a 27 kDa N-terminal myristoylated protein and is one of the first proteins to appear in HIV infected cells.²⁹⁴ Nef was originally given the name HIV type 1 negative factor, as it was thought to down-regulate virus replication and was speculated to play a role in the apparent latency of HIV.²⁹⁵ Later studies have shown, however, that Nef actually plays a central role in the promotion of viral replication.²⁹⁶ The Nef protein is unique to HIV and related SIV.²⁹⁷ Experimental evidence suggests that Nef enhances viral replication and pathogenesis through a combination of mechanisms.²⁹⁸ Nef targets the T cell receptor, thus priming viral replication and also plays a role in viral survival through immune evasion and antiapoptosis.^{279, 299}

In the early 1990s researchers at Harvard Medical School found that the deletion of the Nef gene from SIV lead to poor replication and nonpathogenesis in rhesus monkeys.^{300, 301} This indicated that Nef would be a promising target for antiviral drug development. Furthermore, rhesus monkeys vaccinated with live SIV attenuated by the deletion of the Nef gene were protected against live, pathogenic SIV.³⁰¹ These results suggested that a live, attenuated vaccine to protect against HIV in humans was attainable. However, recent findings in a macaque study, indicate that Nef-deletion mutants have retained virulence; with disease progression being slowed, but not prevented.³⁰² It appears that, due to continuous viral replication and selective pressure, an initially benign viral strain

was able to generate virulent progeny.³⁰² A vaccine capable of inducing an immunoglobulin G response specific to Nef could be of particular utility, since the HIV-1 Nef antigen is found on the surface of infected cells, where it allows the formation of a bridge between an infected cell and an uninfected cell.³⁰³ This function could be blocked by anti-Nef immunoglobulin antibodies.^{303, 304}

A number of acyl modified peptide sequences from Nef have been tested for immunogenic properties and some have been used in clinic trials of vaccines.^{242, 243, 255} Lipopeptides taken from the Nef sequence, modified with palmitoyl and other lipopeptides are being studied. It is assumed these can be safely studied in humans as there is no danger of viral infection with small peptide inoculation.

4.1.3.2 Nef₁₋₆ Peptide and Synthetic Derivatives

Having discovered a class of lipopeptides from *Mycobacterium tuberculosis*, that are presented by CD1a to T cells,³⁰⁵ our collaborators proposed the hypothesis that other lipid-modified peptides may be presented in a similar manner. N-terminal acylated Nef₁₋₆, amongst other similarly acylated peptides, was chosen to be tested for CD1 presentation to T cells. N-terminal acylated Nef₁₋₆ was purchased from a local peptide synthesis company, with (C14:0), palmitic (C16:0), stearic (C18:0) and oleic (C18:1) acids being used to synthesise four different acyl modifications on the Nef₁₋₆ peptide (Anaspec corporation, Boston, USA). Pools of acylated peptides, including the N-terminal acylated Nef₁₋₆ pool, were tested using an antigen-presenting cell (APC) producing four human CD1 proteins (CD1a-d). It was found that the Nef₁₋₆ pool of N-terminal acylated peptides resulted in a T cell activation when presented by CD1c.¹⁷⁶

The mixture of synthetic peptides containing the pool of N-terminal acylated peptides was fractionated using a HPLC with a split interface to allow simultaneous collection,

UV detection and ESI-QIT-MS analysis.¹⁷⁶ Our collaborators identified three major components in the synthetic mixture, with m/z values of 900.7 (singly charged), 867.4 (doubly charged) and 868.5 (doubly charged). The species with an m/z value of 900.7 was identified as Nef₁₋₆ with an N-terminal palmitoyl moiety, (pam)Nef₁₋₆. The species with m/z values of 867.4 (doubly charged) and 868.5 (doubly charged) did not match any of the expected acylated peptide products and our collaborators were unable to ascertain the structures of these components. These two components were subsequently named Nex 1 and Nex 2. Testing of the three major synthetic components for T cell response showed that only the Nex 1 and Nex 2 components stimulated a T cell response, with no response being observed for the (pam)Nef₁₋₆ acyl peptide.¹⁷⁶ For our collaborators to complete studies into the proposed mechanism of the interactions of Nex 1 and Nex 2 with the CD1c molecule, and its presentation of the antigens to T cells, it was critical to know the structures of Nex 1 and Nex 2.

4.1.4 Aims

The aim of this work was to use tandem mass spectrometry techniques developed in Chapter 3 of this thesis to characterise the Nex 1 and Nex 2 molecules. Our collaborators, through the use of ESI-QIT-MSⁿ had determined that Nex 1 and Nex 2 differed by 2 u and were structurally similar according to their mass spectra (personal communications). It was also assumed, since the Nex 1 and Nex 2 entities were found in the synthetic products resulting from the synthesis of *N*-acylated Nef₁₋₆, that they would in some way be structurally related to the Nef₁₋₆ peptide. This gave a starting point to the structural elucidation of these molecules. A number of analytical strategies were employed simultaneously in characterising the structure of Nex 1 and Nex 2 to confirm the structure with the greatest degree of confidence possible. The sample containing the mixture of synthesis products was analysed by MALDI-TOF MS, MALDI-Q-o-TOF

MS/MS, ESI-Q-o-TOF MS/MS and ESI-FTICR MS/MS. The mixture was also subjected to enzymatic digestion using trypsin and subsequently analysed by ESI-Q-TOF MS/MS.

4.2 Results and Discussion

4.2.1 MS Analysis of the Synthetically *N*-acylated Nef₁₋₆ Mixture

The synthetic mixture containing the acylated peptide Nef₁₋₆ and a number of synthetic by-products was obtained from collaborators, Moody *et. al.*, at Brigham and Women's Hospital in Boston, MA, USA.¹⁷⁶ In order to ascertain a general *m/z* profile for the sample, initial analysis of the sample was performed using MALDI-TOF MS. The sample was diluted 1:40 from its original concentration in 50% acetonitrile; 0.1% TFA and subsequently mixed 1:1 with α -cyano matrix (in the same solvent) and allowed to air dry prior to analysis. Figure 4.1 shows the MALDI-TOF mass spectrum of the acylated Nef₁₋₆ synthetic mixture. Proposed identities for the majority of species observed in the sample can be found in Table 4.1. The major component of the sample was (pam)Nef₁₋₆ (*m/z* 900.59), along with sodium and potassium adducts of this peptide. The remaining major components being Nex 2 (*m/z* 1736.02), followed by Nex 1 (*m/z* 1733.99). Small peaks were found in the MALDI-MS spectrum corresponding to the additional anticipated products of the acylated Nef₁₋₆ synthesis (see section 4.1.3.2 above), being (Myr)Nef₁₋₆ (*m/z* 872.54), (ole)Nef₁₋₆ (*m/z* 926.61) and (ste)Nef₁₋₆ (*m/z* 928.63). The majority of the remaining masses can be attributed to sequences based on expected addition, truncation or deletion products resulting from the peptide synthesis process.³⁰⁶ Further characterisation of (pam)Nef₁₋₆, Nex 1 and Nex 2 was carried out by MALDI-Q-o-TOF MS/MS and ESI-Q-o-TOF MS/MS and ESI-FTICR MS/MS analysis.

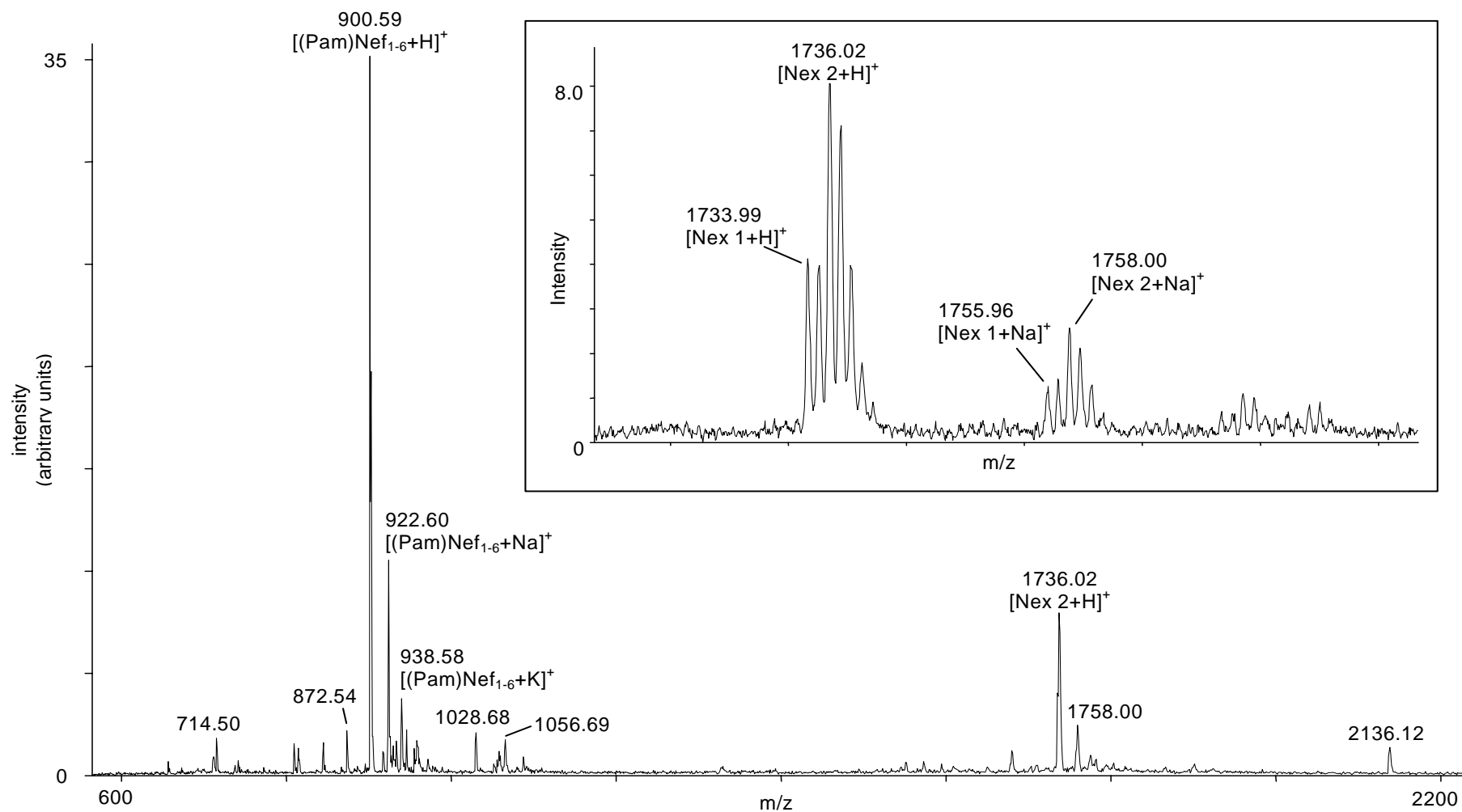


Figure 4.1 MALDI-TOF MS of the acylated Nef₁₋₆ synthetic mixture. Inset: expansion of the m/z region showing Nex 1 and Nex 2 ions. (Table 4.1 provides a detailed list of the masses observed in this spectrum.)

Table 4.1 Proposed identities of species present in the acylated Nef₁₋₆ synthetic mixture.

Experimental [m+H] ⁺ /u	Relative intensity (%)	Proposed assignments [†]	Calculated [m+H] ⁺ /u	Error (ppm)
714.50	6.5	[G(pam)GKSK + H] ⁺	714.51	-18.2
813.59	5.1	[G(pam)GKWK + H] ⁺	813.56	35.3
843.58	5.8	[G(pam)KWSK + H] ⁺	843.57	6.1
872.54	7.5	[(Myr)Nef ₁₋₆ + H]1 ⁺	872.56	-26.8
900.59	100.0	[(pam)Nef ₁₋₆ + H]1 ⁺ [G(pam)GKWSK + H] ⁺	900.59	-0.3
922.60	30.9	[(pam)Nef ₁₋₆ + Na] ⁺	922.57	29.0
926.61	4.1	[(ole)Nef ₁₋₆ + H]1 ⁺	926.61	-1.2
928.63	5.4	[(ste)Nef ₁₋₆ + H]1 ⁺	928.62	10.1
938.58	11.9	[(pam)Nef ₁₋₆ + K] ⁺	938.55	33.9
957.64	6.1	[G(pam)GKWSK + Gly + H] ⁺	957.61	24.3
1028.69	7.2	[G(pam)GKWSK + Lys + H] ⁺	1028.65	33.2
1054.70	3.5	[G(ole)GKWSK + Lys + H] ⁺	1054.78	-76.2
1056.70	4.7	[G(ste)GKWSK + Lys + H] ⁺	1056.80	-92.5
1086.68	3.9	[G(pam)GKWSK + Trp + H] ⁺	1086.67	7.7
1676.89	2.9	[G(ole)KWSKXSKWSK + H] ⁺	1676.98	-51.4
1678.88	4.8	[G(ste)KWSKXSKWSK + H] ⁺	1678.99	-67.7
1707.92	2.7	[G(pam)GKWSKXSKWSK + H] ⁺	1707.98	-37.4
1733.99	12.6	[Nex 1 + H] ⁺ [G(ole)GKWSKXSKWSK + H] ⁺	1734.00	-5.0
1736.02	23.7	[Nex 2 + H] ⁺ [G(ste)GKWSKXSKWSK + H] ⁺	1736.02	-0.1
1755.96	4.6	[Nex 1 + Na] ⁺	1755.98	-9.5
1758.00	8.3	[Nex 2 + Na] ⁺	1758.00	1.7
1771.90	3.0	[Nex 1 + K] ⁺	1771.95	-30.2
1773.88	4.1	[Nex 2 + K] ⁺	1773.97	-51.3
2136.12	4.1	[Nex 2 + WSK + H] ⁺	2137.22	-514.9

[†] pam: palmitoyl (C16:0); Myr: Myristoyl (C14:0); ste: stearoyl (C18:0); ole: oleoyl (C18:1); X: ester linked kynurenine residue (see section 4.2.3.3 for more details).

4.2.2 MS/MS of *N*-glycine Palmitoylated Nef₁₋₆

N-glycine palmitoylated Nef₁₋₆ ((pam)Nef₁₋₆) was analysed by MS/MS as a standard compound for comparison with the MS/MS of the unsequenced Nex 1 and Nex 2. These peptides were predicted to have some sequence in common to the acylated Nef₁₋₆ peptides which were the target of the synthetic process. (pam)Nef₁₋₆ was analysed by MS/MS using three separate techniques, namely, MALDI-Q-o-TOF MS/MS, ESI-Q-o-TOF MS/MS and ESI-FTICR MS/MS (details of the instrumentation can be found in Chapter 2). For MALDI-Q-o-TOF analysis, the sample was diluted 1:40 with 50% acetonitrile, 0.1 % TFA and subsequently mixed 1:1 with α -cyano matrix (in the same solvent) and allowed to air dry prior to analysis. For ESI-Q-o-TOF the sample was diluted 1:40 with 50 % acetonitrile, 0.05 % formic acid and introduced via nanospray using a glass capillary with an orifice of approximately 1 μ m. The sample for analysis by ESI-FTICR MS/MS was prepared and introduced to the mass spectrometer in the same manner as for the sample analysed by ESI-Q-o-TOF MS/MS, except that the solvent used was 50 % methanol, 0.05 % formic acid. The spectra from these three analyses are shown in Figure 4.2 - 4.4, with tabulated assignments in Appendices A and B and Table 4.2. Scheme 4.1 shows the sequence of (pam)Nef₁₋₆, indicating the major sequence ion fragmentation. Scheme 4.2 shows the structure of the palmitoyl moiety.

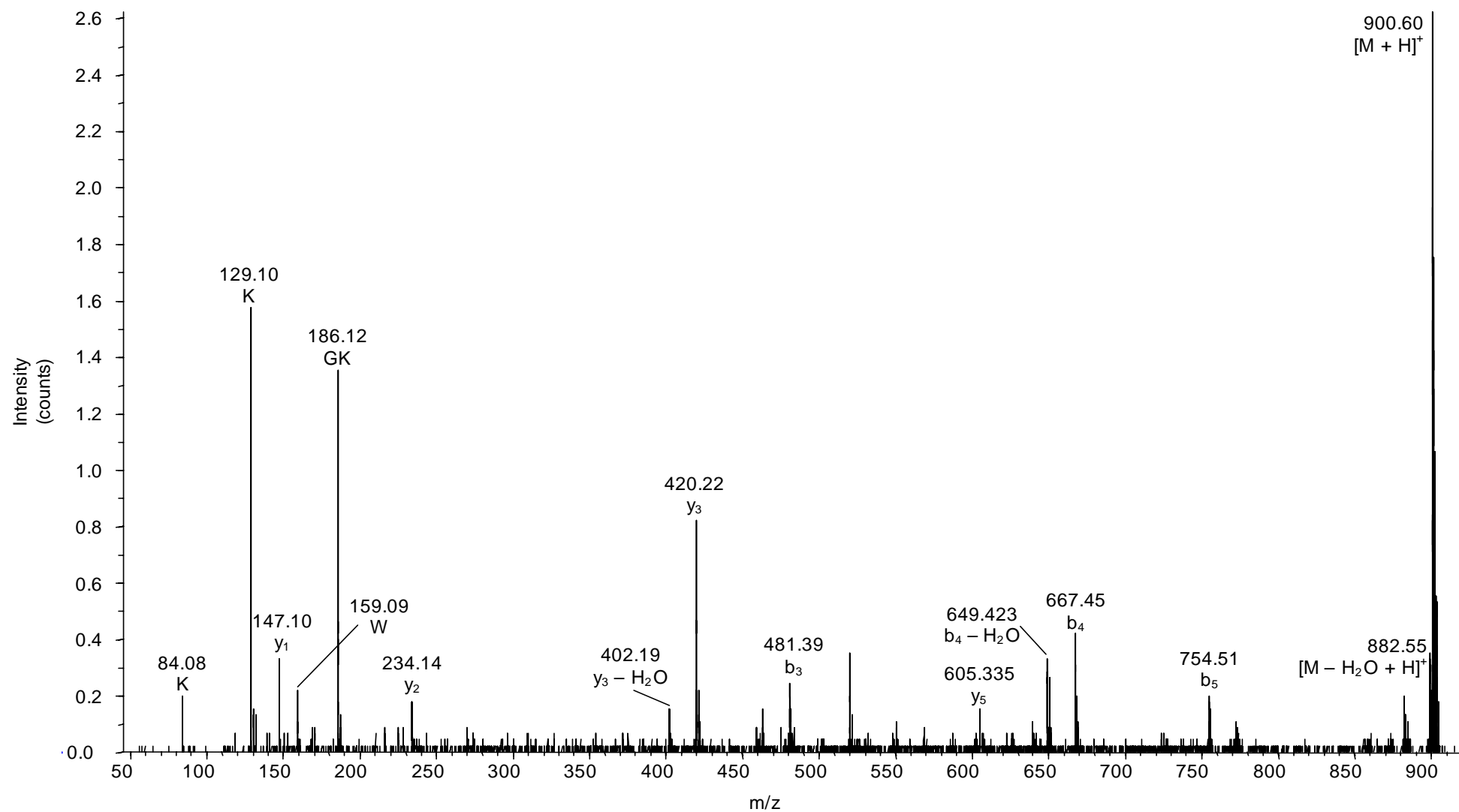


Figure 4.2 MALDI-Q-o-TOF MS/MS of (pam)Nef₁₋₆, m/z 900.6 precursor ion. Collision energy 40 eV. Detailed assignments are given in Appendix A.

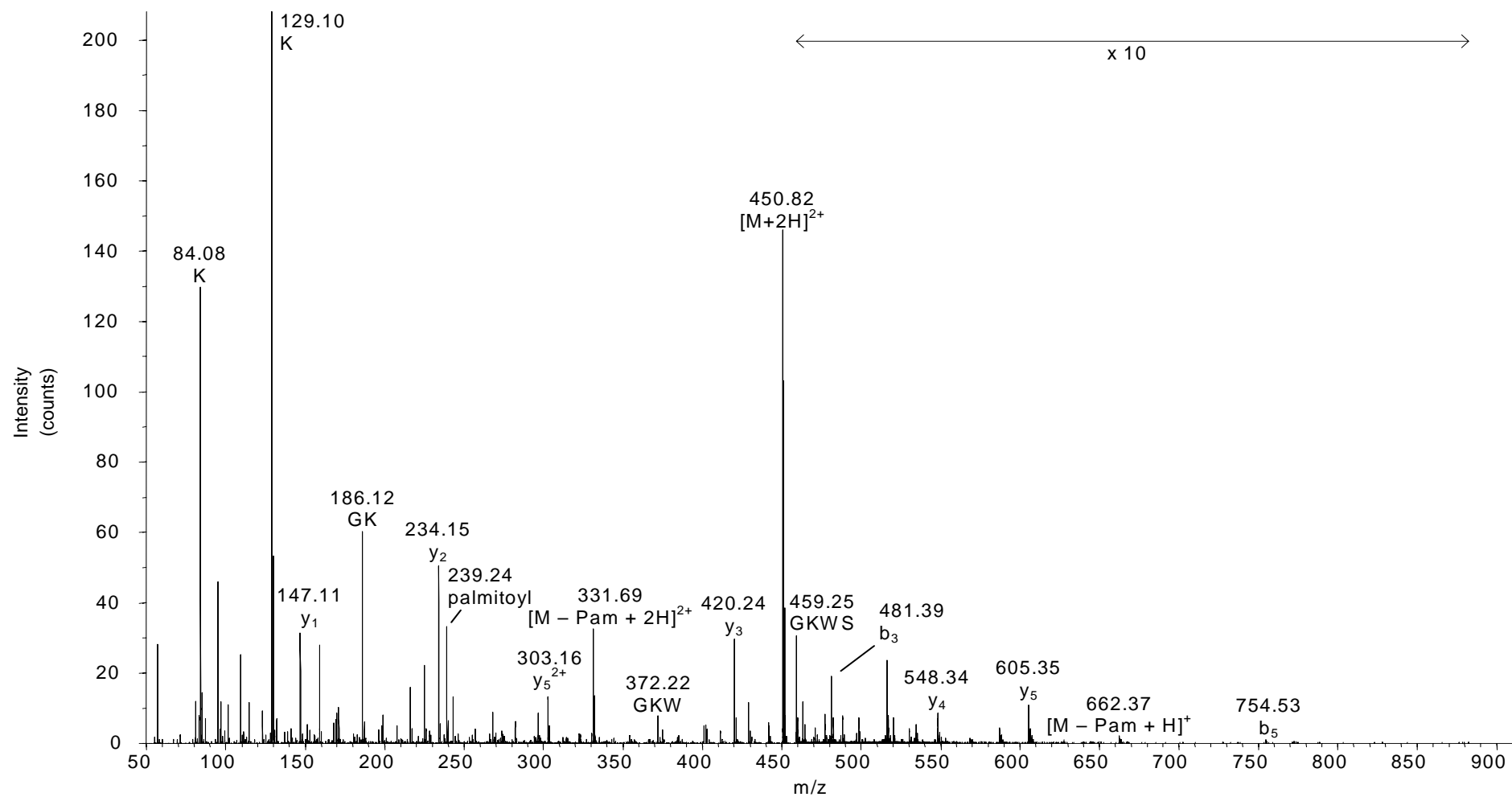


Figure 4.3 ESI-Q-o-TOF MS/MS of (pam)Nef₁₋₆, m/z 450.81 precursor ion. Collision energy 30 eV. Detailed assignments are given in Appendix B.

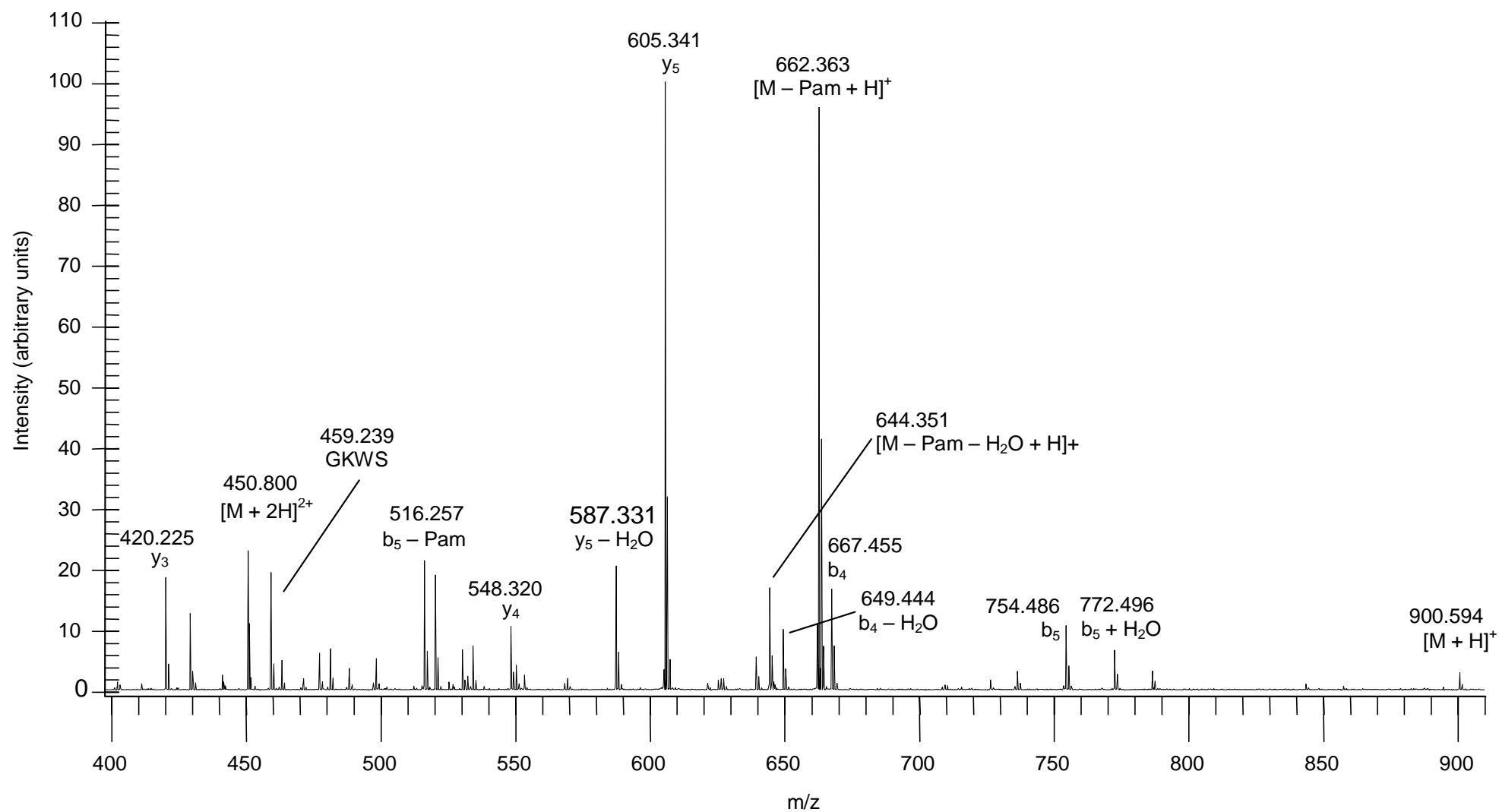


Figure 4.4 ESI-FTICR MS/MS of (pam)Nef₁₋₆ m/z 450.8 precursor ion, Q2 CAD at 10 eV, 1000 ms external accumulation. Detailed assignments are given in Table 4.2.

All three MS/MS techniques gave good sequence coverage for the (pam)Nef₁₋₆ peptide. The MALDI-Q-o-TOF MS/MS resulted in complete sequence ion coverage, with the product ions b₁₋₅ and y₁₋₅ being observed. The ESI-Q-o-TOF MS/MS spectrum shows a complete y-ion series (y₁₋₅) and three b-ions (b₁, b₃ and b₅). The ESI-FTICR MS/MS spectrum showed a partial sequence for both the y-ions and b-ions (y₃₋₅ and b₃₋₅).

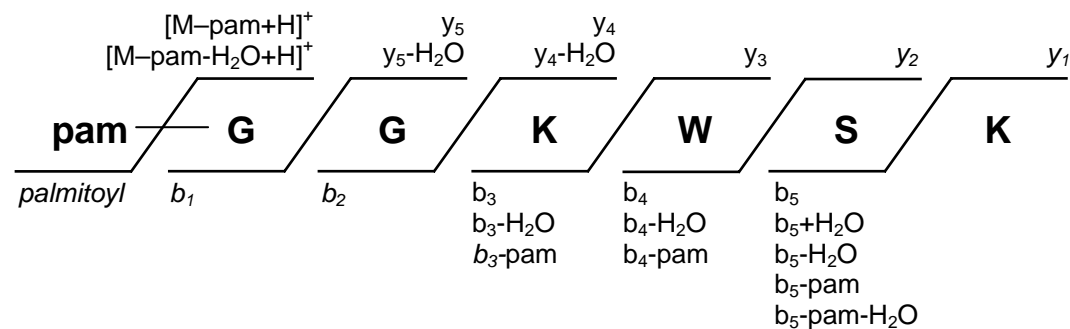
In addition to product amino acid sequence ions, a number of ions are observed as a result of the loss of the palmitoyl moiety from the precursor ion and also from a number of the product sequence ions. For example, the $[M - \text{pam} + H]^+$, m/z 662, ion in the ESI-Q-o-TOF MS/MS and ESI-FTICR MS/MS spectra; as well as similar losses from b₅, b₄ and b₃ ions. The ions which result from the loss of the acyl moiety from the product sequence ions, are useful for determining the structure of the peptide as they indicate the part of the sequence at which the acyl moiety is attached.

Another useful diagnostic ion is the b₁ ion, this being the N-terminal palmitoylated glycine ion, m/z 296. This ion is observed in both the MALDI and ESI-Q-o-TOF spectra. Similar ions were observed Chapter 3, where an acylated b₁ ion was observed when the N-terminal residue of a peptide was acylated. In contrast, when an internal residue was acylated, an acylated immonium ion was observed instead (see Chapter 3). An additional diagnostic ion is the palmitoyl carbenium ion at m/z 239, as observed in the MALDI and ESI-Q-o-TOF mass spectra. This ion, along with the (pam)Nef₁₋₆ b₁ ion, allows for the confirmation of the presence of the palmitoylated N-terminal glycine residue.

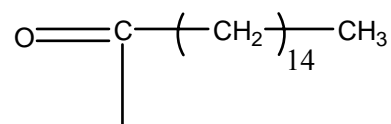
Table 4.2 Assignment of ESI-FTICR MS/MS results for (pam)Nef₁₋₆.

<i>m/z</i>	<i>z</i>	Experimental [<i>m</i> + <i>H</i>] [†] / <i>u</i>	% base peak height	assignment [†]	Theoretical [<i>m</i> + <i>H</i>] [†] / <i>u</i>	error (ppm)
450.800	2	900.592	23.5	[<i>M</i> +2 <i>H</i>] ²⁺	900.592	0.1
900.594	1	900.594	3.3	[<i>M</i> + <i>H</i>] ⁺	900.592	1.8
772.496	1	772.496	6.6	<i>b</i> ₅ + <i>H</i> ₂ O	772.497	-1.1
754.486	1	754.486	10.7	<i>b</i> ₅	754.487	-0.7
736.476	1	736.476	3.5	<i>b</i> ₅ - <i>H</i> ₂ O	736.476	-0.5
726.490	1	726.490	2.0	<i>a</i> ₅	726.492	-2.5
667.454	1	667.454	17.0	<i>b</i> ₄	667.455	-1.0
662.363	1	662.363	96.0	[<i>M</i> -pam+ <i>H</i>] ⁺	662.363	0.0
649.444	1	649.444	10.1	<i>b</i> ₄ - <i>H</i> ₂ O	649.444	-0.9
644.351	1	644.351	17.4	[<i>M</i> -pam- <i>H</i> ₂ O+ <i>H</i>] ⁺	644.352	-1.3
639.460	1	639.460	5.6	<i>a</i> ₄	639.460	-0.1
605.340	1	605.340	100.0	<i>y</i> ₅	605.341	-1.2
587.331	1	587.331	21.4	<i>y</i> ₅ - <i>H</i> ₂ O	587.331	0.4
548.320	1	548.320	11.3	<i>y</i> ₄	548.320	0.5
534.267	1	534.267	7.5	<i>b</i> ₅ + <i>H</i> ₂ O-pam	534.270	-6.2
530.309	1	530.309	7.0	<i>y</i> ₄ - <i>H</i> ₂ O	530.309	-0.8
516.257	1	516.257	22.4	<i>b</i> ₅ -pam	516.260	-5.7
498.246	1	498.246	5.5	<i>b</i> ₅ - <i>H</i> ₂ O-pam	498.246	-0.7
488.262	1	488.262	3.9	<i>a</i> ₅ -pam	488.260	3.4
481.376	1	481.376	7.3	<i>b</i> ₃	481.375	1.2
477.246	1	477.246	6.5	GKWS+ <i>H</i> ₂ O	477.246	-1.0
463.365	1	463.365	5.2	<i>b</i> ₃ - <i>H</i> ₂ O	463.365	-0.1
459.239	1	459.239	3.6	GKWS	459.236	7.3
441.224	1	441.224	2.5	GKWS- <i>H</i> ₂ O	441.225	-1.8
429.225	1	429.225	13.3	<i>b</i> ₄ -pam	429.225	0.0
420.225	1	420.225	19.1	<i>y</i> ₃	420.225	0.6

[†] pam refers to the palmitoyl adduct, located at the N-terminal glycine of Nef₁₋₆



Scheme 4.1 Annotated sequence of (pam)Nef1-6. Plane text indicates product ions seen in the ESI-FTICR MS/MS spectrum. Italicised text indicates additional peaks seen in MALDI or ESI-Q-o-TOF spectra. Charge states are omitted for clarity; see Appendices A and B and Table 4.2 for complete assignments.



pam (16:0)
f.w. 238.2297

Scheme 4.2 Structure of the palmitoyl moiety.

Of the three MS/MS techniques utilised in the analysis of (pam)Nef₁₋₆ ESI-FTICR MS/MS resulted in the spectra with the best sensitivity (as determined by the signal-to-noise ratio), highest resolution and least complicated spectrum, with the majority of ions being easily assigned to b or y sequence ions. This is in comparison to ESI and MALDI-Q-o-TOF, which had a greater number of internal fragment ions and losses of both the atoms of H₂O and NH₃ from product sequence ions, complicating the interpretation of the spectra. The successful mass spectrometric analysis of (pam)Nef₁₋₆, which was expected to have sequence in common with Nex 1 and Nex 2, gave a good understanding of how this type of peptide fragments, using three different MS/MS techniques.

4.2.3 MS/MS of Nex 1 and Nex 2

Nex 1 and Nex 2 were found to differ in mass by 2 u (See Figure 4.1, inset). Using MALDI-Q-o-TOF MS/MS, where the predominant charge state for the precursor ions was 1+, unit resolution isolation was used for the isolation of Nex 2, the more abundant of the two Nex species. This isolation was achieved without a significant loss in sensitivity, resulting in an individual MS/MS spectrum for Nex 2. The MALDI-Q-o-TOF MS/MS spectra of Nex 2, and of a combined Nex 1 and Nex 2, are shown in Figure 4.5 below. The MALDI-Q-o-TOF MS/MS spectra resulted in fragment ions that are consistent with the fragmentation of a peptidic molecule. For example, the low mass ions m/z 84 and m/z 129 are associated with the presence of lysine residues. The m/z 186 corresponds to residue combination, GK. Additionally, the difference between several of the ions corresponds with the expected mass difference for a number of amino acid residues, as shown in Figure 4.5 a). These ions suggest that C-terminal sequence is KWSK. The apparent presence of lysine residues, the internal sequence GK and C-

terminal sequence KWSK were strong indicators that the sequences of Nex 1 and Nex 2 have components in common with the Nef₁₋₆ sequence: GGKWSK.

The MALDI-Q-o-TOF MS/MS spectra show that the Nex 1 and Nex 2 spectra differ by 2 u only for specific fragment ions, as indicated by an asterisk, and otherwise have isometric fragment ions. This indicates that Nex 1 and Nex 2 have very similar sequences, differing by 2 u in only one section of the peptide. The presence of relatively high m/z product ions, for example, m/z 1228, 1357 and 1413, that do not differ by 2 u suggested that the difference of 2 u can be attributed to a species present towards one end or side-chain branch of the peptide, rather than in the middle of the sequence. Whilst these spectra gave a strong indication that Nex 1 and Nex 2 were essentially identical they were not sufficiently detailed to allow a full sequence characterisation of the molecules.

Nex 1 and Nex 2 were also analysed using both ESI-Q-o-TOF MS/MS and ESI-FTICR MS/MS (data shown below). As was found for the analysis of (pam)Nef₁₋₆, the spectra resulting from FTICR analysis was more detailed, with greater resolution, than that obtained using the Q-o-TOF instrument. When analysed by ESI-MS/MS (Q-o-TOF or FTICR), Nex 1 and Nex 2 were found to have a predominant charge state of 3+. This resulted in a difference of approximately m/z 0.6 between the monoisotopic forms of the two species, differing in mass by 2 u. Attempts to isolate the two components individually for ESI-MS/MS analysis resulted in significant losses in sensitivity and concomitant loss of charge state information that is usually acquired when the isolation window encompasses an isotopic cluster. Consequently, for the purposes of ESI-MS/MS analysis, Nex 1 and Nex 2 were isolated simultaneously by Q1 when analysed by ESI-MS/MS.

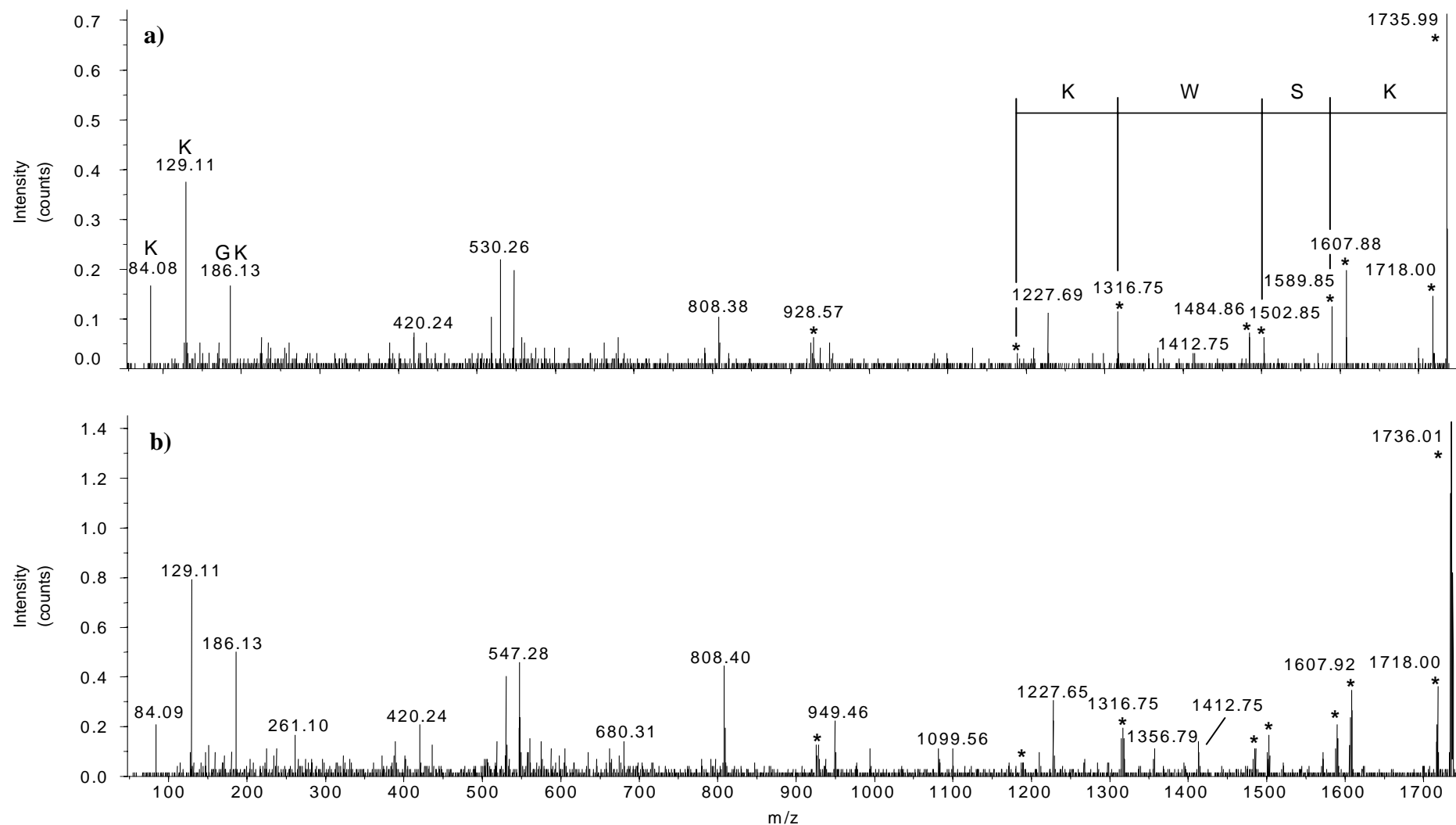


Figure 4.5 MALDI-Q-o-TOF MS/MS of Nex 1 and Nex 2. MALDI-Q-o-TOF MS/MS of Nex 2; m/z 1736 precursor ion, using a high resolution setting for the isolation, allowing for the isolation of only the monoisotopic ion of Nex 2 a). Combined MALDI-Q-o-TOF MS/MS of Nex 1 and Nex 2 using a precursor isolation value of m/z 1734 with a low resolution setting, allowing for the isolation of both Nex species b). Collision energy 65 eV. A table of assignments can be found in Appendix C.

4.2.3.1 Identification of the Acyl Component in Nex 1 and Nex 2.

Critical to the successful characterisation of Nex 1 and Nex 2 was the identification of the components resulting in a mass difference of 2 u. Calculation of the mass difference between Nex 1 and Nex 2 revealed that these two peptides differed, most likely, by two hydrogen atoms. The delta mass between Nex 1 and Nex 2 was determined to be 2.013 ± 0.004 u, which compares favourably with the calculated value for the sum of two hydrogen atoms of 2.015 u. (See Table 4.3 for calculations.) These findings suggested that the difference between Nex 1 and Nex 2 was likely to result from a double carbon bond being present in Nex 1, but not in Nex 2. Since stearic (C18:0) and oleic (C18:1) acids were used in the initially planned synthesis of *N*-acylated Nef₁₋₆, it was concluded that Nex 1 was most likely acylated by oleoyl and Nex 2 by stearoyl.

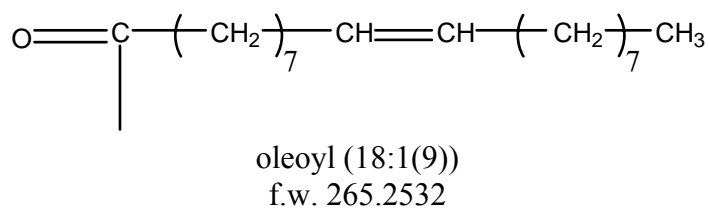
Table 4.3 Calculation of the mass difference between Nex1 and Nex 2.

Data Source	Charge state	Mass of Nex 1 [m+H] ⁺ / u	Mass of Nex 2 [m+H] ⁺ / u	Difference / u
ESI-FTICR MS	2+	1733.99346	1736.00670	2.01324
ESI-FTICR MS	3+	1733.99837	1736.01128	2.01291
MALDI-Q-o-TOF MS	1+	1734.00413	1736.01243	2.00830
ESI-Q-o-TOF MS	3+	1733.95503	1735.97383	2.01880
average				2.013
standard deviation				0.004
mass of 2xH				2.015
difference / u				0.002

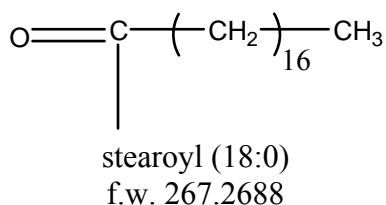
When comparing the ESI-Q-o-TOF MS/MS spectrum of (pam)Nef₁₋₆ (see Figure 4.6) with that of Nex 1 and Nex 2 (Figure 4.6) it was noted that the ion at *m/z* 239, assigned to a palmitoyl carbenium ion in the spectrum of (pam)Nef₁₋₆, was absent from the tandem mass spectra of Nex 1 or Nex 2. The ion was replaced by an oleoylium ion (*m/z* 265, Nex 1) and a stearoylium ion (*m/z* 267, Nex 2). Additionally, the palmitoylated b₁ ion found in the spectrum of (pam)Nef₁₋₆ (*m/z* 296) was replaced by two proposed b₁

ions at m/z 322 and m/z 324 for Nex 1 and Nex 2 respectively (see Figure 4.6 inset). Structures of the oleoyl and stearoyl moieties are shown in Scheme 4.3. The presence of b_1 ions analogous to that found in (pam)Nef₁₋₆ indicates that the acylations for Nex 1 and Nex 2 are also present on an N-terminal glycine residue. Hence Nex 1 was predicted to contain an oleoyl acyl modification and Nex 2 to contain a stearoyl acyl modification to the N-terminal glycine as an analogues to the (pam)Nef₁₋₆ peptide.

a)



b)



Scheme 4.3 Structure for a) oleoyl and b) stearoyl moieties.

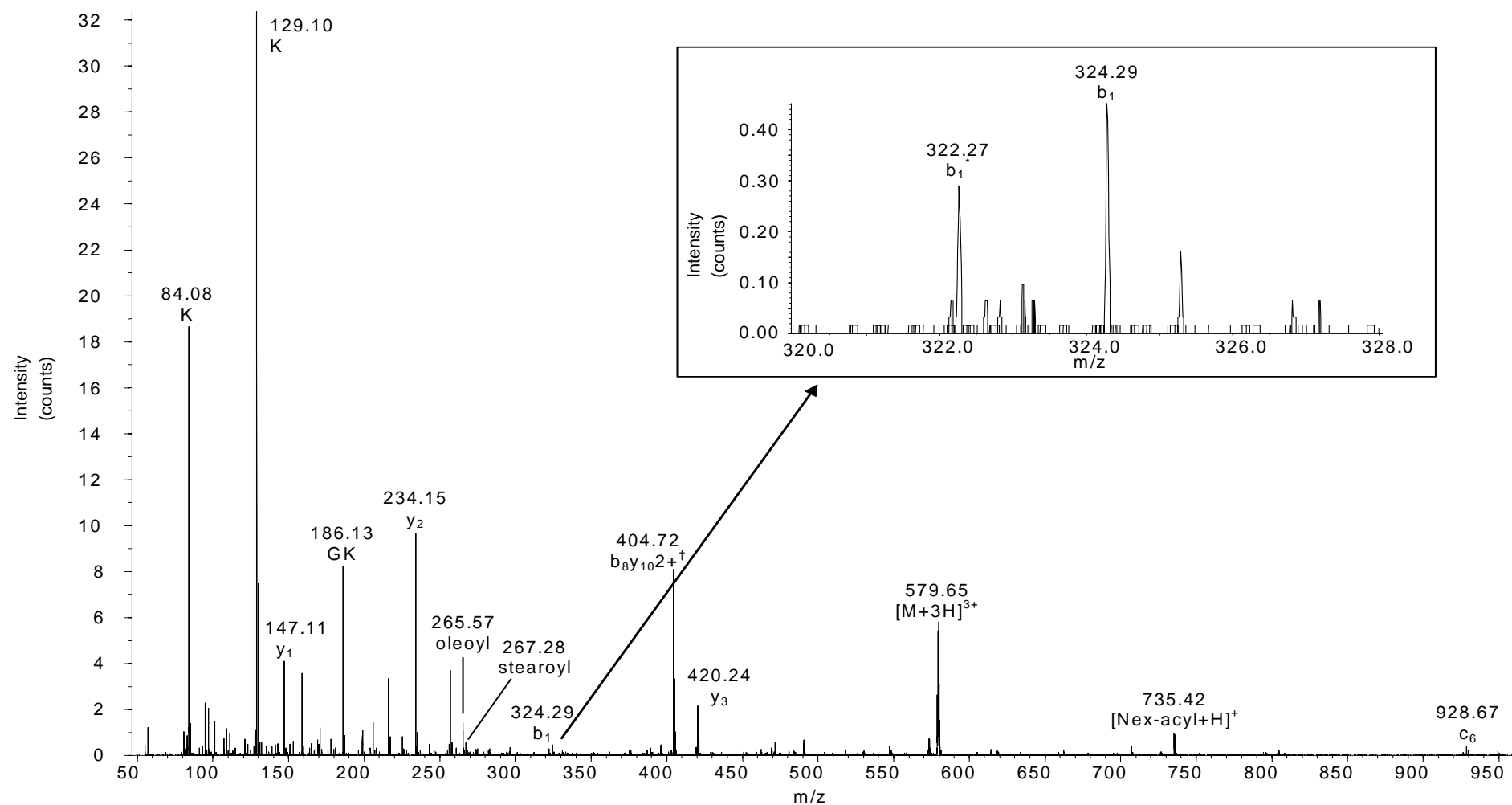


Figure 4.6 Combined ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 using a precursor isolation value of m/z 578.7 with a low resolution setting, allowing for the isolation of both Nex species. † indicates species with four possible sequence combinations. (Detailed assignments can be found in Appendix D.)

4.2.3.2 Sequence Determination of Nex 1 and Nex 2.

Analysis of the ESI-FTICR MS/MS spectrum of Nex 1 and Nex 2 (Figure 4.7) revealed consistent delta mass differences of 264 and 266 u from the respective precursor ions and product sequence ions of Nex 1 and Nex 2. Based on the assumption that Nex 1 and Nex 2 were acylated at their N-termini (due to the presence of m/z values attributed to b_1 ions, see section 4.2.3.1), ion pairs differing by 2 u, which were accompanied by a loss of 264 u or 266 u, were assigned as b-type fragment ions. B-type sequence ions with the presence and absence of the acyl group were found for the $b_8 - b_{11}$ ions. For example, b_9 ions m/z 1314.8 and 1316.8 and a corresponding b_9 -acyl ion can be seen in the ESI-FTICR MS/MS spectrum (Figure 4.7) of Nex 1 and Nex 2. Proposing a structure for Nex 1 and Nex 2 comprising of a (pam)Nef₁₋₆ like sequence at the N-terminal end, with the palmitoyl group having been replaced with an oleoyl group (Nex 1) or a stearyl group (Nex 2), allowed for the rapid assignment of a number of the remaining sequence ions. Fragment ions corresponding to predicted m/z for y_{9-11} and b_{4-6} were observed in the ESI-FTICR MS/MS spectrum.

After the identification of the N-terminal and C-terminal amino acid residues, there remained an internal mass difference of 191.06 u that did not match a naturally occurring or an expected acyl-modified amino acid. Table 4.4 shows the data and calculations used to determine this value. The mass value of the unidentified amino acid was calculated by taking the precursor or product ion experimental masses and subtracting the theoretical masses of the precursor or product ions excluding the unidentified component. The resultant values were averaged to determine an empirical mass for the unidentified component. The data selected for this calculation were for ions

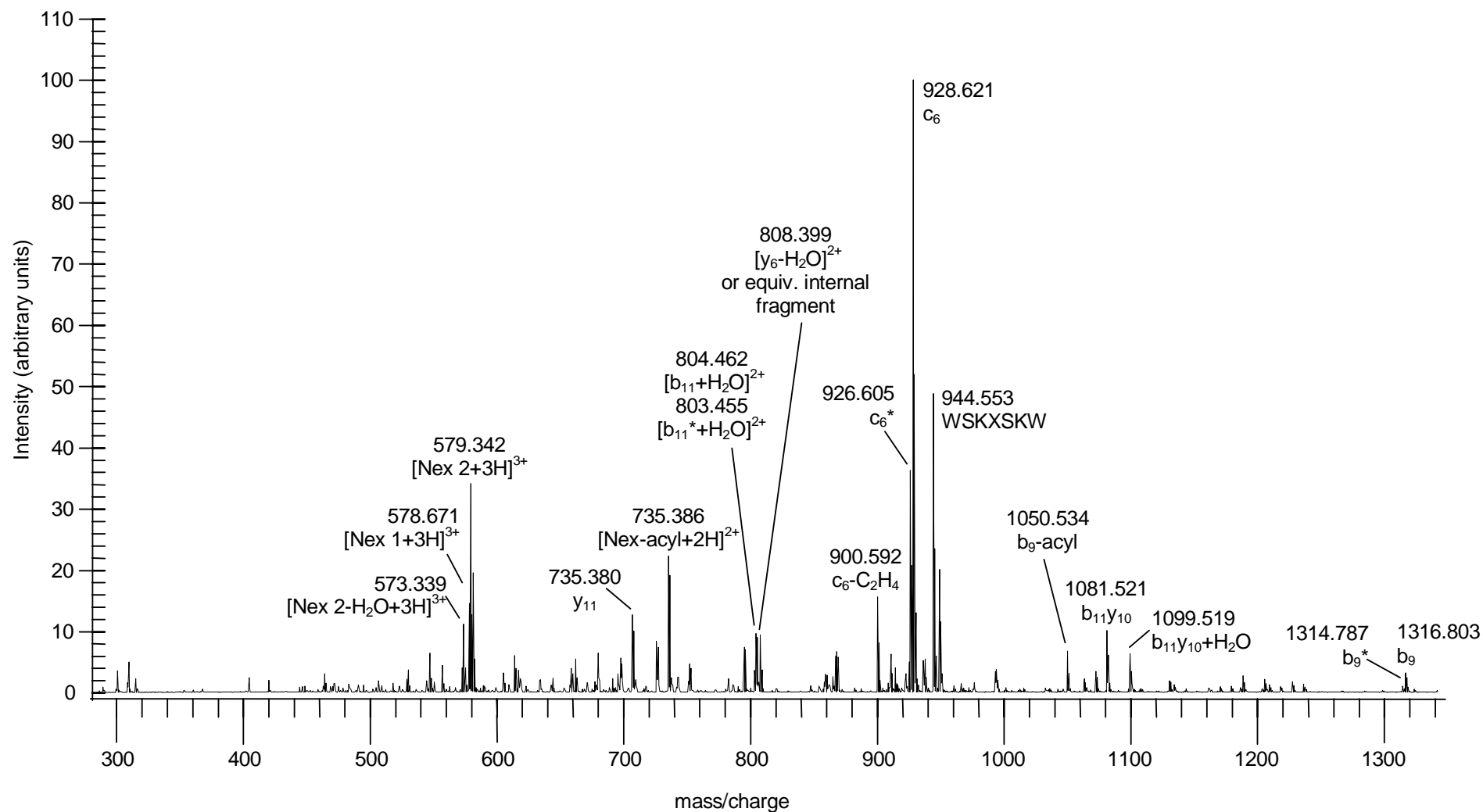


Figure 4.7 ESI-FTICR MS/MS of Nex 1 and Nex 2; m/z 579 precursor ion, Q2 CAD at 10 eV, 1000 ms external accumulation. A summary of the assignments can be found in Table 4.5. More detailed assignments are given in Appendix E. * indicates the oleoyl (18:1) form of the product ion, unmarked refers to ions derived from Nex 2 or either form of precursor ion. “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from the product ion. X denotes an ester linked kynurine (see text for explanation).

of relatively high intensity, which also encompassed the portion of peptide sequence containing the unidentified residue. (The possibility that this mass difference could be attributed to two separate amino acid type species was explored, however there were no product ion species that supported this possibility. For example, when looking for m/z differences due to the loss of water or the loss of the acyl group, there were no additional m/z values corresponding to any smaller components.)

Table 4.4 Calculations to determine the mass of amino acid-7.

Peptide sequence	Experimental [m+H] ⁺ mass / u [†]	Theoretical [m+H] ⁺ mass excluding amino acid residue-7 / u	Experimental mass of amino acid residue-7 / u
[Nex 1+2H] ²⁺	1733.9935	1542.9412	191.0523
[Nex 1+3H] ³⁺	1733.9984	1542.9412	191.0572
[Nex 2+2H] ²⁺	1736.0067	1544.9568	191.0499
[Nex 2+3H] ³⁺	1736.0113	1544.9568	191.0545
[Nex-acyl+2H] ²⁺	1469.7532	1278.6959	191.0573
[Nex-acyl+3H] ³⁺	1469.7561	1278.6959	191.0602
b ₉ [*] 1 ⁺	1314.7874	1123.7243	191.0631
b ₉ [*] 2 ⁺	1314.7837	1123.7243	191.0594
b ₉ 1 ⁺	1316.8028	1125.7400	191.0628
b ₉ 2 ⁺	1316.7980	1125.7400	191.0580
y ₉ 1 ⁺	1227.6197	1036.5580	191.0617
y ₉ 2 ⁺	1227.6149	1036.5580	191.0570
Average mass difference			191.0578 ± 0.004^{††}

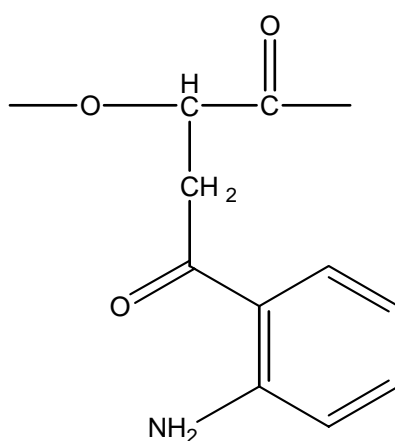
[†] [m+H]⁺ masses taken from ESI-FTICR MS/MS results, as tabulated in Table 4.5 and Appendix E. ^{††} standard deviation of the value.

* Indicates fragment ion derived from oleoyl modified Nex; no marking indicates fragment ions derived from stearoyl modified Nex.

On inspection of the ESI-FTICR MS/MS spectrum it was observed that two major product ions had not been assigned, namely m/z 926.6 and m/z 928.6 (see Figure 4.7 and Table 4.5). These ions correspond to the expected b₆ ions of Nex 1 and Nex 2 with the addition of 18 u. The closest match to an expected sequence ion was to that of c₆, with calculated m/z values of 925.6 and 927.6 for Nex 1 and Nex 2 respectively. This

suggested the usual amide linkage between Lys-6 and the unidentified amino acid at position 7 may have been replaced by an ester linkage. This would give c_6 ions with m/z values of 926.6 or 928.6 for Nex 1 or Nex 2 respectively.

Replacing an amide with an ester link resulted in a remaining unexplained mass of 190 u. It was noticed that this value is 4 u higher than expected for a tryptophan residue (186 u). There have been previous reports of the oxidative cleavage of the indole ring of tryptophan, increasing the mass by 4 u, to produce kynurenine (Kyn).^{51, 307-309} The process by which this Kyn residue could have been created as a by-product in the synthesis of acylated Nef₁₋₆ is discussed in the text below. The structure of a proposed ester-linked kynurenine is shown in Scheme 4.4 below. The mass of this residue determined experimentally (Table 4.4) was 191.0578 u, and the calculated mass is 191.0582 u, giving a mass difference of 0.0004 u (2 ppm).



residue mass: 191.0582 u

Scheme 4.4 Structure of ester-linked kynurenine residue (denoted X throughout text).

A summary of the ions observed in the ESI-FTICR MS/MS spectrum of Nex 1 and Nex 2 can be found in Table 4.4, with a complete analysis of the spectrum in Appendix E. Scheme 4.5 shows the major fragmentation ions observed for Nex 1 and Nex 2 as found with ESI-FTICR, MALDI-Q-o-TOF and ESI-Q-o-TOF MS/MS. (The data from the ESI-Q-o-TOF MS/MS analysis is presented in Appendix D.) A number of isometric fragment ions were observed in the ESI-FTICR MS/MS of Nex 1 and Nex 2, such as the ions y_1 and internal fragment ions WSK+H₂O, SKW+H₂O and KWS+H₂O. Similar longer chain y-ion series ions had isometric internal fragment ions. This redundancy is as a result of the repeating WSK motif.

As described in the introduction to this chapter (Acyl)Nef₁₋₆ was synthesised by a local supplier using a mixture of acylating agents. It appears, however, that Nex 1 and Nex 2 were unexpected by-products of the synthetic scenario. It is possible that the oleoyl and stearoyl modified Nex peptides were initially synthesised correctly, but reacted post-synthetically during the steps required to remove the completed peptide from the synthesis medium. The MALDI-TOF mass spectrum of the acylated Nef₁₋₆ synthetic mixture (Figure 4.1 and Table 4.1) shows two low abundance ions that can be attributed to oleoylated Nef₁₋₆ and stearoylated Nef₁₋₆, with m/z values of 926.6 and 928.6 respectively, indicating that they were synthesised correctly, in small quantities. The peptides were synthesised using a solid state column with the C-terminal Lys-6 of Nef₁₋₆ being the first amino acid laid down, and acylation of the N-terminal glycine the final synthetic step. It is postulated that following the final synthetic step, Nex 1 and Nex 2 were produced as a result of the differing hydrophobicity of their slightly longer N-terminal acyl chains compared with (pam)Nef₁₋₆, and the elution conditions used. This resulted in the two peptides being eluted at a slightly slower rate from the synthesis column when compared to (pam)Nef₁₋₆. It is possible that the peptide WSKWSK, which

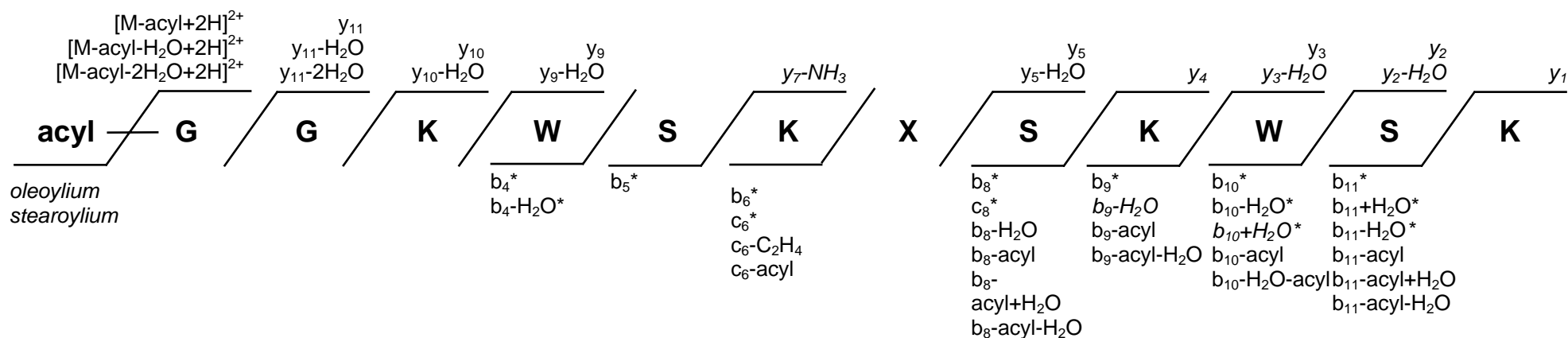
forms the basis of the truncation of Nef₁₋₆ to Nex, was produced as a side reaction on the synthesis column. In the presence of trifluoroacetic acid, commonly used in the elution of peptides from solid-state synthesis columns, the α -amine of the N-terminal tryptophan (Trp) may have become a leaving group in a reaction with the C-terminal lysine of (ole)Nef₁₋₆ and (ste)Nef₁₋₆, with the intermediary possibly being stabilised by the aromaticity of the tryptophan side group. Prior or subsequent to this reaction, the new Trp-7 was converted to Kyn. A small amount of palmitoylated Nex peptide was also detected in the sample at m/z 1707.9 in the MALDI-TOF mass spectrum (see Figure 4.1 and Table 4.1).

Transformation of Trp to Kyn (4 u higher in mass than Trp) has been reported elsewhere,^{51, 307-309} with the Kyn being formed from Trp via oxidative cleavage of the indole ring. This process has also been found to occur posttranslationally in biological systems.³⁰⁷ It is known that chemical treatment of Trp containing peptides and proteins with hydrogen peroxide can cause the formation of Kyn,³⁰⁹ and it is postulated that in this synthetic case, the transformation may have been caused by TFA or some other oxidative species present in the elution mixture. In the ESI-FTICR MS/MS spectrum of Nex 1 and Nex 2, cleavage adjacent to the ester-linked Kyn results in the product ion c_6 . The c_6 ions, with m/z values of 926.6 and 928.6 are the base peak ions in their respective ESI-FTICR MS/MS product ion spectra. The ester linked kynurenine has been denoted X throughout this chapter.

Table 4.5 Summary of the assignment of ESI-FTICR MS/MS results for Nex 1 and Nex 2.

<i>m/z</i>	<i>z</i>	experimental [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	relative intensity (%)	assignment [†]	theoretical [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	error (ppm)
579.342	3	1736.011	33.4	[Nex2+H] ⁺	1736.015	-2.2
578.671	3	1733.998	14.2	[Nex1+H] ⁺	1733.999	-0.6
795.457	2	1589.907	7.3	b ₁₁	1589.910	-1.7
794.451	2	1587.894	2.7	b ₁₁ [*]	1587.894	0.1
751.941	2	1502.873	4.6	b ₁₀	1502.878	-2.9
750.934	2	1500.861	1.9	b ₁₀ [*]	1500.862	-0.8
742.937	2	1484.867	2.5	b ₁₀ -H ₂ O	1484.867	0.1
741.930	2	1482.852	1.0	b ₁₀ -H ₂ O [*]	1482.851	0.6
735.380	2	1469.753	22.1	[Nex-acyl+H] ⁺	1469.754	-0.6
706.870	2	1412.732	12.7	y ₁₁	1412.733	-0.6
678.360	2	1355.711	0.9	y ₁₀	1355.711	0.1
662.329	2	1323.649	3.2	b ₁₁ -acyl	1323.649	0.3
658.903	2	1316.798	3.9	b ₉	1316.798	-0.2
657.896	2	1314.784	1.3	b ₉ [*]	1314.783	0.5
618.813	2	1236.617	2.3	b ₁₀ -acyl	1236.617	0.6
614.311	2	1227.615	6.1	y ₉	1227.616	-1.0
1205.728	1	1205.728	2.2	c ₈	1205.730	-1.6
1203.711	1	1203.711	0.7	c ₈ [*]	1203.714	-2.4
1188.689	1	1188.689	2.7	b ₈	1188.703	-12.0
1186.675	1	1186.675	0.7	b ₈ [*]	1186.688	-10.5
1050.534	1	1050.534	7.6	b ₉ -acyl	1050.537	-2.6
928.621	1	928.621	100.0	c ₆	928.624	-2.9
926.605	1	926.605	37.3	c ₆ [*]	926.608	-2.8
922.440	1	922.440	1.6	b ₈ -acyl	922.442	-2.8
910.611	1	910.611	6.7	b ₆	910.613	-2.7
908.594	1	908.594	1.6	b ₆ [*]	908.597	-3.2
782.517	1	782.517	2.3	b ₅	782.518	-1.3
780.501	1	780.501	0.7	b ₅ [*]	780.502	-2.1
695.485	1	695.485	3.1	b ₄	695.486	-1.4
693.471	1	693.471	1.0	b ₄ [*]	693.470	0.3
420.225	1	420.225	1.7	y ₃ or WSK+H ₂ O SKW+H ₂ O KWS+H ₂ O	420.225	1.6

[†] Nex 1 refers to the Nex peptide with an oleoyl (18:1(9)) residue at the N-terminal glycine, Nex 2 refers to the Nex peptide with a stearoyl (18:0) residue at the N-terminal glycine; * indicates the oleoyl (18:1) form of the product ion, unmarked refers to ions derived from Nex 2 or either form of precursor ion. “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from the product ion. The spectra was internally calibrated based on ions [M+3H]³⁺, y₉, c₆ and their isotopes. A table comprising the complete results can be found in Appendix E.



Scheme 4.5 Annotated sequence of Nex 1 and Nex 2. Plane text indicates cleavage seen in ESI-FTICR MS/MS spectrum. Italicised text indicates additional peaks seen in MALDI- or ESI-Q-o-TOF spectra (see Appendices C and D for assignments). * indicates the presence of product ions with both oleoyl and stearoyl acylations at the N-terminal glycine. See section 4.2.3.1 for a discussion of the identification of the N-terminal acyl component. X represents an ester linked kynurenine residue derived from a tryptophan residue (see section 4.2.3.2). Charge states are omitted for clarity; see Table 4.1 and Appendix E for complete assignments of the ESI-FTICR MS/MS data.

4.2.4 Tryptic Digestion of Nex 1 and Nex 2

In order to confirm the predicted structures of Nex 1 and Nex 2 the sample was digested enzymatically using trypsin. Trypsin was used as the predicted Nex 1/2 sequence contains several lysine residues (trypsin cleaving at the C-terminal side of Lysine residues). The sample was incubated with trypsin overnight at 37°C. The mixture was desalted using a zip-tip™ then extracted with 50% acetonitrile, 0.5% formic acid. The resultant peptide mixture was analysed by ESI-Q-o-TOF MS and MS/MS, using a nanospray source as described previously in Chapter 2.1.1. An automated MS/MS (IDA) experiment was carried out, where peptides with a charge of +2 or +3 were isolated for MS/MS analysis (See Chapter 2.4 for more details). The mass spectrometric results obtained from the tryptic digest are summarised in Table 4.6 and in Scheme 4.6. Scheme 4.6 shows the tryptic peptides derived from acylated Nex 1 and Nex 2 that were observed in the mass spectrum of the trypsin digested synthesis mixture. Sequences underlined in bold indicate tryptic peptides for which tandem spectra were obtained. Sequences underlined with a dotted line indicate tryptic peptides that were observed in the mass spectrum, but for which no tandem mass spectrum was obtained. The tryptic peptide WSK is present at two distinct positions in the proposed Nex 1 and Nex 2 sequences, hence its presence in the digest mix is inconclusive. The ion at m/z 826 corresponds to two possible isometric sequence components, T4-9 and T7-12. The data for the MS/MS of peptides T1-9 and T1-6 for Nex 1 and Nex 2 are shown and discussed below.

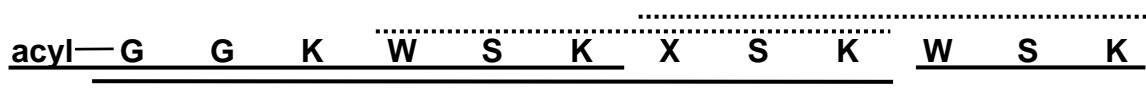
The presence of tryptic peptides T1-3, T1-6 and T1-9 indicate that there is no modification of the ϵ -amino group of lys-3, lys-6 or lys-9, as lysine modified at the ϵ -amino group of the side chain become resistant to tryptic digestion at the adjacent peptide amine bond.¹⁶⁸

Table 4.6 Peptides resulting from the tryptic digestion of Nex 1 and Nex 2.

Experimental [m+H] ⁺ /u	Assignment [†]	Theoretical [m+H] ⁺ /u	error (ppm)	MS/MS obtained
1334.84	Nex 2 T1-9	1334.81	25.6	yes
1332.83	Nex 1 T1-9	1332.79	25.3	yes
928.65	Nex 2 T1-6	928.62	30.1	yes
926.68	Nex 1 T1-6	926.61	76.2	yes
826.42	Nex T7-12; T4-9	826.41	18.1	no
527.41	Nex 2 T1-3	527.42	-7.2	no
525.43	Nex 1 T1-3	525.40	50.2	no
420.24	Nex T4-6; Nex T10-12	420.22	36.2	yes*

[†] T indicates tryptic peptide, with proceeding numbers indicating position in sequence.

*MS/MS spectra were obtained, but are not presented as they did not aid in the sequence determination for the Nex peptides.



Scheme 4.6 Tryptic digestion of Nex 1 and Nex 2. ‘—’ denotes tryptic peptides for which m/z values were observed and additional MS/MS results were obtained ‘. . .’ denotes peptides for which m/z values only were observed; X represents an ester linked kynurenine residue.

The MS/MS results for the analysis of Nex 1 and Nex 2 T1-9 are shown in Figure 4.8 and Table 4.7, with Scheme 4.7 showing the major product ion sequence fragmentation. The sequence ions $y_{1,2}$, y_{6-8} , b_1 , c_6 and b_{7-8} were observed in the spectrum. A number of useful ions were observed. In particular, the presence of c_6 and b_7 ions aid the confirmation of the assignment of the ester-linked kynurenine residue at position-7 in the Nex sequence (as described in section 4.2.3.2 above). A b_7 ion was not observed in the tandem mass spectra of the full-length Nex 1 or Nex 2 peptides; hence its presence in the MS/MS of the tryptic peptide T1-9 is of particular utility in confirming the sequences of Nex 1 and Nex 2.

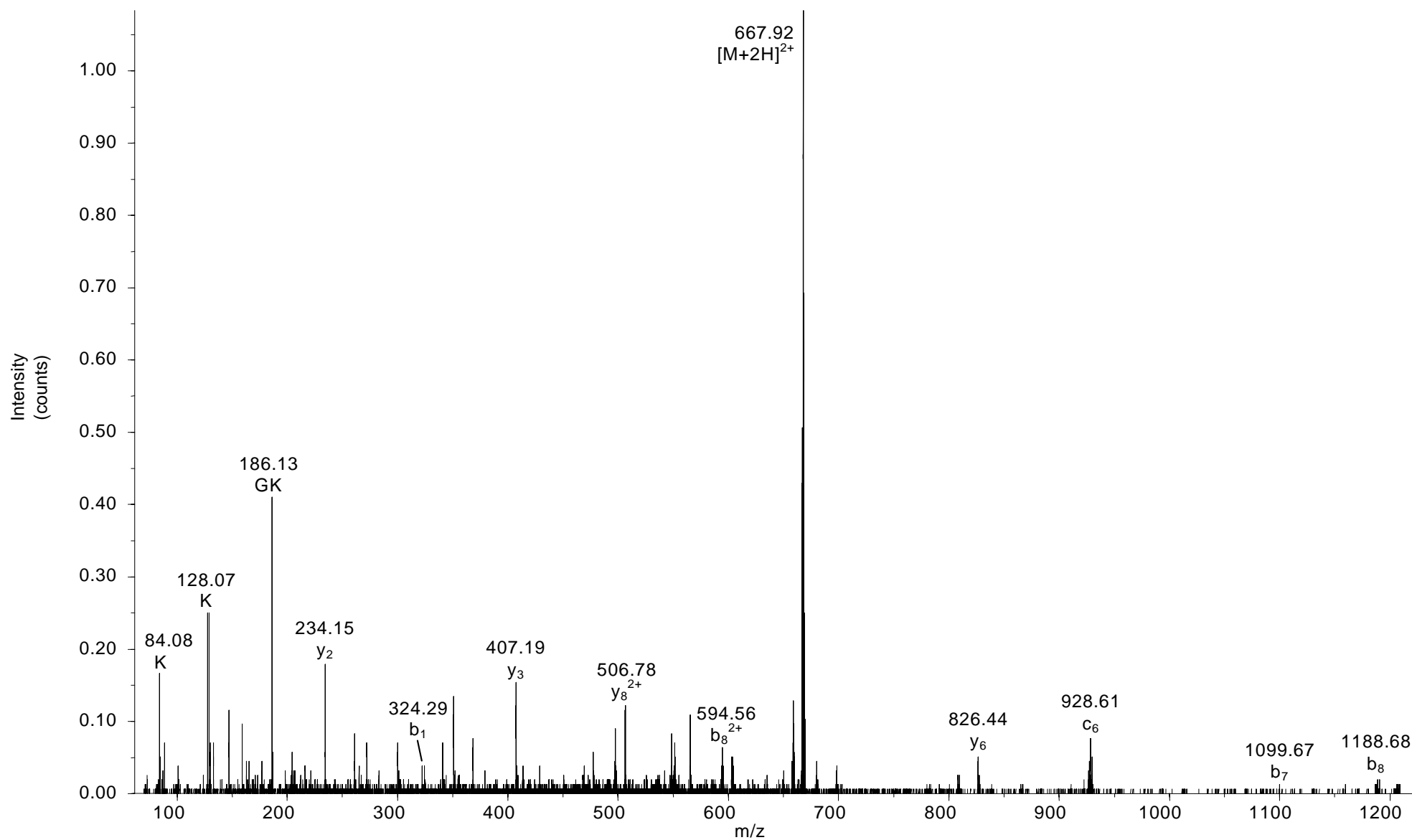
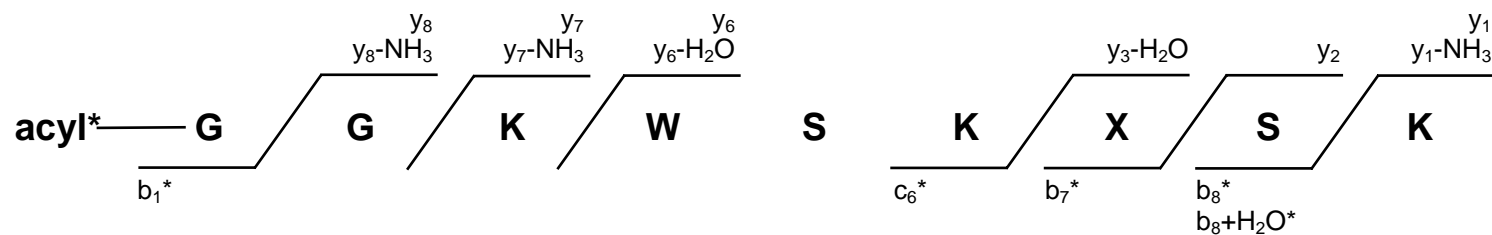


Figure 4.8 ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 tryptic peptides T1-9, m/z 666, low resolution.

Table 4.7 Assignments for ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 combined tryptic peptides T1-9.

<i>m/z</i>	<i>z</i>	Experimental [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	Relative Area (%)	Assignment [†]	Calculated [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	Error (ppm)
667.92	2	1334.84	100.0	[Nex 2] ⁺	1334.81	23
666.93	2	1332.85	42.9	[Nex 1] ⁺	1332.79	41
658.91	2	1316.82	11.9	[Nex 2-H ₂ O+H] ⁺	1316.80	17
657.92	2	1314.83	4.4	[Nex 1-H ₂ O+H] ⁺	1314.78	36
1206.79	1	1206.79	0.7	b ₈ +H ₂ O	1206.71	66
603.87	2	1206.74	3.1	b ₈ +H ₂ O	1206.71	22
602.87	2	1204.73	2.2	b ₈ +H ₂ O*	1204.70	29
1188.74	1	1188.74	0.7	b ₈	1188.70	31
594.86	2	1188.72	2.8	b ₈	1188.70	12
593.85	2	1186.70	1.9	b ₈ *	1186.69	12
551.34	2	1101.67	3.1	b ₇	1101.67	-1
550.33	2	1099.66	2.4	b ₇ *	1099.66	2
506.28	2	1011.55	8.1	y ₈	1011.53	22
497.75	2	994.50	5.3	y ₈ -NH ₃	994.50	-3
477.77	2	954.54	3.0	y ₇	954.50	35
469.26	2	937.52	1.3	y ₇ -NH ₃	937.48	44
928.64	1	928.64	3.6	c ₆	928.62	15
926.42	1	926.62	2.1	c ₆	926.61	13
413.72	2	826.43	1.6	y ₆	826.41	20
826.42	1	826.42	3.5	y ₆	826.41	14
808.41	1	808.41	1.5	y ₆ -H ₂ O <i>or</i> KWSKXS	808.40	15
698.33	1	698.33	1.5	WSKXS+H ₂ O	698.31	27
680.33	1	680.33	1.8	WSKXS	680.30	43
565.29	1	565.29	3.9	WSKX-28	565.28	25
548.30	1	548.30	4.7	KWSK+H ₂ O	548.32	-29
407.20	1	407.20	7.2	y ₃ -H ₂ O <i>or</i> SKX KXS	407.19	10
324.26	1	324.26	1.6	b ₁	324.29	-79
322.24	1	322.24	1.8	b ₁ *	322.27	-123
261.10	1	261.10	2.6	XS-H ₂ O	261.09	40
234.15	1	234.15	6.7	y ₂	234.15	-1
204.12	1	204.12	3.3	GK+H ₂ O	204.13	-90
186.12	1	186.12	11.9	GK	186.12	3
159.10	1	159.10	3.2	W	159.09	61
147.11	1	147.11	4.1	y ₁	147.11	1
133.09	1	133.09	1.8	GK-NH ₃ +2H ₂ O	133.08	89
130.09	1	130.09	2.6	y ₁ -NH ₃	130.09	-4
129.10	1	129.10	7.2	K	129.10	-3
101.11	1	101.11	1.8	K	101.11	-18
84.09	1	84.09	5.8	K	84.08	149

[†] X denotes ester linked kynurine; * indicates the oleoyl (18:1) form of the product ion, as derived from Nex 1, unmarked refers to ions derived from Nex 2 or either form of precursor ion.



Scheme 4.7 MS/MS of tryptic peptide 1-9 from Nex 1 and Nex 2. Results are combined for the MS/MS of the doubly and triply charged species. * indicates the presences of product ions with both oleoyl and stearoyl acylations at the N-terminal glycine. See section 4.2.3.1 for a discussion of the identification of the N-terminal acyl component; X represents an ester linked kynurenine residue (see section 4.2.3.2). Charge states are omitted for clarity; see Table 4.7 for complete assignments.

The MS/MS results for the analysis of Nex 1 and Nex 2 T1-6 are shown in Figure 4.9 and Table 4.8, with Scheme 4.8 showing the major product ion sequence fragmentation. The sequence ions y_{1-3} , y_5 and b_1 were observed. The MS/MS spectrum for simultaneously fragmented Nex 1 and Nex 2 T1-6 is very similar to that of $[(\text{pam})\text{Nef}_{1-6} + 2\text{H}]^{2+}$, m/z 450.8 (see inset), indicating that their sequences are analogous. The predominant differences between the spectra of $(\text{pam})\text{Nef}_{1-6}$, and Nex 1 and Nex 2, is the presence of a palmitoyl carbenium ion at m/z 239.2 in the $(\text{pam})\text{Nef}_{1-6}$ spectrum, and the presence of oleoylium (m/z 265.3) and stearoylium (m/z 267.3) ions in the Nex 1 and Nex 2 T1-6 spectrum. This result aids considerably in confirming the identity of the six N-terminal amino acids of the Nex 1 and Nex 2 sequences.

The tryptic digestion of the Nex 1 and Nex 2 peptides was of use in confirming the sequence that had been predicted from the intact MS/MS of these two peptides. In particular, the MS/MS of the T1-9 sequence supported the assignment of an ester-linked kynurenine at amino acid position-7. Similarity of the T1-6 MS/MS spectrum and the $(\text{pam})\text{Nef}_{1-6}$ spectrum supported the proposed sequence for the six N-terminal amino acid residues of Nex 1 and Nex 2.

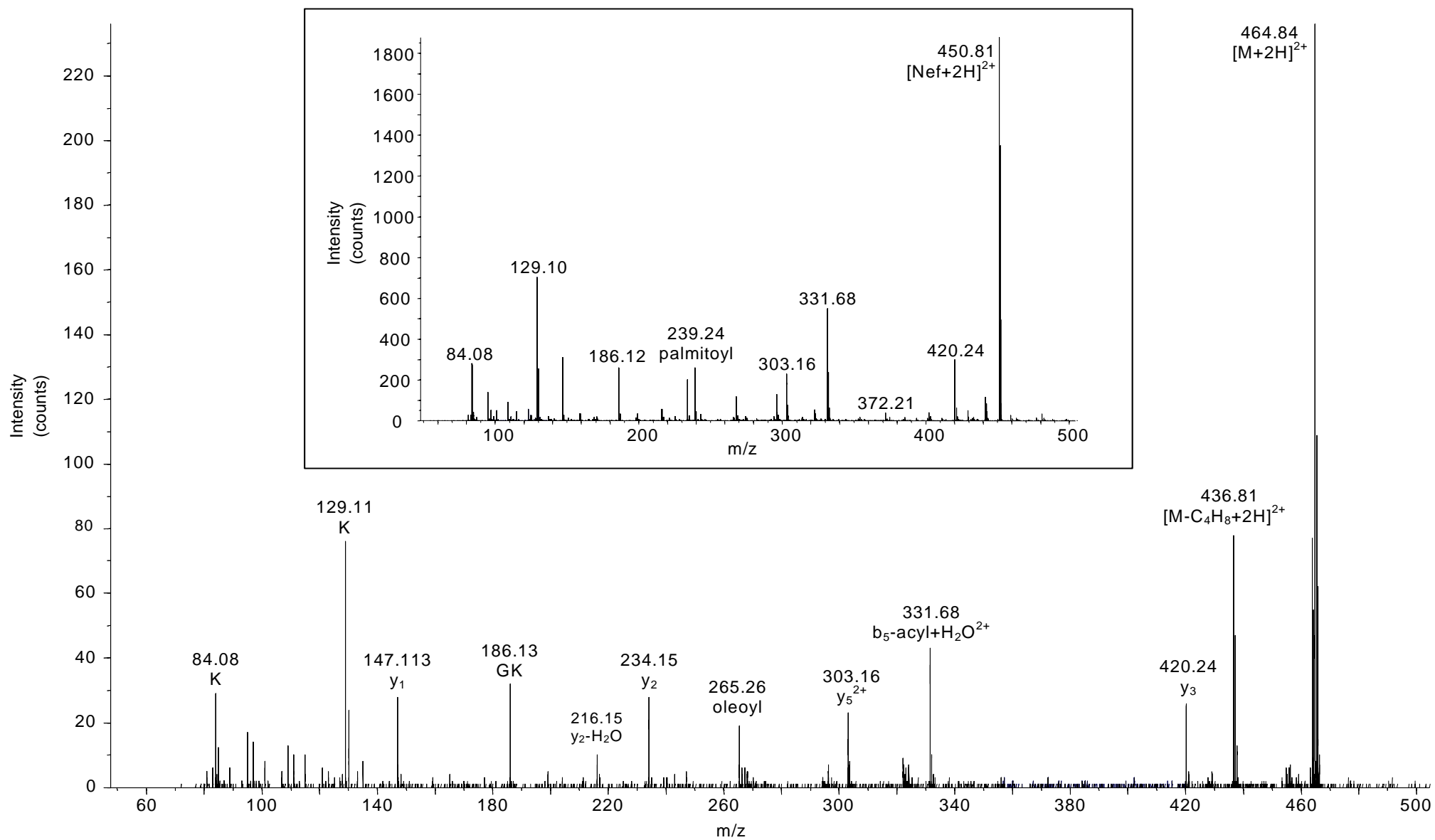
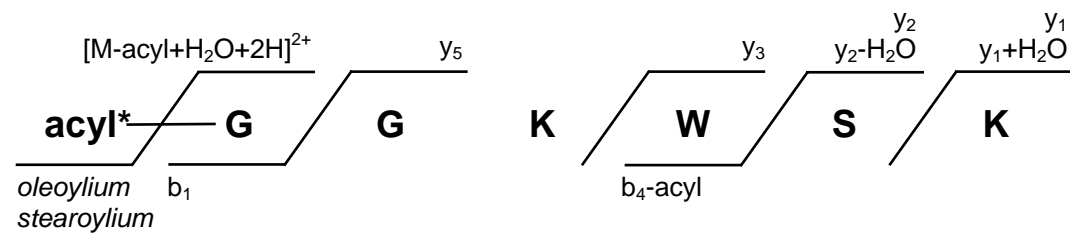


Figure 4.9. ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 $[T1-6+2H]^{2+}$, m/z 463/465. Inset ESI-Q-o-TOF MS/MS of $[Nef_{1-6} + 2H]^{2+}$, m/z 450.

Table 4.8 Assignments for ESI-Q-o-TOF MS/MS of Nex 1 and Nex 2 combined tryptic peptides T1-6.

<i>m/z</i>	<i>z</i>	Experimental [<i>m</i> + <i>H</i>] [†] / <i>u</i>	Relative Area (%)	Assignment [†]	Calculated [<i>m</i> + <i>H</i>] [†] / <i>u</i>	Error (ppm)
464.84	2	928.67	100.0	[Nex 2] ⁺	928.624	53
463.83	2	926.66	32.6	[Nex 1] ⁺	926.608	56
455.83	2	910.66	4.4	[Nex 2-H ₂ O+H] ⁺	910.613	50
454.81	2	908.61	1.3	[Nex 1-H ₂ O+H] ⁺	908.597	19
436.81	2	872.61	31.3	[M-C ₄ H ₈ +H] ⁺	872.561	60
331.69	2	662.37	15.7	[Nex-acyl+H ₂ O+H] ⁺	662.360	13
303.18	2	605.35	11.1	<i>y</i> ₅	605.340	25
605.47	1	605.47	1.7	<i>y</i> ₅	605.340	214
429.25	1	429.25	1.5	<i>b</i> ₄ -acyl	429.220	64
420.24	1	420.24	7.0	<i>y</i> ₃	420.220	45
324.28	1	324.28	2.4	<i>b</i> ₁	324.290	-30
296.32	1	296.32	3.0	G(ste) immonium	296.295	75
267.28	1	267.28	3.4	stearoyl	267.279	5
265.27	1	265.27	10.8	oleoyl	265.263	33
234.15	1	234.15	6.9	<i>y</i> ₂	234.150	9
216.17	1	216.17	2.5	<i>y</i> ₂ -H ₂ O	216.130	182
186.14	1	186.14	7.8	GK	186.120	99
165.13	1	165.13	1.0	<i>y</i> ₁ +H ₂ O	165.121	42
147.12	1	147.12	9.3	<i>y</i> ₁	147.110	70
129.10	1	129.10	23.5	K	129.100	42
84.08	1	84.08	9.4	K	84.080	-1

[†] “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from the product ion.



Scheme 4.8 MS/MS of tryptic peptide T1-6 from Nex 1 and Nex 2. Results are combined for the MS/MS of the doubly and triply charged species. * indicates the presences of product ions with both oleoyl and stearoyl acylations at the N-terminal glycine. See section 4.2.3.1 for a discussion of the identification of the N-terminal acyl component. Charge states are omitted for clarity; see Table 4.8 for complete assignments.

4.3 Conclusions

Sequences for the previously unidentified immunogenic peptides Nex 1 and Nex 2 have been successfully proposed. The peptides were found to be modified by an oleoyl moiety in the case of Nex 1, and by a stearoyl moiety in the case of Nex 2, at an N-terminal glycine residue. The first six N-terminal amino acid residues of the Nex peptides were found to be consistent with the sequence of Nef₁₋₆, a fragment of an HIV protein. The Nex peptides were found to have six additional residues at the C-terminal end. The initial link to the additional residues was found to be via an ester linkage. Also, a modified tryptophan residue was identified, where the tryptophan residue had undergone oxidative cleavage of its indole ring, resulting in the formation of a kynurenine residue. The sequence for the acylated Nex peptides is proposed to be G(ste/ole)GKWSK-O-kyn-SKWSK.

Knowledge gained in Chapter 3 of this thesis, regarding the nature of collisionally induced dissociation of acylated peptides, aided significantly in the elucidation of the sequences of Nex 1 and Nex 2. In particular, the presence of b₁ acylated product ions and the absence of acylated immonium (or similar) ions, enabled the identification of the oleoyl and stearoyl moieties. Additionally, the use of a variety of analysis techniques allowed for the confirmation of the sequence assignments, where one technique on its own did not provide sufficient information to confirm the structure of this unusual synthetic peptide.

Subsequent to this work our collaborators were able to synthesis a Nex 2 analogue, with a point substitution at the ester linked kynurenine, which was replaced by an amine linked tryptophan residue, giving the sequence G(ste)GKWSKWSKWSK. This peptide was also found to elicit a T cell response when presented by CD1c. The corresponding

unacylated peptide did not elicit a T cell response.¹⁷⁶ The characterisation of Nex 1 and Nex 2 has allowed the discovery of a novel class of T cell antigens. As well as having potential for vaccine development, these antigens could be useful for modulating T cell activity, or be used as a diagnostic tool in evaluating immune responses.³¹⁰

5 Characterisation of Monodisperse Polymers by Mass Spectrometry

5.1 Introduction

5.1.1 Monodisperse Polymers

Monodisperse polymers are polymers where all molecules in a sample have the same mass and functionality/functional groups. Traditional methods of polymer synthesis, such as the production of polyethylene glycol, typically result in a non-homogenous product, with many products of varying masses being produced.¹⁶ Monodisperse polymers can reliably be produced in a semi-synthetic way by reacting a functional group with the various reactive residues of a protein, thus using the protein as an amine-based backbone for polymer synthesis.¹⁷ Semisynthetic modified proteins allow for the control of molecular weight and functionality, and may be useful where well-defined macromolecular structures are desired, for example, as therapeutic agents.^{311, 312} They may also be used where controlled chemical-physical properties are required, for example, in molecular weight standards, or for biological or spectroscopic studies.¹⁷

Functionalised proteins can be produced in such a way as to modify specific amino acid residues in a protein (such as all exposed lysine residues) or by using site-specific labelling to modify one or more specific residues within a protein molecule. The complete (or per-) functionalisation of a protein can be achieved by first denaturing the protein of interest using an agent such as SDS.¹⁷ Site-specific functionalisation has been performed by Broo and co-workers who demonstrated the ligand-directed labelling of a single tyrosine^{313, 314} and a single mutant lysine residue within the human glutathione transferase A1-1 protein with a range of acyl groups.⁵⁹ An alternative option is to produce novel proteins that mimic the native state of natural proteins, as suggested by Raleigh *et. al.*³¹⁵

In this chapter two proteins of different size, ubiquitin (8.6 kDa) and bovine carbonic anhydrase II (BCA, 29 kDa), were acylated with a selection of functional groups at their available amines. The range of acyl functional groups used is shown in Table 5.1.

Table 5.1 Acyl groups used to perfunctionalise ubiquitin and bovine carbonic anhydrase.

Functional Group	Formula	Delta Mass ^a (monoisotopic/u)	Modifying Agent ^b
Acetyl	CH ₃ CO-	42.011	(CH ₃ CO) ₂ O
Benzoyl	C ₆ H ₅ CO-	104.026	C ₆ H ₅ COCl
Trifluoropropionoyl	CF ₃ CH ₂ CO-	110.000	CF ₃ CH ₂ CO-(NHS)
Glutaryl	COOH(CH ₂) ₃ CO-	114.032	$\text{CO(CH}_2\text{)}_3\text{COO}$ └──────────┘
Iodoacetyl	ICH ₂ CO-	167.907	ICH ₂ CO-(NHS)
PEG ^c	CH ₃ O(CH ₂ CH ₂ O) ₃ CH ₂ CH ₂ CO-	218.115	CH ₃ O(CH ₂ CH ₂ O) ₃ CH ₂ CH ₂ CO-(NHS)

^a In calculating the delta mass it is assumed that there is the abstraction of one hydrogen from the modified amino acid residue. Delta mass is given per functional unit.

^b NHS: *N*-hydroxysuccinimide.

^c For the purposes of this study the specific molecule defined by the formula given in this table (methoxypolyethylene glycol propionic acid NHS ester) will be referred to as PEG.

5.1.2 Characterisation of Monodisperse Polymers

In developing a novel class of molecules it is important to be able to characterise the products of the functionalisation, to confirm that all targeted reaction sites are functionalised and the product is monodisperse in nature. Mass spectrometry can be used to characterise these molecules, confirming the homogeneity in molecular weight, and the residues in the protein at which functional groups have been added.

There are a limited number of standard analytical techniques that can be employed for the analysis of synthetically modified peptides and proteins. Available methods include gel electrophoresis,³¹⁶ capillary electrophoresis³¹⁷ and a range of mass spectrometry

techniques.¹⁵⁵ Gel electrophoresis separates proteins primarily on the basis of mass. Therefore, dependant on the mass of the protein modifying agent, the extent of modification can be approximated from the intact mass of the modified protein. Capillary electrophoresis separates based on charge. This allows proteins that have been modified on an otherwise charge holding amino acid (such as lysine), to be characterised by comparison with a pre-run charge ladder for the same protein. Both gel electrophoresis and capillary electrophoresis have the advantages of being both relatively inexpensive to perform and easily portable.

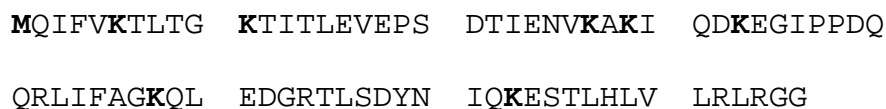
Mass spectrometry, involves expensive instrumentation and generally a dedicated laboratory. Mass spectrometry, however, provides data that enables accurate characterisation of the modified groups on a modified peptide or protein. In addition, mass spectrometric analysis is not reliant on the presence of specific charge holding amino acid side-chains. Mass spectrometry can be coupled with liquid chromatography, allowing analysis of smaller peptides resulting from enzymatic digestion of the intact modified protein enabling confirmation of the functionalised sites within the protein.

The aim of this chapter is to characterise a number of monodisperse polymers. The experience in analysing acylated peptides gained in Chapter 3 of this thesis will be applied and extended, by analysing a range of synthetic protein acylations utilised in producing monodisperse polymers. Specifically, the proteins will be examined in terms of the sites at which acylation has occurred and to identify any spurious modifications. This analysis was completed in parallel with collaborators who used capillary electrophoresis to analyse the same polymers.

5.2 Results and Discussion

5.2.1 Analysis of Peracylated Ubiquitin by Mass Spectrometry

Ubiquitin was selected for the synthesis and study of monodisperse polymers as an example of a relatively low molecular weight protein (8.6 kDa) containing eight amine sites available for peracylation under the reaction conditions employed (See Chapter 2.5). These sites are found at seven lysine residues and the N-terminal amine. The sequence for native ubiquitin is given in Scheme 5.1.



Scheme 5.1 Amino acid sequence for Ubiquitin (human and bovine). (Lysines and the N-terminal methionine are indicated in bold.)

Ubiquitin was peracylated with acetyl, benzoyl, glutaryl, iodoacetyl and PEG as shown in Table 5.1 above, by our collaborators Jerry Yang *et. al.*¹⁷ A number of the peracylated ubiquitin peptides were analysed by MALDI-TOF MS to determine their intact masses. Each of the peracylated ubiquitin polymers were subsequently digested enzymatically and analysed by HPLC-ESI-Q-o-TOF MS/MS in order to confirm the sites of acylation.

5.2.1.1 Intact Mass Analysis of Acylated Ubiquitin by MALDI-TOF MS

Initial intact mass analysis of the functionalised ubiquitin involved the MALDI-TOF MS analysis of acetylated ubiquitin. Capillary electrophoresis of the acetylated ubiquitin had indicated the presence of one product; the peracetylated product containing eight acetyl groups, as determined by comparison to a charge ladder. Mass spectrometry of the acetylated ubiquitin, however, indicated the presence of a small amount of an additional acetyl group (see Figure 5.1 part a)). Given the presence of a tyrosine residue at position 59 and three serine residues (positions 20, 57 and 65) containing phenolic hydroxyl, and hydroxyl groups respectively, it was possible that any of these could be an additional site of acetylation. Treatment of the acetylated ubiquitin with lithium hydroxide (100 μ M perfunctionalised ubiquitin in 0.1 M LiOH for 1 h at 4°C, followed by dialysis against water) resulted in a significant reduction of the additional acetate adduct, as shown in Figure 5.1 part b). Subsequent samples of peracylated ubiquitin and bovine carbonic anhydrase were pre-treated with lithium hydroxide to remove any spurious additions at hydroxyl residues within the protein molecule prior to analysis.

Analysis of three of the perfunctionalised ubiquitin molecules, along with the native ubiquitin, were performed using MALDI-TOF MS. The resultant mass spectra are shown in Figure 5.2. The results are summarised in Table 5.2 below.

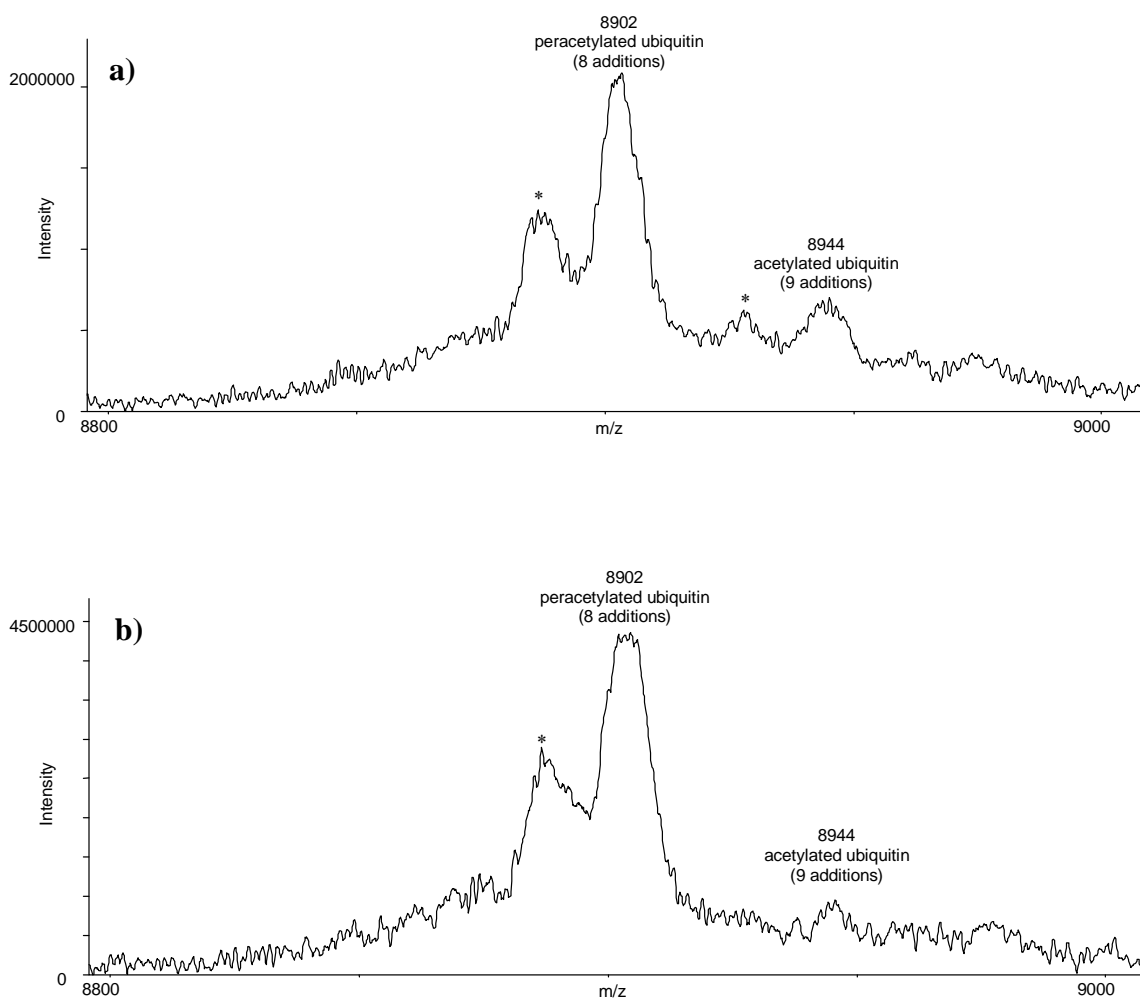


Figure 5.1 MALDI-TOF MS analysis of peracetylated ubiquitin prior to treatment with lithium hydroxide a) and after lithium hydroxide treatment b). (Reflectron mode used. * indicates peaks resulting from metastable decay.)

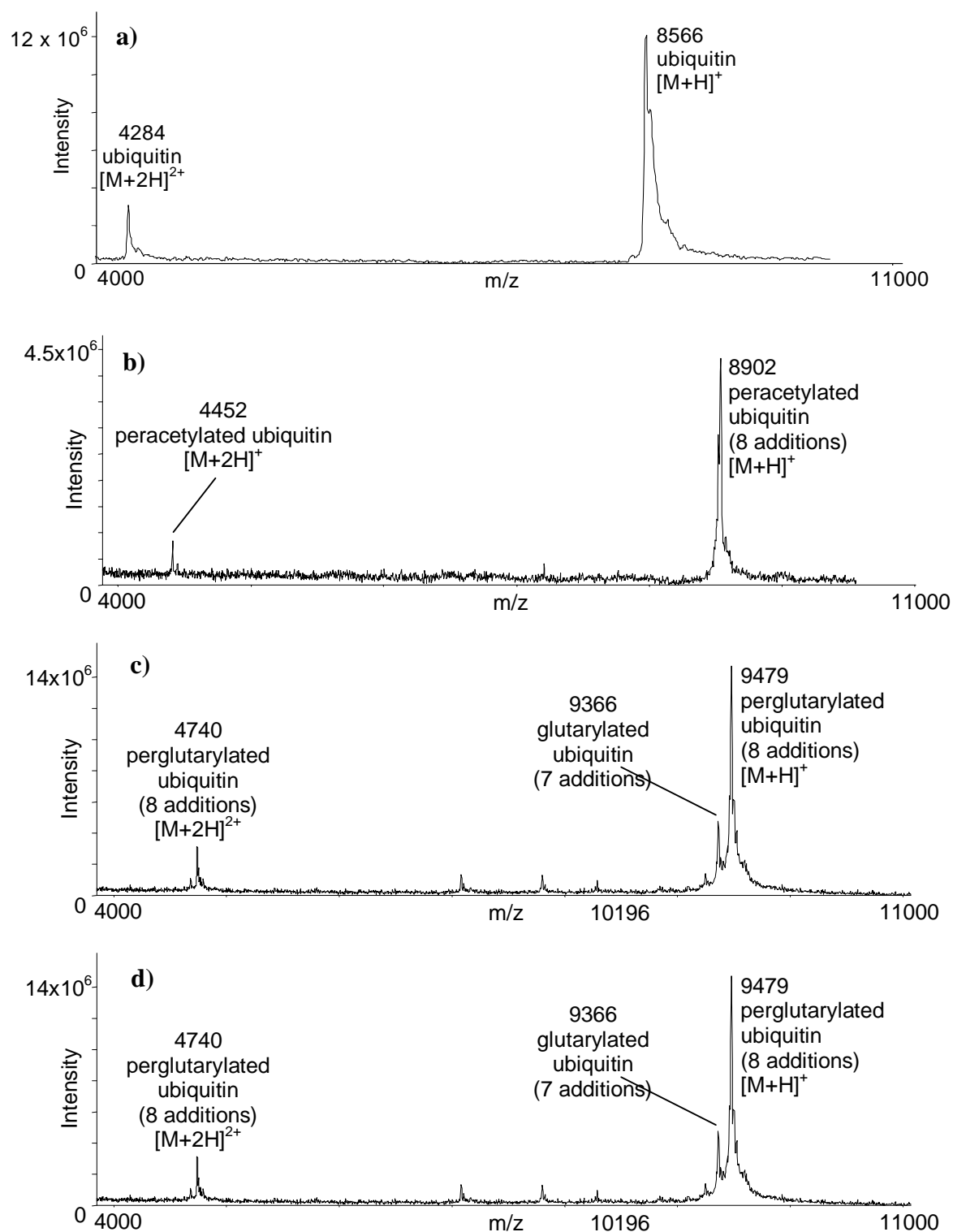


Figure 5.2 MALDI-TOF MS analysis of native ubiquitin a), peracetylated ubiquitin b), perglutarylated ubiquitin c) and perPEGylated ubiquitin d). (Note: the MS of ubiquitin was acquired in linear mode, with the perfunctionalised molecules being acquired in reflectron mode.)

Table 5.2 Summary of Intact MS observed for functionalised ubiquitin.

Functionalised ubiquitin	Experimental Mass [m+H]⁺ (average/u)	Theoretical Mass [m+H]⁺ (average/u)	Error (ppm)
nil (native)	8565.76	8565.92	-19
peracetylated (8 additions)	8902.58	8902.22	41
9 acetyl additions	8944.52	8944.25	30
perglutarylated (8 additions)	9478.96	9478.72	25
7 glutarate additions	9365.53	9364.62	97
perPEGylated (8 additions)	10312.56	10311.94	60
7 PEG additions	10093.65	10093.69	-4
7 PEG additions + 1 truncated PEG	10195.69	10195.74	-5
6 PEG additions + 1 truncated PEG	9976.90	9977.49	-59

Each of the MALDI-TOF MS spectra shows a clear molecular ion for the expected mass of the respective perfunctionalised ubiquitin molecules. A small shoulder to the lower mass side of the spectra acquired in reflectron mode can be observed, which is caused by metastable decay of the protein molecules. Both the perglutarylated and perPEGylated spectra show a peak due to the addition of one less functional group to the ubiquitin backbone (m/z 9365 and m/z 10094 respectively) than in the perfunctionalised molecule. The additional peaks observed in the perPEGylated ubiquitin spectra correspond with the substitution of one PEG molecule (as defined in Table 5.1) with a truncated form of PEG with a possible molecular formula of $C_4H_7O_3$ (delta mass 102.032 u, monoisotopic; 102.050 u, average). Based on the peak heights observed in the MALDI-TOF MS it appears that approximately 45 % of the PEGylated species contain one truncated PEG molecule and approximately 40 % of the molecules are missing one PEG moiety. There may be some PEGylated ubiquitin molecules containing more than one truncated PEG molecule, however this is not clear from the

MALDI-TOF MS data, as the position of the expected mass for the presence of two truncated molecules approximately corresponds with the position of the metastable decay peak for the PEGylated ubiquitin containing seven PEG groups. (The appearance of four distinct peaks for PEGylated ubiquitin was also observed in the linear MALDI-TOF MS spectrum, data not shown, indicating that the phenomenon is not a result of the MALDI-reflectron detection process.) The identification of possible specific lysine residues, either missing a functional group or containing a truncated PEG, is discussed in sections 5.2.1.3 and 5.2.1.4 below.

These results clearly indicate that capillary electrophoresis is not reliable in determining the number and type of functionalised residues within a given molecule, given the conditions used by our collaborators.¹⁷ For example, the presence of a truncated form of PEG is not discerned using capillary electrophoresis where charge is the only means of determining the addition of a functional group to a lysine residue.

5.2.1.2 HPLC-ESI-Q-o-TOF MS/MS Analysis of Enzymatically Digested Peracetylated Ubiquitin

Peracetylated, perbenzoylated, perglutarated, periodated and perPEGylated ubiquitin were enzymatically digested using GluC and AspN enzymes. GluC cleaves C-terminal to glutamic acid (E) and aspartic acid (D) residues. The enzyme AspN cleaves at the N-terminal side of aspartic acid (D) residues. (The enzymes trypsin and LysC, both commonly used enzymes, were not appropriate for this study due to the functionalisation of the lysine residues; trypsin cleaving at lysine and arginine and LysC at lysine. The presence of only two arginine residues, both proximal to the C-termini of ubiquitin limited the utility of the trypsin enzyme.) Samples were digested and subsequently diluted to an appropriate concentration using Buffer A and analysed by HPLC-ESI-Q-o-TOF MS/MS, as described in Chapter 2. Results were obtained using

Intelligent Data AcquisitionTM (IDATM; Applied Biosystems), with peptides of 2+ or 3+ charge state in the range m/z 350 – 1500 being selected for analysis. An example of the total ion current (TIC) produced as a result of the HPLC-ESI-Q-O-TOF MS/MS analysis, along with the combined mass spectra for the peptide elution is given in Figure 5.3. The example given is for perPEGylated ubiquitin digested with AspN. The majority of the enzymatic peptides eluted between 29 and 37 minutes, with a 6-mer eluting earlier, at approximately 26 minutes and a 20-mer at 55 minutes. The peptides eluting between 29 and 27 minutes were predominately ten to twelve amino acids in length. A sub-stoichiometric addition of PEG was observed for AspN peptide 21-31 which is discussed further in section 5.2.1.3. Unlabelled peaks are due to peptides produced by missed-cleavages and/or degradation cleavage products.

The enzymatic peptides observed in the HPLC-ESI-Q-O-TOF MS/MS analysis of GluC digested peracylated ubiquitin are given in Table 5.3, and resultant sequence coverage maps are shown in Scheme 5.2. The GluC digestion produced a number of useful peptides containing modified lysine residues. Most of the enzymatic peptide ions observed were of a 2+ or 3+ charge state, with only a few being singly charged. The majority of the peptides produced were successfully identified and fragmented using the IDATM software, as described in Chapter 2.4. A range of acylating agents were used and comparative MS/MS spectra for the various functional groups are shown in Figure 5.4 - 5.6. The theoretical masses for the enzymatic peptides were determined using Protein Prospector as described in Chapter 2.6,¹⁸¹ assignments for the spectra obtained are discussed below. The MS/MS spectra obtained covered seven of the eight acylated sites on the ubiquitin peptide. An m/z value, was observed for peptide containing the K(glutarate)-63 adduct, however the peptide was not analysed by MS/MS.

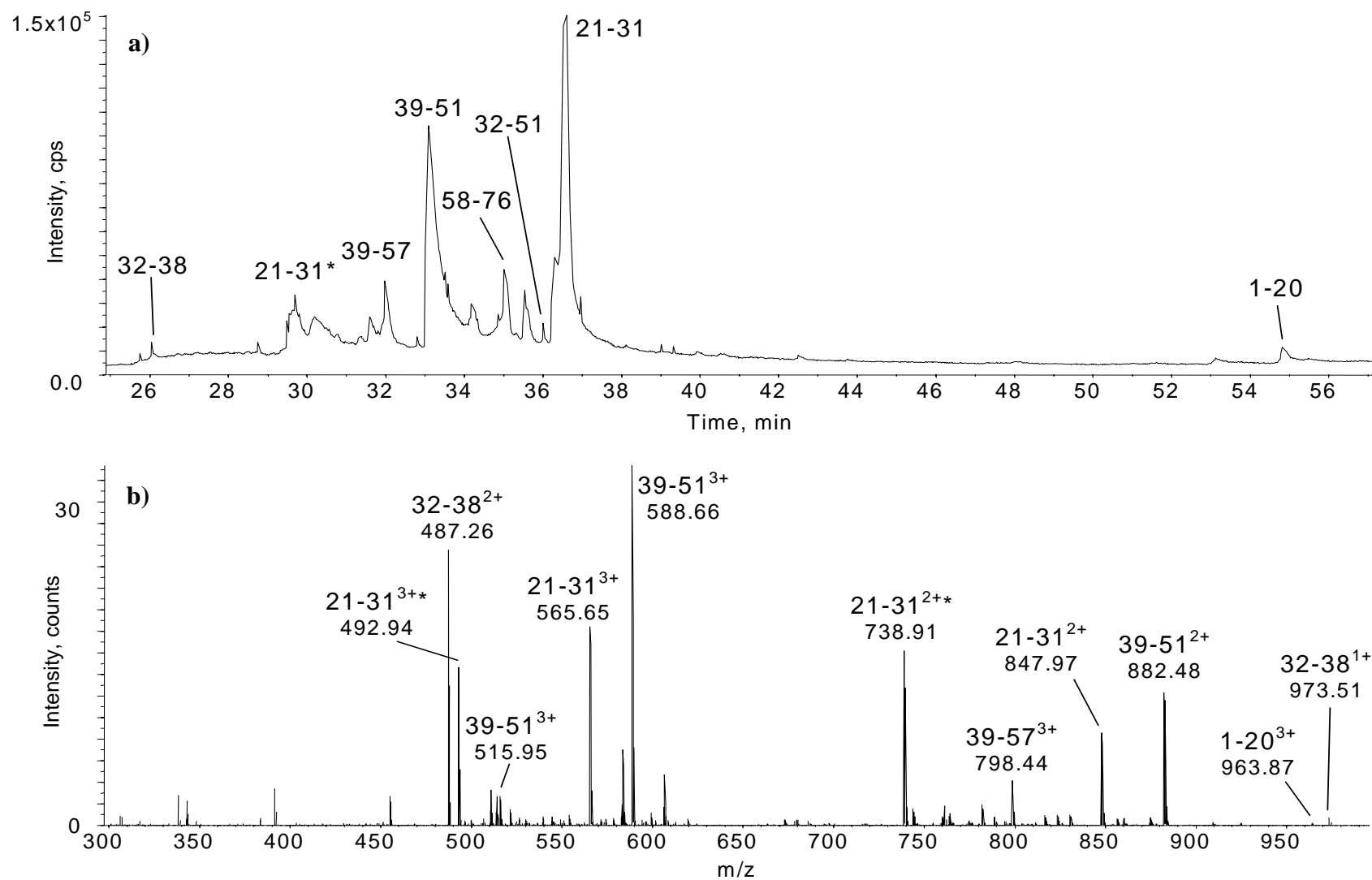
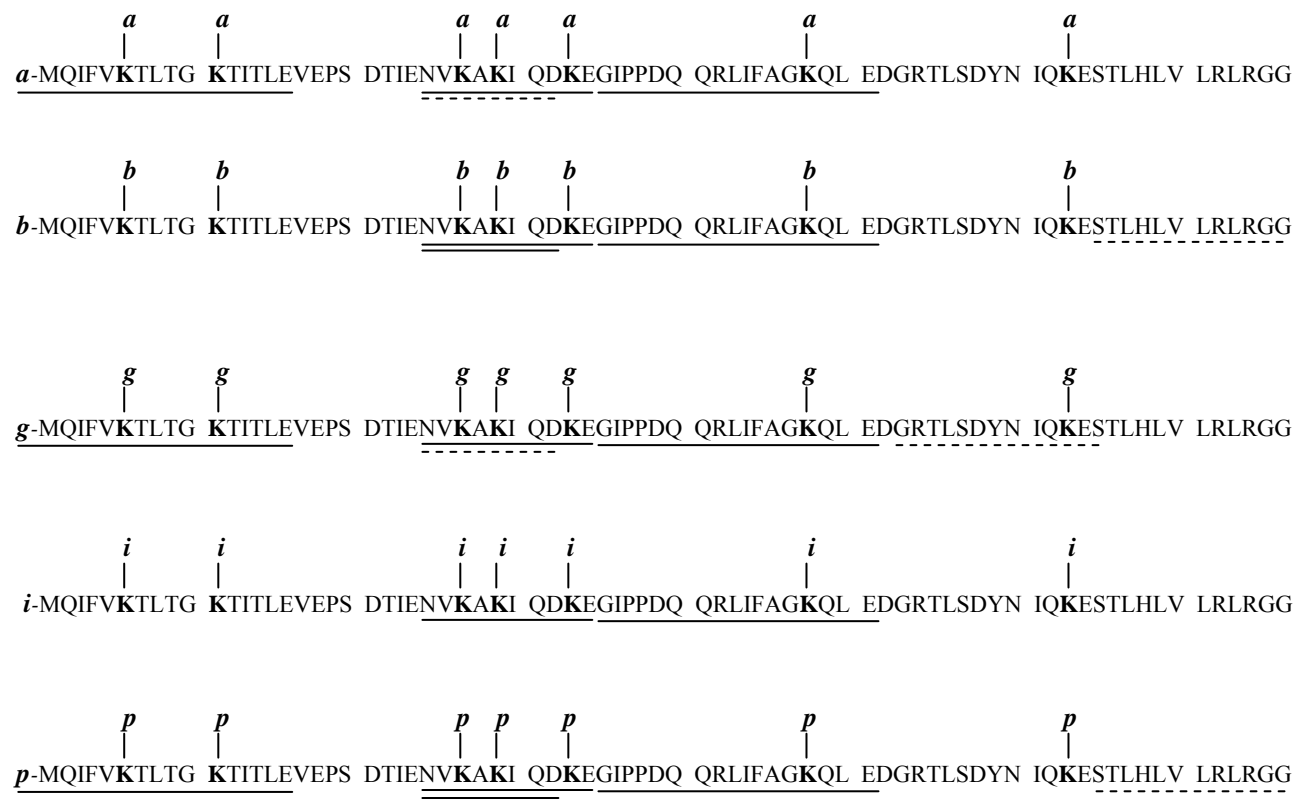


Figure 5.3 HPLC-ESI-Q-o-TOF MS/MS analysis of AspN digested perPEGylated ubiquitin: TIC a), combined mass spectrum for the peptide analysis b). Numerals indicate the sequence position of the AspN produced peptides. (* indicates sub-stoichiometric addition of PEG, discussed in section 5.2.1.3. below.)



Scheme 5.2 Sequence coverage maps for the GluC digestion of perfunctionalised ubiquitin. Where a indicates acetate; b indicates benzoate; g indicates glutarate; i indicates iodoacetate and p indicates PEG addition. Peptides analysed by MS/MS are indicated by an unbroken line; peptides detected by MS only are indicated by a dashed line.

Table 5.3 Peptide assignments for the GluC enzymatic digest of perfunctionalised ubiquitin.

m/z ^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment	Calculated [m+H] ⁺ /u	Error (ppm)
Peracetylated Ubiquitin						
500.278	2	999.548	26.0	25-32	999.547	1
1298.714	1	1298.714	30.5	25-34	1298.695	15
649.854	2	1298.700	30.5	25-34	1298.695	3
433.576	3	1298.713	30.5	25-34	1298.695	14
975.010	2	1949.012	55.5	1-16	1949.066	-28
650.362	3	1949.069	55.5	1-16	1949.066	2
976.530	2	1952.052	34.8	35-51	1952.060	-4
651.367	3	1952.084	34.8	35-51	1952.060	12
Perbenzoylated Ubiquitin						
562.292	2	1123.576	37.1	25-32	1123.578	-2
441.279	3	1321.821	34.5	65-76	1321.806	11
742.869	2	1484.731	43.8	25-34	1484.742	-7
495.587	3	1484.746	43.8	25-34	1484.742	3
1007.536	2	2014.064	38.3	35-51	2014.076	-6
672.031	3	2014.076	38.3	35-51	2014.076	0
Perglutarylated Ubiquitin						
572.300	2	1143.593	27.8	25-32	1143.589	3
1514.750	1	1514.750	32.2	25-34	1514.758	-5
757.878	2	1514.749	32.2	25-34	1514.758	-6
505.594	3	1514.768	32.2	25-34	1514.758	6
769.381	2	1537.754	27.4	53-64	1537.749	3
1012.526	2	2024.044	35.4	35-51	2024.081	-18
675.368	3	2024.088	35.4	35-51	2024.081	3
1083.060	2	2165.111	55.8	1-16	2165.130	-9
722.394	3	2165.167	55.8	1-16	2165.130	17
Periodoacetylated Ubiquitin						
838.693	2	1676.377	37.5	25-34	1676.385	-4
1039.476	2	2077.944	37.3	35-51	2077.957	-6
693.331	3	2077.979	37.3	35-51	2077.957	11
PerPEGylated Ubiquitin						
661.371	2	1321.735	38.4	65-76	1321.806	-54
676.380	2	1351.752	33.0	25-32	1351.757	-3
847.975	2	1694.942	35.2	22-32	1694.931	6
565.658	3	1694.958	35.2	22-32	1694.931	16
914.014	2	1827.021	36.3	25-34	1827.010	6
609.680	3	1827.024	36.3	25-34	1827.010	8
1064.593	2	2128.179	36.5	35-51	2128.165	7
710.068	3	2128.189	36.5	35-51	2128.165	11
1239.192	2	2477.376	58.2	1-16	2477.381	-2
826.469	3	2477.392	58.2	1-16	2477.381	5

^a Bold-face type indicates peptides for which MS/MS data was obtained.

Figure 5.4 shows the MS/MS spectra for peracetylated, perglutarylated and perPEGylated ubiquitin GluC enzymatic peptide 1-16. These peptides contain an

acylated N-terminal methionine and two acylated lysine residues. The MS/MS spectra obtained show very similar fragmentation for the three acylating agents used. This is in accord with the results observed in Chapter 3, where very similar fragmentation patterns were observed for peptides modified with varying acyl chain lengths. The presence of an acylated N-terminal methionine is confirmed by the presence of a b_1 ion in each of the fragmentation spectra. A b_1 ion is not normally observed unless there is a modification to the N-terminal amino acid allowing for the formation of a 5-membered ring and consequent fragmentation.²³⁸ Similar b_1 ions were observed in Chapter 3, when N-terminally acylated peptides were fragmented and in Chapter 4 for the fragmentation of Nef₁₋₆, Nex 1 and Nex 2.

For each of the acylated peptides fragmented a corresponding acylated lysine carbenium ion²³⁵ was observed, being K(acetate) at m/z 126.09; K(glutarate) at m/z 198.12 and K(PEG) at m/z 302.21 (which corresponds to 17 u less than an acylated lysine immonium ion, resulting from the neutral loss of NH₃). There is a corresponding absence of any lysine carbenium ion at m/z 84 or amino-acylium ion at m/z 129, as normally expected in the fragmentation spectrum of a peptide containing lysine residues.³¹⁸ These observations are in agreement with Kim *et. al.* who found that the acetylated lysine carbenium ion at m/z 126 was a more useful marker ion than one at m/z 143, which corresponds to an acetylated lysine immonium ion.³¹⁹ The elimination of ammonium from the immonium ion has also been observed for several amino acids, including native lysine, by Dookeran *et. al.*³²⁰ In this present study there were no appreciable ions observed at m/z 143 corresponding to an acetylated lysine immonium ion that could not otherwise be accounted for; nor were acylated lysine immonium ions observed for any of the other acyl moieties studied.

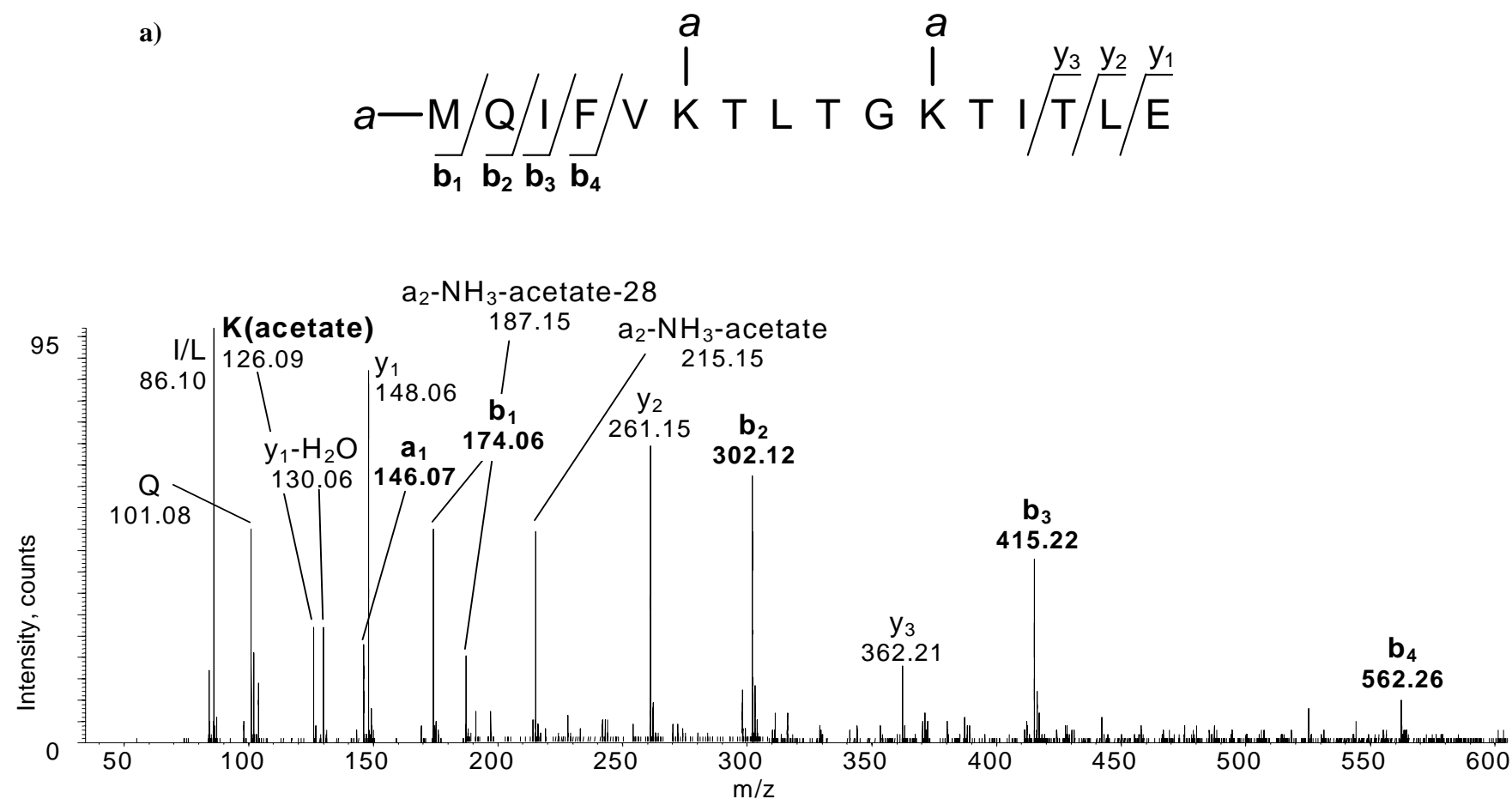


Figure 5.4 HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 1-16. Peracetylated ubiquitin a). *a* indicates acetyl group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

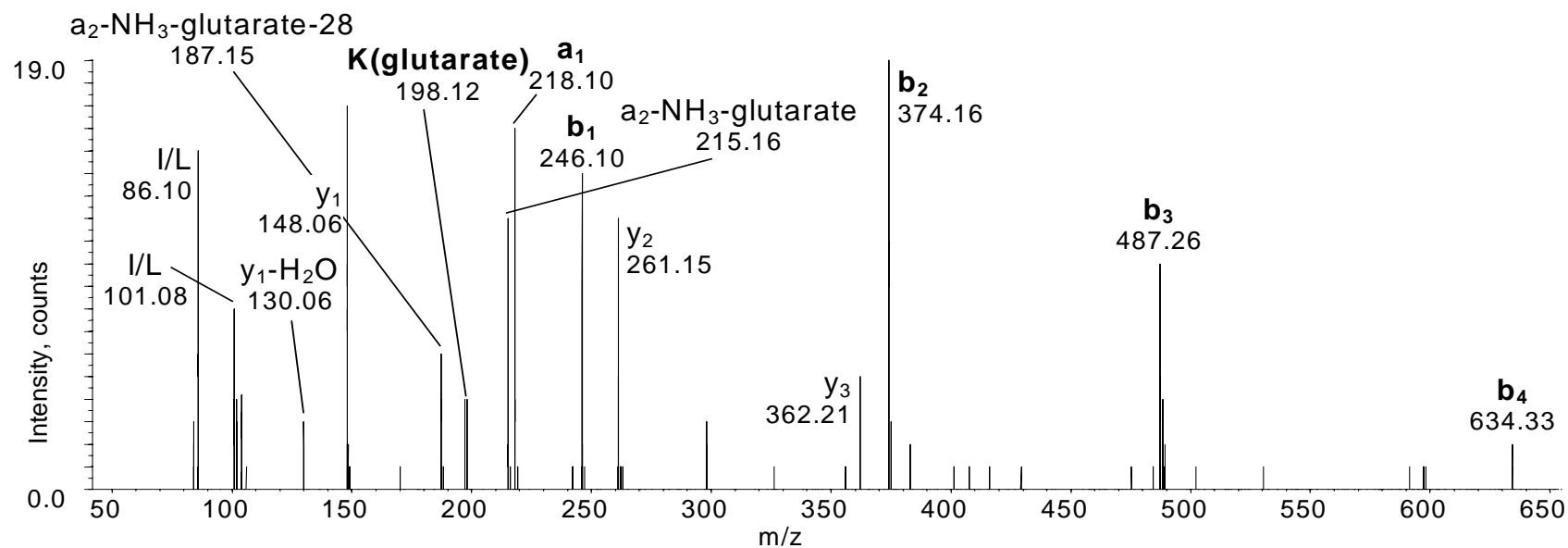
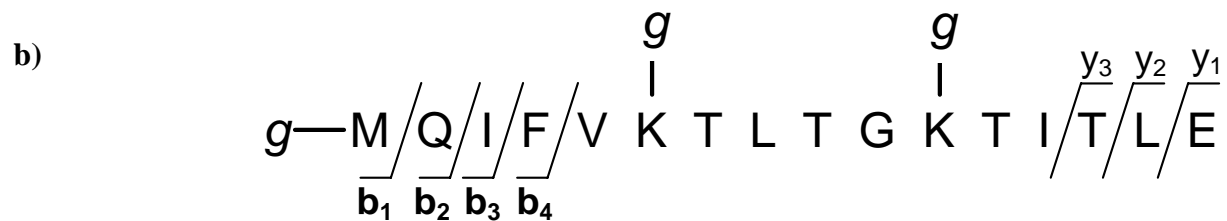


Figure 5.4 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 1-16. Perglutarylated ubiquitin b). *g* indicates glutaryl group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

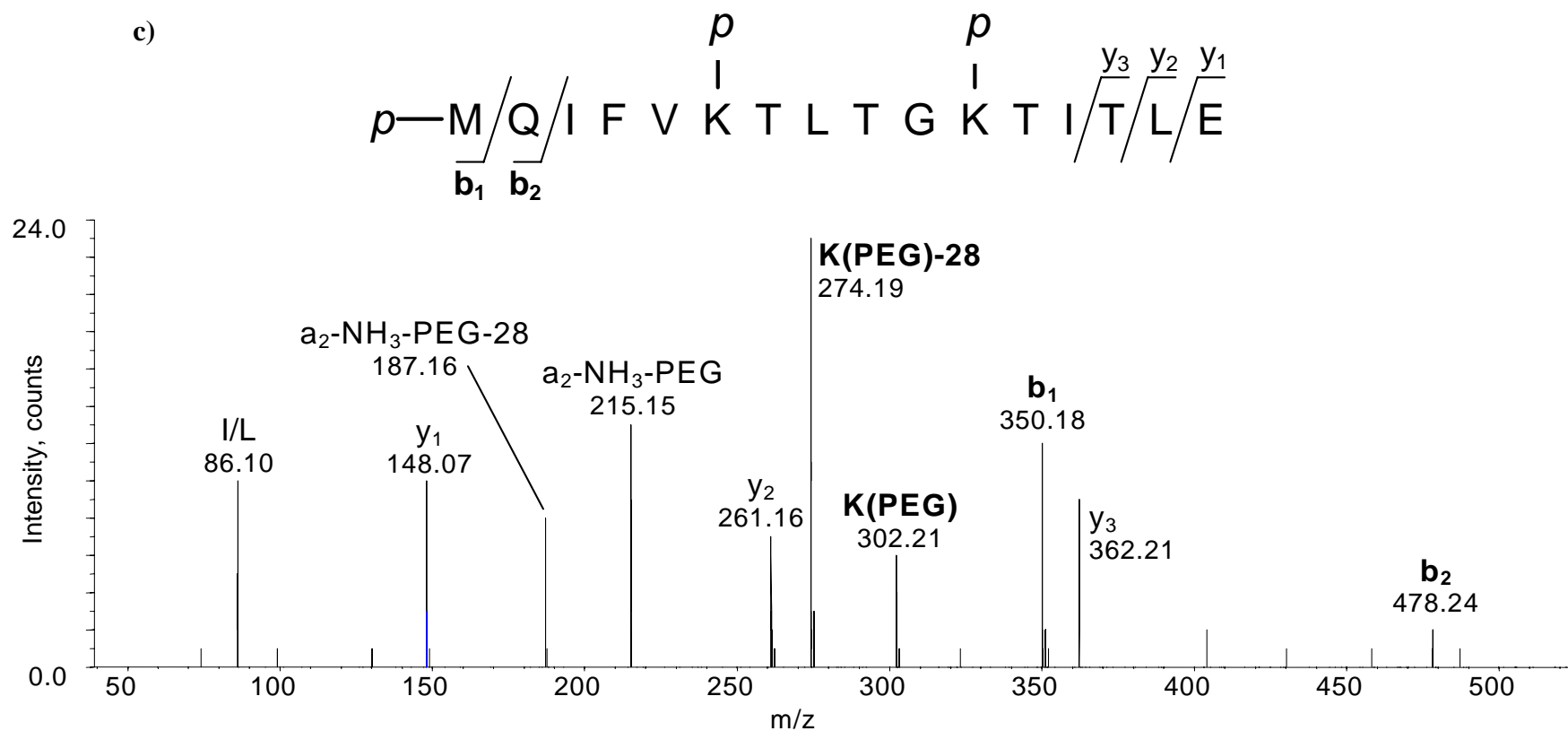


Figure 5.4 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 1-16. PerPEGylated ubiquitin c). *p* indicates PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

This GluC 1-16 peptide was observed late in the HPLC elution; between *ca.* 55 and 58 minutes, depending on the identity of the modifying agent. This late elution is most likely due to both the length of the amino acid chain (a 16-mer) and the presence of three acyl moieties. This result is similar to that observed for AspN peptide 1-20, as shown in Figure 5.3, which elutes at *ca.* 55 minutes.

The MS/MS spectra for peracetylated, perbenzoylated, perglutarylated, periodoacetylated and perPEGylated ubiquitin GluC enzymatic peptide 25-34 are shown in Figure 5.5. The sequence ions observed for all of the peptides include b_2 - b_5 and y_1 - y_4 with an additional b_6 ion observed in the perPEGylated peptide. These sequence ions clearly show the presence of the acyl moieties bound at K-27, K-28 and K-29. In addition, acylated lysine carbenium ions were observed corresponding to K(acetate) at m/z 126.09; K(benzoate) at m/z 188.11; K(glutarate) at m/z 198.12; K(iodoacetate) at m/z 252.00 and K(PEG) at m/z 302.21. The peptides eluted between *ca.* 30 minutes and 44 minutes, being largely dependent on the identity of the acyl moiety. The peptide modified by acetyl eluted at 30.5 minutes, whereas the peptides modified by benzoyl eluted at 43.8 minutes, with the remaining peptides in between these two times.

Figure 5.6 shows the MS/MS spectra for peracetylated, perbenzoylated, perglutarylated, periodoactylated and perPEGylated ubiquitin GluC enzymatic peptide 35-51. These spectra show very similar fragmentation to each other. A small number of sequence ions are observed, with the spectra being dominated by the presence of two proline residues towards the N-terminal end of the peptide. The presence of a proline in a peptide subjected to MS/MS analysis is known to strongly affect the fragmentation pattern observed.¹⁶¹ As a result, an unusual number of internal fragment ions are seen in the spectra for peracylated ubiquitin GluC peptide 35-51. The predominant y-cleavage ions

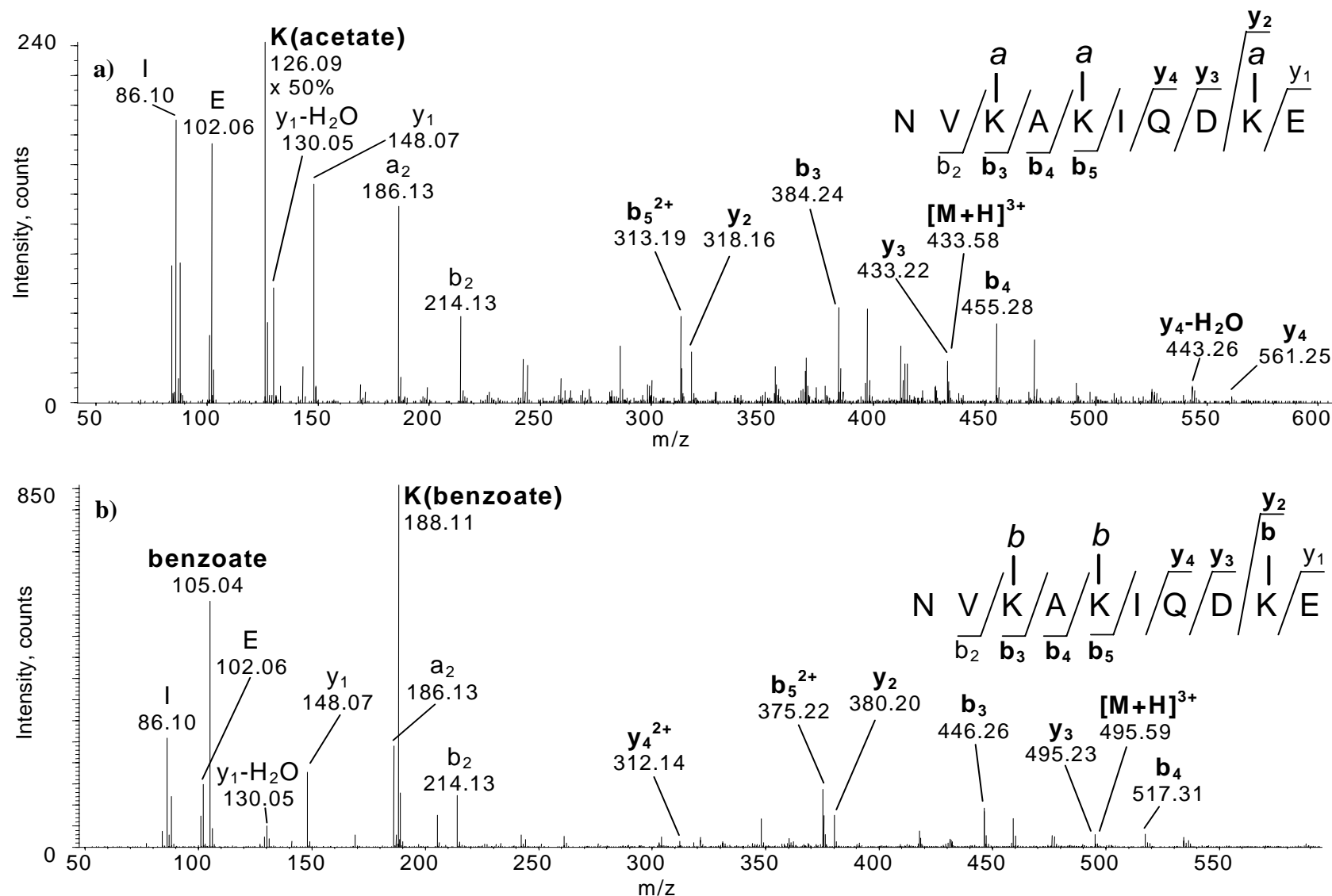


Figure 5.5 HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 25-34. Peracetylated ubiquitin a), perbenzoylated ubiquitin b). *a* indicates acetyl group and *b* indicates benzoyl group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

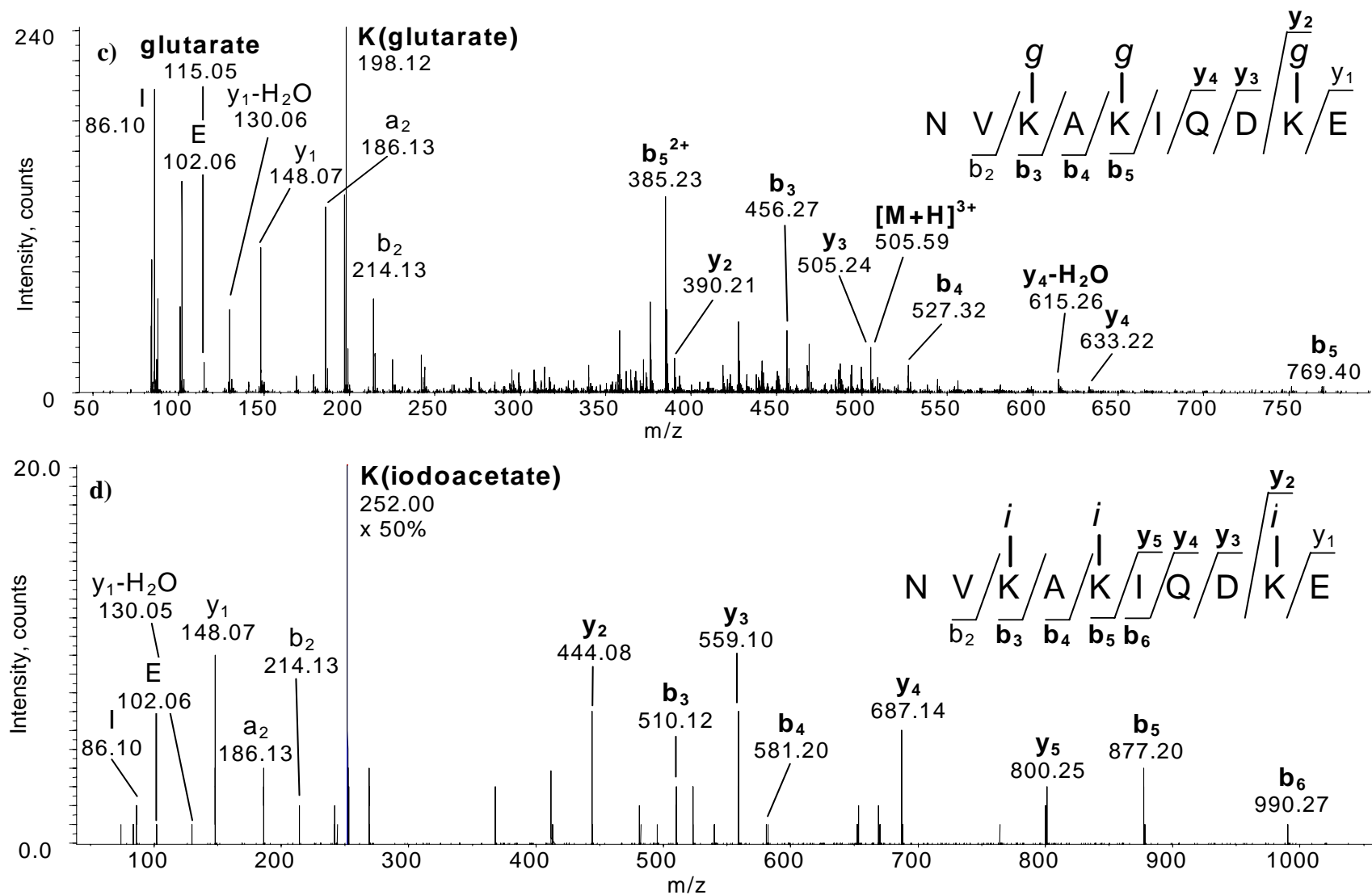


Figure 5.5 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 25-34. Perglutarylated ubiquitin c), periodoacetylated ubiquitin d). *g* indicates glutaryl group, *i* indicates iodoacetyl group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

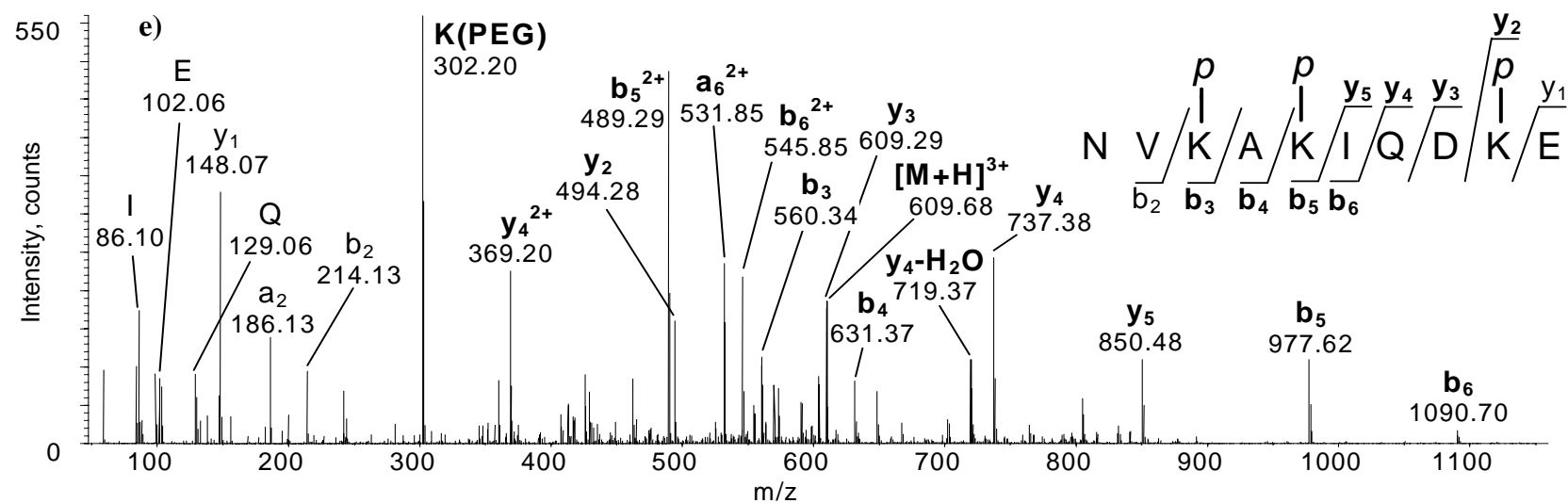


Figure 5.5 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 25-34. PerPEGylated ubiquitin e). *p* indicates PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

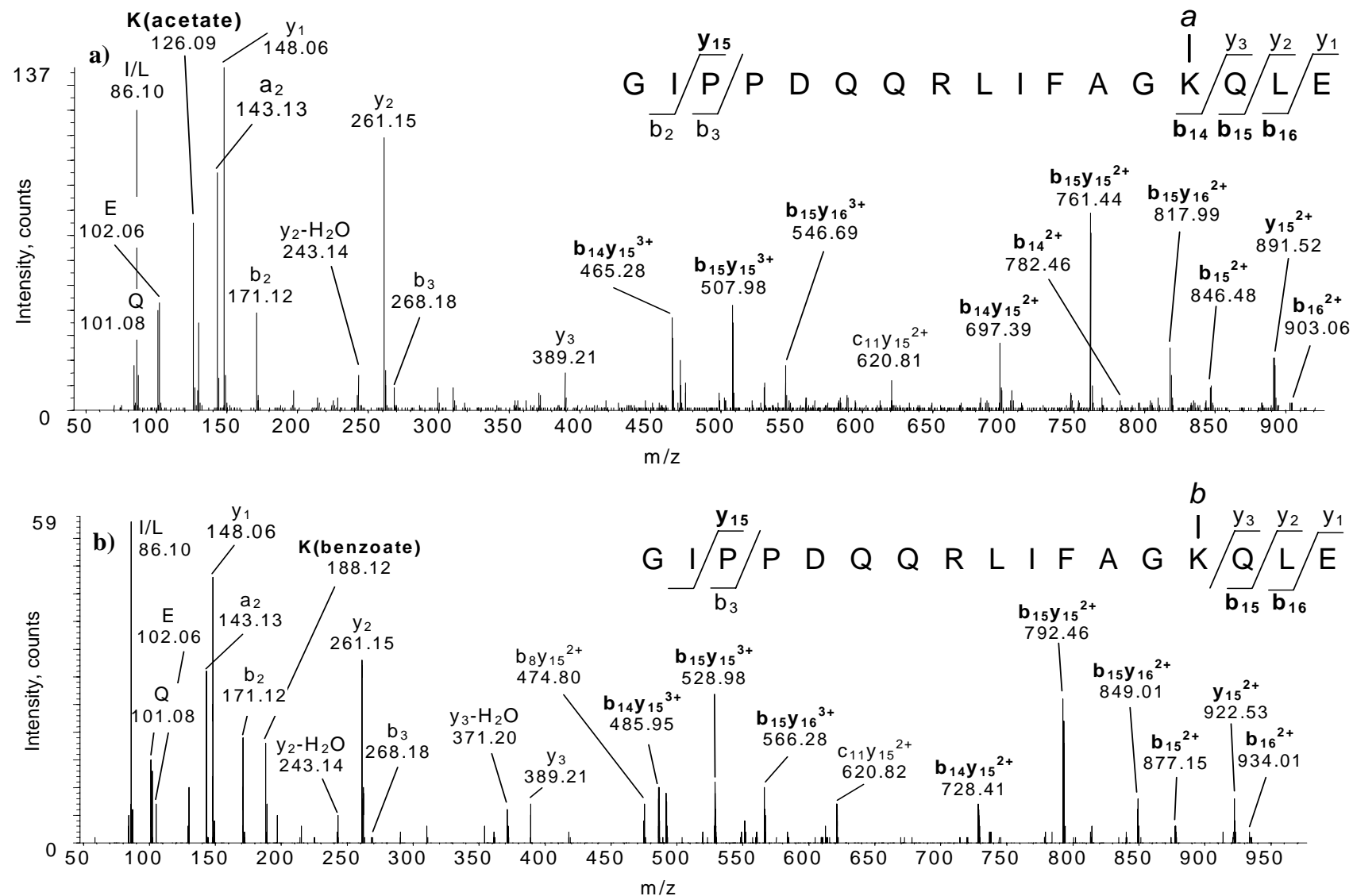


Figure 5.6 HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 35-51. Peracetylated ubiquitin a), perbenzoylated ubiquitin b). *a* indicates acetate group, *b* indicates benzoate group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

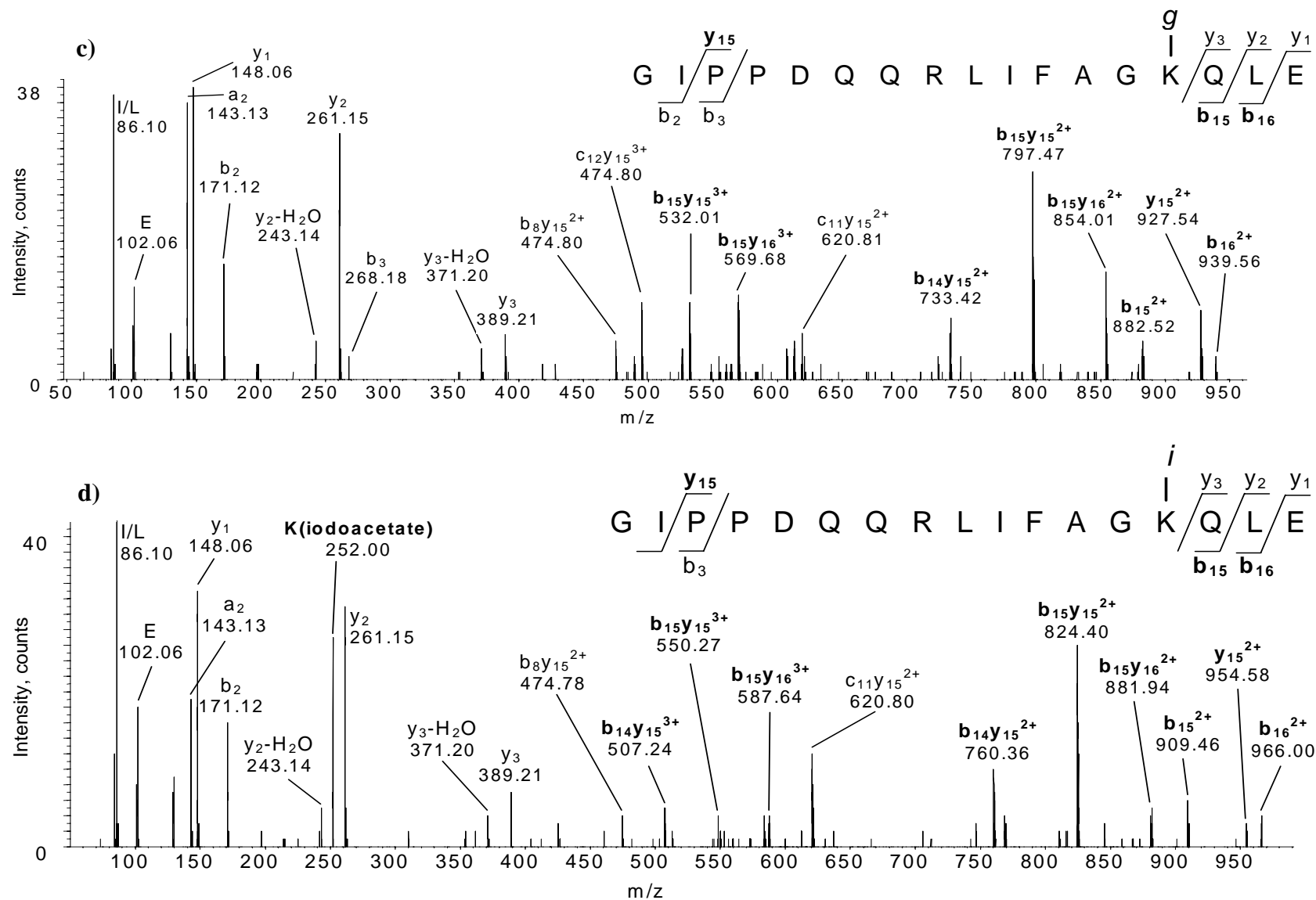


Figure 5.6 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 35-51. Perglutarylated ubiquitin a), periodacetylated ubiquitin b). *g* indicates glutarate group, *i* indicates iodoacetate group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

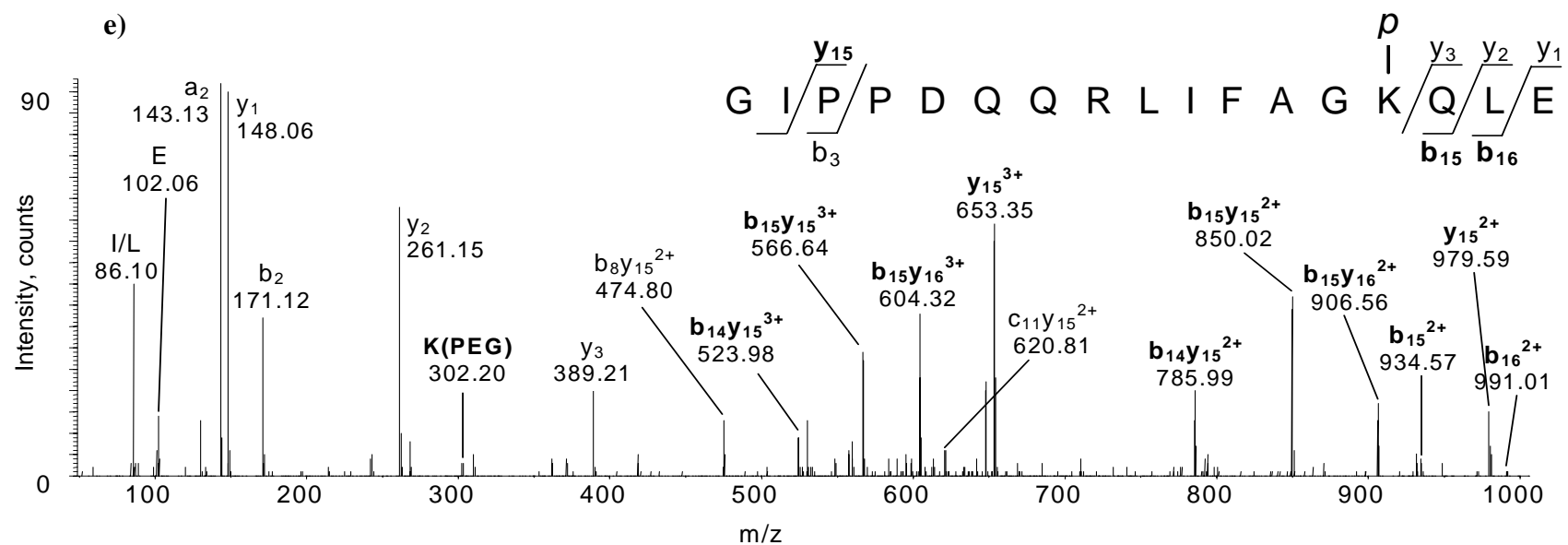
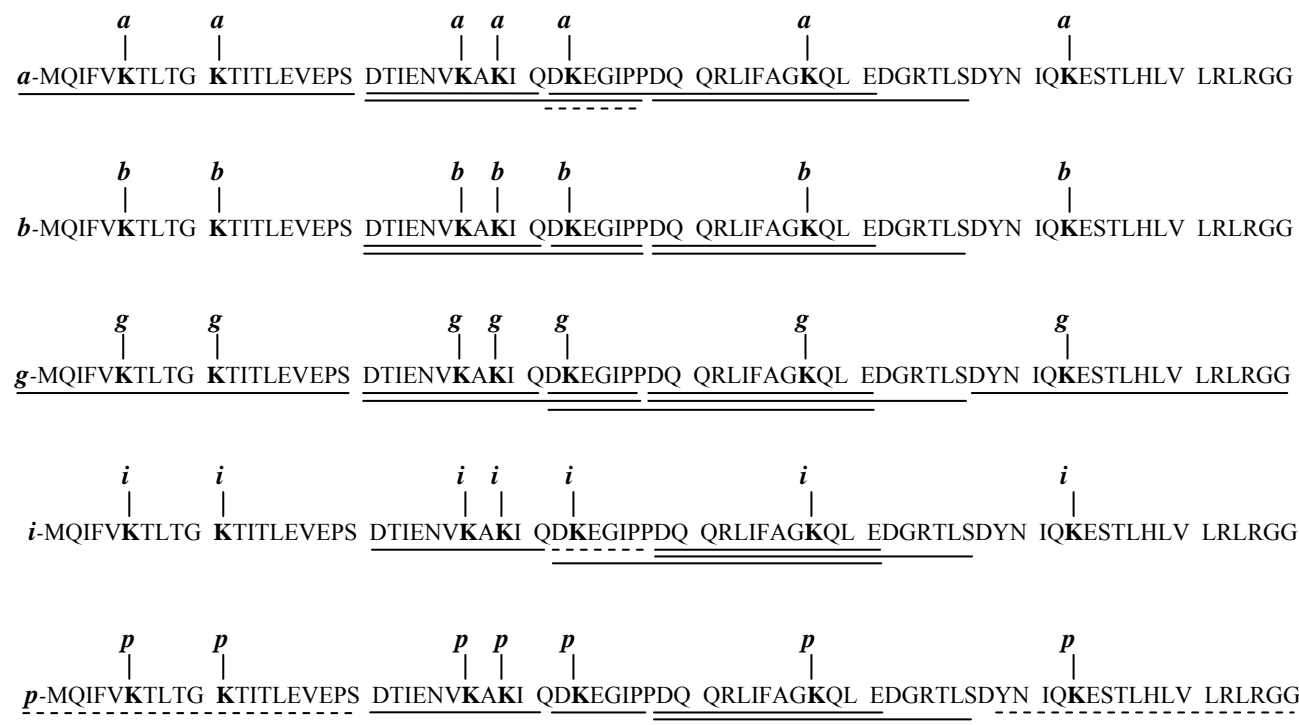


Figure 5.6 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perfunctionalised ubiquitin, GluC peptide 35-51. PerPEGylated ubiquitin e). *p* indicates PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

are to the N-terminal side of the two proline residues and the predominant b-cleavage ions are to the C-terminal side of the acyl moiety. Acylated lysine carbenium ions are observed for each of the peracylated peptides, confirming the presence of the acyl moiety in the peptide. These peptides eluted between *ca.* 34 minutes and 38 minutes, and, similar to the GluC enzymatic peptides 25-34, the peptide modified by acetyl eluted first and the peptide modified by benzoyl last. This peptide contains only one acyl moiety, so it is interesting to note that the range of elution times is significantly less than that observed for the GluC enzymatic peptides 25-34.

The enzymatic peptides observed for the AspN digestion of peracylated ubiquitin are given in Table 5.4 and resultant sequence coverage maps are shown in Scheme 5.3. The majority of m/z values corresponding to peptide ions were of a 2+ or 3+ charge state, with a small number being of a 1+ charge state. High sequence coverage was observed for these AspN enzymatic peptides, with MS/MS spectra being obtained for the complete perglutarated ubiquitin peptide. Selected MS/MS spectra representing the five AspN enzymatic peptides that cover the complete sequence are shown in Figure 5.7 – Figure 5.11.



Scheme 5.3 Sequence coverage maps for the AspN digestion of perfunctionalised ubiquitin. Where a indicates acetate; b indicates benzoate; g indicates glutarate; i indicates iodoacetate and p indicates PEG addition. Peptides analysed by MS/MS are indicated by an unbroken line; peptides detected by MS only are indicated by a dashed line.

Table 5.4 Peptide assignments for the AspN enzymatic digest of perfunctionalised ubiquitin.

m/z ^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment	Calculated [m+H] ⁺ /u	Error (ppm)
Peracetylated Ubiquitin						
797.393	1	797.393	24.9	32-38	797.404	-14
1342.708	1	1342.708	33.4	21-31	1342.721	-10
671.865	2	1342.723	33.4	21-31	1342.721	1
448.245	3	1342.718	33.4	21-31	1342.721	-2
794.427	2	1587.846	31.5	39-51	1587.849	-2
529.959	3	1587.863	31.5	39-51	1587.849	9
707.712	3	2121.121	37.7	21-38	2121.107	6
739.725	3	2217.160	30.7	39-57	2217.162	-1
1181.139	2	2361.271	52.9	1-20	2361.262	4
787.758	3	2361.259	52.9	1-20	2361.262	-1
1183.603	2	2366.198	34.5	32-51	2366.235	-15
789.418	3	2366.239	34.5	32-51	2366.235	1
Perbenzoylated Ubiquitin						
859.408	1	859.408	30	32-38	859.420	-13
430.211	2	859.413	30.0	32-38	859.420	-7
1466.732	1	1466.732	42.1	21-31	1466.753	-14
733.876	2	1466.745	42.1	21-31	1466.753	-5
489.592	3	1466.760	42.1	21-31	1466.753	5
1649.858	1	1649.858	36.0	39-51	1649.865	-4
825.433	2	1649.857	36.0	39-51	1649.865	-4
550.631	3	1649.877	36.0	39-51	1649.865	8
760.409	3	2279.210	34.3	39-57	2279.178	14
769.715	3	2307.130	48.3	21-38	2307.154	-11
Perglutarylated Ubiquitin						
869.416	1	869.416	25.5	32-38	869.425	-1
435.216	2	869.425	25.5	32-38	869.425	-2
1486.761	1	1486.761	33.5	21-31	1486.764	-2
743.886	2	1486.765	34.5	21-31	1486.764	1
496.259	3	1486.762	34.5	21-31	1486.764	-1
1659.868	1	1659.868	31.0	39-51	1659.870	-1
830.434	2	1659.860	32.0	39-51	1659.870	-6

m/z ^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment	Calculated [m+H] ⁺ /u	Error (ppm)
Perglutarylated Ubiquitin continued						
553.963	3	1659.874	32.0	39-51	1659.870	3
763.729	3	2289.172	30.8	39-57	2289.183	-5
776.094	3	2326.266	34.0	58-76	2326.251	6
1169.090	2	2337.172	38.3	21-38	2337.171	0
779.730	3	2337.174	38.3	21-38	2337.171	1
837.425	3	2510.259	32.4	32-51	2510.277	-7
1289.173	2	2577.338	53.2	1-20	2577.326	5
859.784	3	2577.335	53.2	1-20	2577.326	4
Periodoacetylated Ubiquitin						
923.284	1	923.284	26.0	32-38	923.301	-18
462.153	2	923.298	26.0	32-38	923.301	-2
1594.497	1	1594.497	38.1	21-31	1594.515	-11
797.756	2	1594.505	38.1	21-31	1594.515	-6
857.384	2	1713.759	34.0	39-51	1713.746	8
571.926	3	1713.762	34.0	39-51	1713.746	10
1172.018	2	2343.029	32.6	39-57	2343.059	-13
781.694	3	2343.068	32.6	39-57	2343.059	4
873.348	3	2618.027	37.8	32-51	2618.028	-1
PerPEGylated Ubiquitin						
973.499	1	973.499	27.0	32-38	973.509	-10
487.257	2	973.505	27.0	32-38	973.509	-4
1694.894	1	1694.894	36.5	21-31	1694.931	-22
847.966	2	1694.924	36.5	21-31	1694.931	-4
565.645	3	1694.920	36.5	21-31	1694.931	-6
882.476	2	1763.945	33.2	39-51	1763.954	-5
588.657	3	1763.955	33.2	39-51	1763.954	1
798.438	3	2393.298	32.0	39-57	2393.267	13
810.790	3	2430.353	35.0	58-76	2430.335	7
906.815	3	2718.430	36.0	32-51	2718.445	-6
963.867	3	2889.586	55.1	1-20	2889.577	3

^a Bold-face type indicates peptides for which MS/MS data was obtained

The MS/MS spectrum of the AspN enzymatic peptide 1-20 of peracetylated ubiquitin is shown in Figure 5.7. The presence of b_1 - b_5 and y_1 - y_3 confirm the identity of the peptide. The b_1 ion is the same as that observed for the GluC 1-16 peptide and is observed due to the presence of a modification to the N-terminal amino acid. The spectrum is dominated by an intense y_2 ion, which is shown at 10 % relative to its actual peak height. An acetylated lysine carbenium ion is observed in the spectrum at m/z 126.1, which is the second most intense peak observed. This peptide elutes at 52.9 minutes with an elution time of up to *ca.* 55 minutes observed for the PEG-modified peptide. This is consistent with the elution time for the GluC 1-16 peptide discussed earlier.

Figure 5.8 shows the spectrum of the AspN enzymatic peptide 21-31 of peracetylated ubiquitin. An almost complete series of b-ions observed and a series of five y-ions, both of which give confirmation of the sites of acetylation. An acetylated lysine carbenium ion is observed in the spectrum. The acetylated, glutarated, iodoacetylated and PEGylated peptides elute in the range *ca.*33-37 minutes, with the benzoylated peptide eluting appreciably later at 42.1 minutes. A similar observation was made for the GluC peptide 24-34, which covers a similar sequence of the ubiquitin polypeptide chain. This indicates that the presence of a benzoyl moiety on an enzymatic peptide significantly alters the elution time of a given peptide sequence relative to the other acylating agents examined.

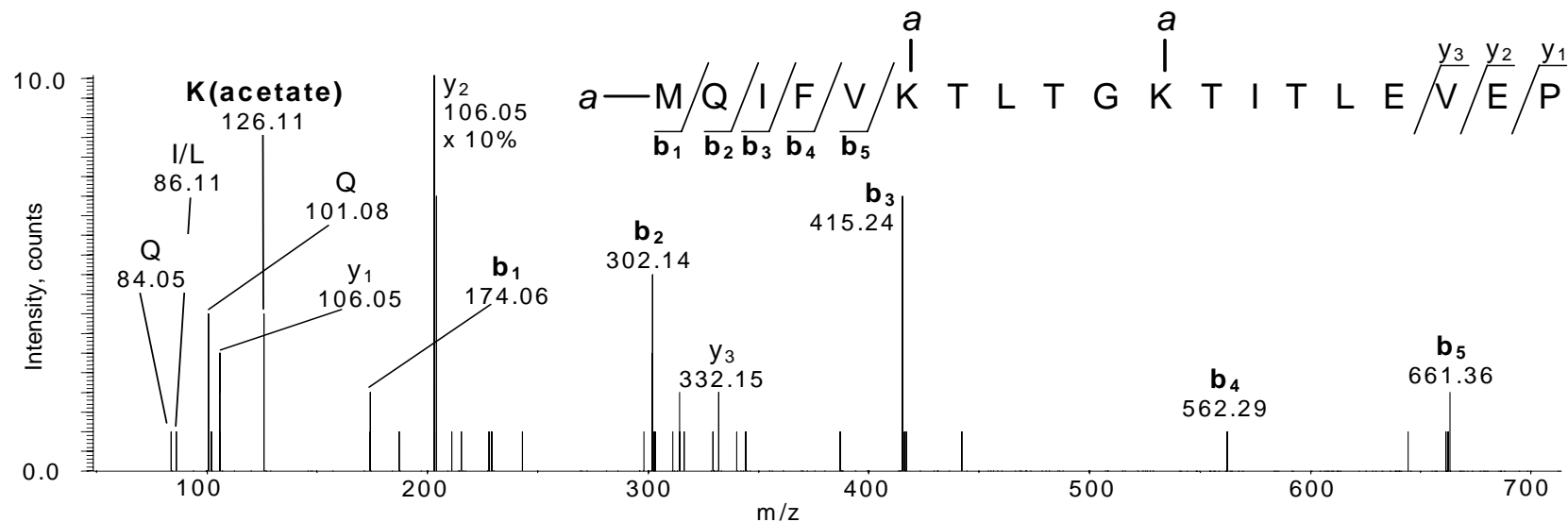


Figure 5.7 HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 1-20. **a** indicates acetate group; bold type indicates the presence of a product ion containing the acyl moiety.

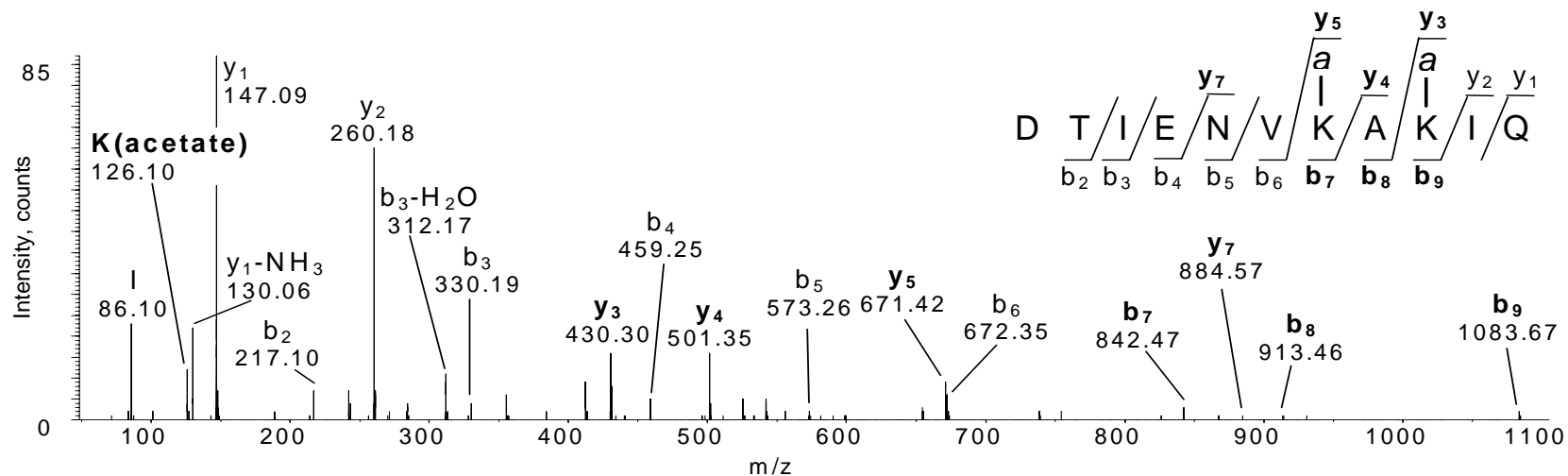


Figure 5.8 HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 21-31. **a** indicates acetate group; bold type indicates the presence of a product ion containing the acyl moiety.

The MS/MS spectrum for the AspN benzoylated peptide 32-38 is shown in Figure 5.9. A complete b-ion series is observed for this peptide. A relatively intense benzoylated lysine carbenium ion (third most intense ion), is also observed. These observations confirm the presence of a lysine on K-33, the second amino acid in the sequence of the peptide. The most intense ions in the spectrum are y_1 and y_2 , which have been reduced to 60 % and 25 % respectively from their original intensities for clarity of the spectrum. These ions are due to cleavages adjacent to proline residues, which commonly result in the preferred cleavage locations.²³⁸ There is also a small peak assigned to a benzoate ion (at m/z 105.0). This ion is in keeping with the observation of ions formed by the removal of the acyl moiety from the peptide backbone, as also observed for straight chain acyl moieties as discussed in Chapter 3. The observed acylated AspN enzymatic peptides 32-38 eluted between *ca.* 25 and 30 minutes. These peptides probably elute earlier than other peptides due to both the short amino acid chain (a 7-mer) and the presence of only one acyl moiety.

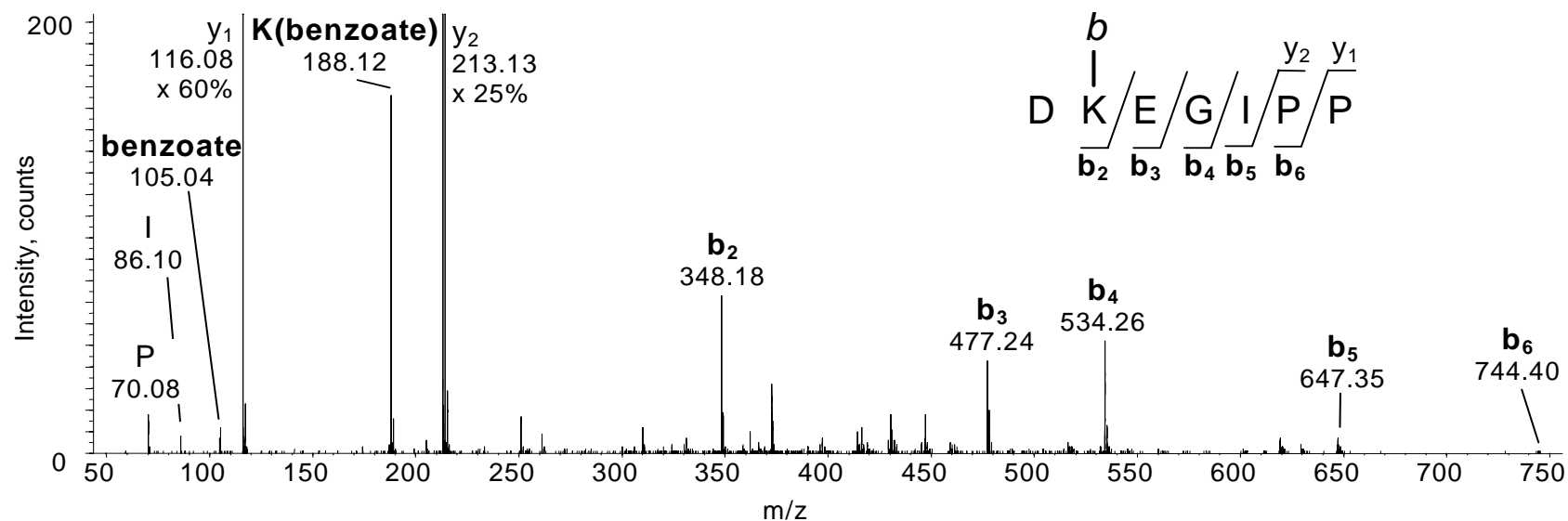


Figure 5.9 HPLC-ESI-Q-o-TOF MS/MS analysis of perbenzoylated ubiquitin, AspN peptide 32-38. *b* indicates benzoate group; bold type indicates the presence of a product ion containing the acyl moiety.

Figure 5.10 shows the MS/MS spectrum of the acetylated ubiquitin AspN enzymatic peptide 39-51. A complete b-ion series is observed and y-ions y_1 - y_7 are also observed. An acetylated lysine carbenium ion is observed at m/z 126.1. These ions confirm the presence of the acetylated lysine K-48. A number of internal fragment ions with relatively high ion abundances are observed, being $b_{11}y_5$, $b_{11}y_6$ and $b_{12}y_5$. An m/z value corresponding to a c_9 fragment ion (m/z 1046.64) appears in the MS/MS spectrum, which is an unusual ion to observe in an ESI-Q-o-TOF MS/MS spectrum. This c_9 cleavage is adjacent to the acetylated lysine residue. The proximity to the modified lysine residue suggests that the acetyl moiety may aid in the formation of the c-ion through localised stabilisation of the neutral loss fragment, or of an intermediate. The ubiquitin AspN enzymatic peptide 39-51 eluted between *ca.* 31 and 34 minutes. This peptide compares with that of AspN enzymatic peptide 21-31 which contains eleven amino acid residues and two acyl moieties. Of these two peptides, the one containing two acyl moieties elutes later (by *ca.* 2-3 minutes), despite being of a shorter chain length, when compared to the peptide containing thirteen amino acids and only one acyl moiety.

The MS/MS spectrum of the perglutarylated ubiquitin AspN peptide 58-76 is shown in Figure 5.11. This relatively long peptide (19-mer) shows predominately N-terminal fragmentation, with b_2 - b_4 and y_{12} - y_{17} ions being observed. These ions cover the region of the peptide containing glutarylated lysine residue (K-63). The two peptides for which this peptide was observed (perglutarylated and perPEGylated ubiquitin) elute at 34.0 minutes and 35.0 minutes, significantly earlier than for the AspN enzymatic peptide 1-20 with elution times of 53.2 minutes and 55.1 minutes respectively. This significant difference in elution time can be attributed to the presence of one acyl moiety in AspN enzymatic peptide 58-76 and the presence of three acyl moieties in the AspN enzymatic peptide 1-20.

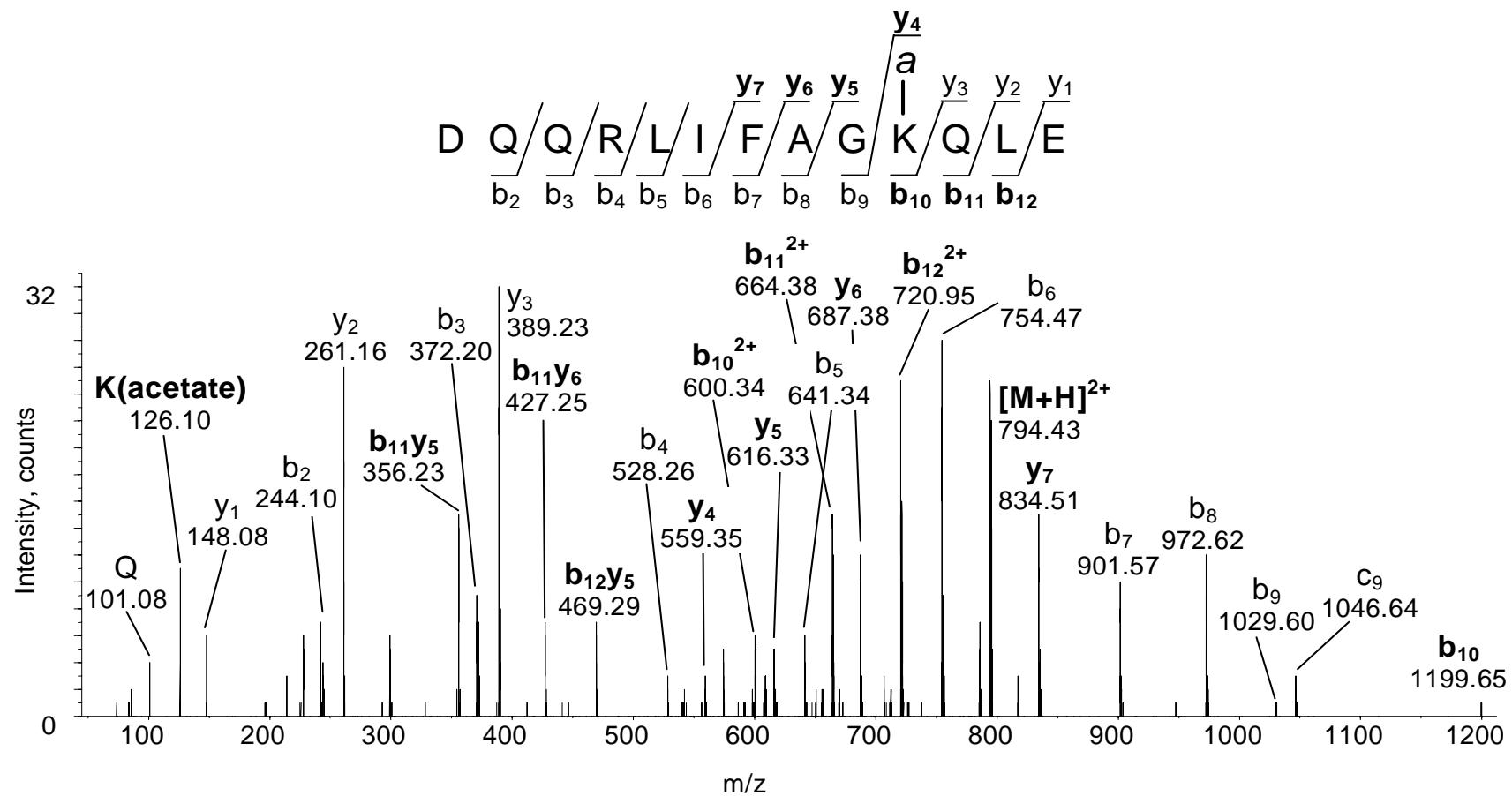


Figure 5.10 HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated ubiquitin, AspN peptide 39-51. **a** indicates acetate group; bold type indicates the presence of a product ion containing the acyl moiety.

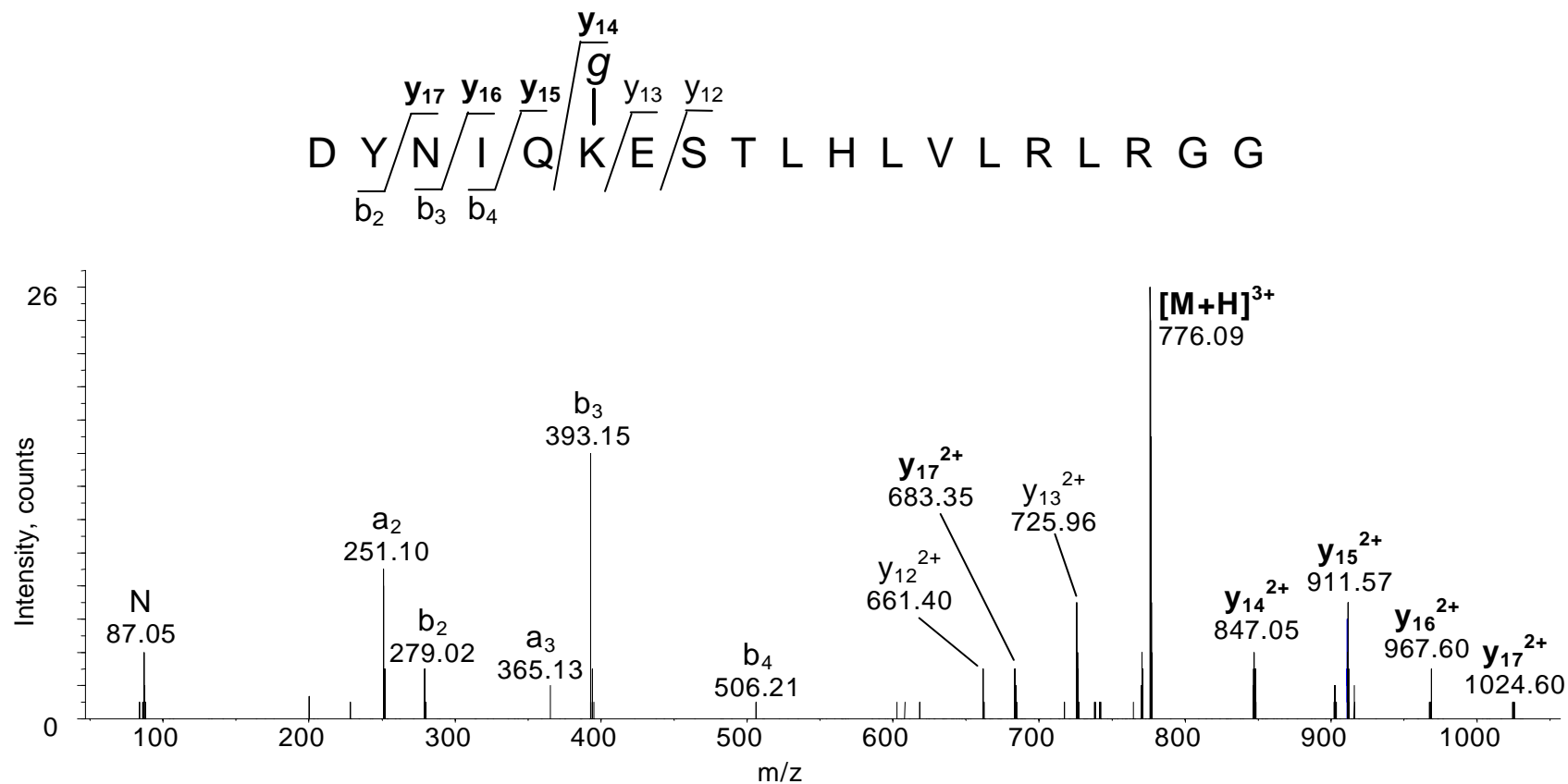


Figure 5.11 HPLC-ESI-Q-o-TOF MS/MS analysis of perglutared ubiquitin, AspN peptide 58-76. **g** indicates glutarate group; bold type indicates the presence of a product ion containing the acyl moiety.

5.2.1.3 Identification of unreacted lysine residues in acylated ubiquitin.

When acylated ubiquitin was analysed by MALDI-TOF MS it was noted that a small proportion of both the glutarylated and PEGylated forms of ubiquitin were missing one functional moiety from the perfunctionalised ubiquitin molecule (see Figure 5.2). Based on peak heights, the proportion of the glutarylated ubiquitin missing one glutaryl moiety corresponds to approximately 20 % of the total glutarylated ubiquitin, and approximately 40 % of the total PEGylated ubiquitin is missing one PEG moiety. There was no m/z value corresponding to a sub-stoichiometric addition of acetyl to ubiquitin in the MALDI-TOF MS spectrum. Additionally, there was no evidence for sub-stoichiometric addition of the functional groups in the HPLC-ESI-Q-o-TOF MS/MS data for the acetylated, benzoylated or iodoacetylated ubiquitin samples.

In order to determine whether any specific sites were involved in the sub-stoichiometric addition of glutaryl or PEG to the ubiquitin molecule the HPLC-ESI-Q-O-TOF MS/MS data was analysed further. The results are summarised in Table 5.5. The enzymatic peptides identified by HPLC-ESI-Q-O-TOF MS/MS analysis give evidence that the sites at which lysine has not been functionalised include K-29 and K-48. The evidence also supports an additional missing adduct at either K-6 or K-11. Data supporting these sites was found in the form of intact masses corresponding to enzymatic peptides missing one functional group and also by tandem mass spectrometry of several peptides (as indicated by bold-face type in Table 5.5). Tandem mass spectra for the enzymatic digestion of sub-stoichiometrically PEGylated ubiquitin confirming unmodified sites K-29 and K-48 are shown in Figure 5.12 and Figure 5.13 respectively.

The MS/MS spectrum for sub-stoichiometrically PEGylated ubiquitin GluC peptide 25-34 is shown in Figure 5.12 part a). This spectrum gives a clear indication that K-29 is unmodified in this particular peptide when compared to the corresponding perPEGylated spectrum (Figure 5.5 part e)). There are two new peaks corresponding to the di-PEGylated b_5 and b_6 PEGylated peptide, accompanied by an absence of the b_5 and b_6 present in the MS/MS spectrum of the tri-PEGylated peptide. There is an additional appearance of an ion at m/z 129.10, corresponding to a lysine amino-acylium ion, which is absent from the MS/MS spectra of per-acylated peptides. Correspondingly, the relative intensity of the K(PEG) carbenium ion at m/z 302.20 is reduced in the MS/MS spectrum of the di-PEGylated peptide when compared to the MS/MS spectrum of the tri-PEGylated peptide. There was no m/z value corresponding to an unmodified b_3 ion (indicating an unmodified K-27) and also no m/z value corresponding to an unmodified y_2 ion (indicating an unmodified K-33). The MS/MS spectrum for unmodified ubiquitin AspN peptide 39-51 is shown in Figure 5.13. There is a distinct series of y -ions from y_1 to y_7 that indicate the absence of a PEGyl moiety at K-48.

The MS data as presented in Table 5.5 indicates the presence of a sub-stoichiometrically PEGylated Glu-C enzymatic peptide 1-16. The tandem mass spectrum (not shown) is similar to that obtained for the equivalent per-PEGylated peptide shown in Figure 5.4. Both MS/MS spectra show only b_1 , b_2 and y_{1-3} product ions which do not allow for a distinction between the K-6 and K-11 sites at which the acyl moiety could be absent. It is possible that both sites are sub-stoichiometrically acylated throughout the sample. There was no evidence to support the sub-functionalisation of sites other than K-6 or K-11, K-29 and K-48. In addition, there was no evidence for the loss of two acyl moieties from any single enzymatic peptide, as detected by HPLC-ESI-Q-O-TOF MS/MS.

Table 5.5 HPLC-ESI-Q-o-TOF MS/MS analysis of glutarylated and PEGylated ubiquitin containing sub-stoichiometric functionalisation of ubiquitin.

m/z^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment	Calculated [m+H] ⁺ /u	Error (ppm)
Sub-stoichiometric addition of Glutaryl to Ubiquitin						
AspN digestion						
686.866	2	1372.725	28.4	21-31 missing one Glutaryl	1372.732	-5
458.249	3	1372.731	28.4	21-31 missing Glutaryl at K-29	1372.732	-1
515.951	3	1545.837	27.0	39-51 missing one Glutaryl	1545.838	-1
821.777	3	2463.316	46.8	1-20 missing one Glutaryl	2463.294	9
GluC digestion						
700.870	2	1400.732	27	25-34 missing one Glutaryl	1400.727	4
467.585	3	1400.738	27	25-34 missing Glutaryl at K-29	1400.727	8
684.378	3	2051.117	49.5	1-16 missing one Glutaryl	2051.098	9
Sub-stoichiometric addition of PEG to Ubiquitin						
AspN digestion						
738.907	2	1476.806	29.7	21-31 missing PEG at K-29	1476.816	-7
492.940	3	1476.806	29.7	21-31 missing PEG at K-29	1476.816	-7
773.415	2	1545.822	26.1	39-51 missing one PEG	1545.838	-10
515.951	3	1545.838	26.1	39-51 missing K-48	1545.838	0
725.723	3	2175.153	26.7	39-57 missing one PEG	2175.152	1
811.112	3	2431.319	35.4	21-38 missing one PEG	2431.307	5
834.111	3	2500.318	32.4	32-51 missing 1 PEG	2500.329	-5
GluC digestion						
804.959	2	1608.910	31.7	25-34 missing PEG at K-29	1608.894	10
536.979	3	1608.920	31.7	25-34 missing PEG at K-29	1608.894	16
637.346	3	1910.023	32.5	35-51 missing PEG at K-48	1910.049	-14
1130.13 3	2	2259.258	45.0	1-16 missing one PEG	2259.266	-3
753.772^b	3	2259.301	45.0	1-16 missing one PEG	2259.266	16

^a bold-face type indicates peptides for which MS/MS data was obtained.

^b fragmentation data is not indicative of the position at which the PEG moiety is missing.

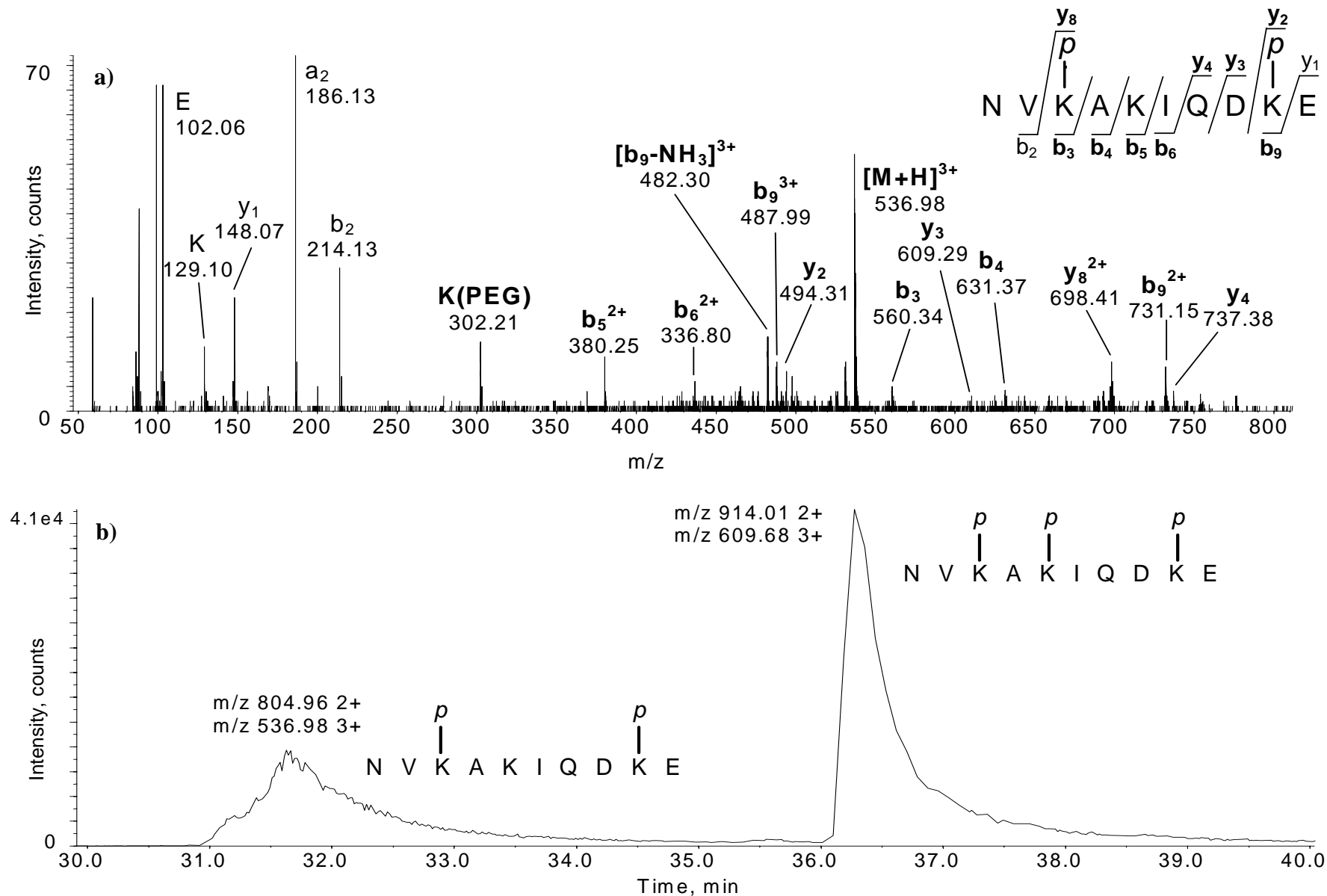


Figure 5.12 HPLC-ESI-Q-o-TOF MS/MS analysis of the sub-stoichiometrically PEGylated ubiquitin GluC peptide 25-34 a). Selected ion chromatogram for the combined 2+ and 3+ states of substoichiometrically PEGylated and perPEGylated ubiquitin AspN peptide 25-34 b). *p* indicates PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

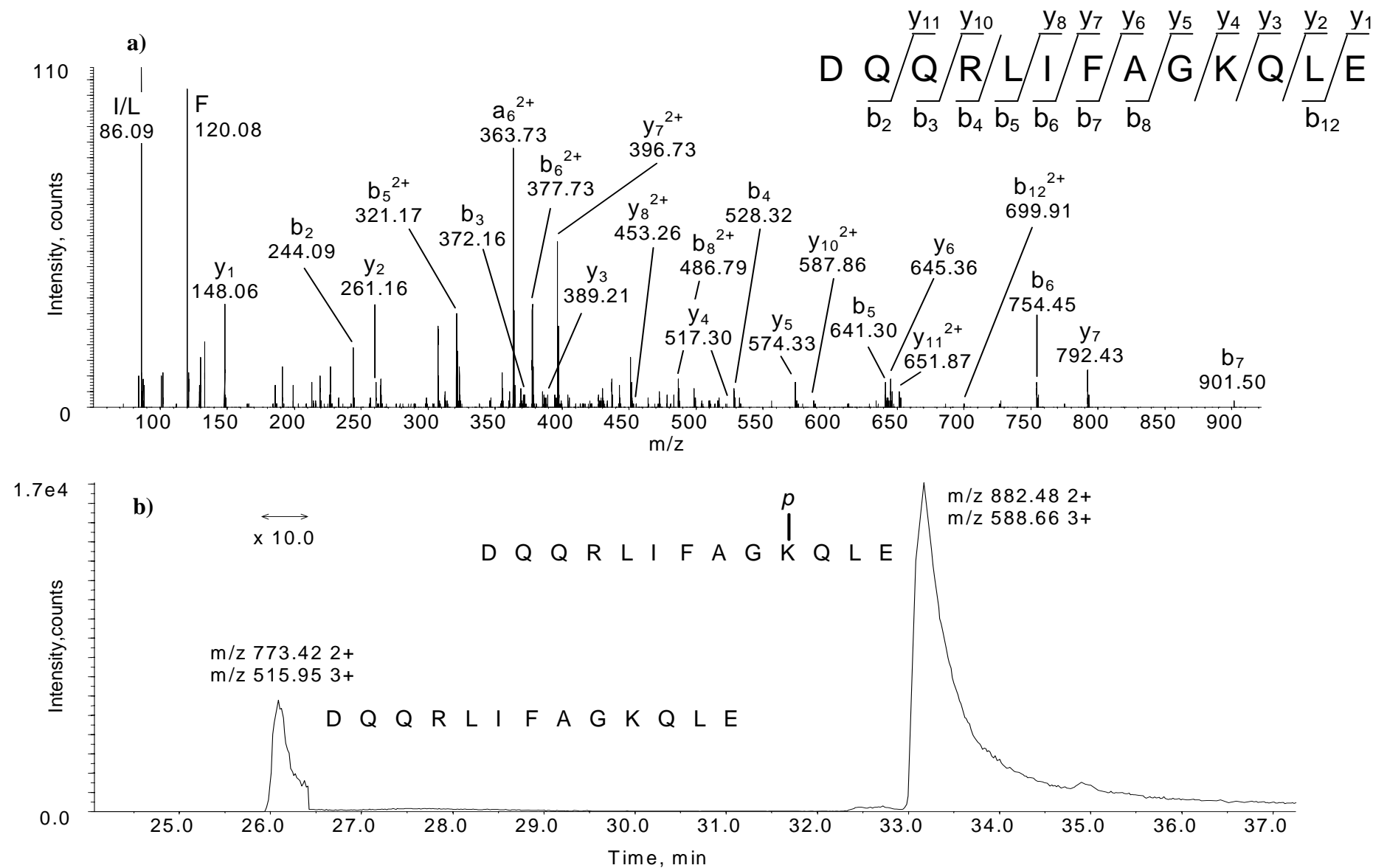


Figure 5.13 HPLC-ESI-Q-o-TOF MS/MS analysis of the unmodified ubiquitin AspN peptide 39-51 a). Selected ion chromatogram for the combined 2+ and 3+ states of the native and PEGylated ubiquitin AspN peptide 39-51 b).

For the two ESI-Q-o-TOF MS/MS spectra shown, the corresponding selected ion chromatogram for both the peptide missing one PEG, and the fully functionalised peptide, is also shown. From the selected ion chromatogram there is a clear indication that the retention time for peptides missing one PEG group is less than that for the fully functionalised peptides. This phenomenon is also shown by comparing the retention times for individual enzymatic peptides of ubiquitin (given in Table 5.5) with the corresponding retention time given in Table 5.3 or Table 5.4, indicating that this observation is consistent for all of the sub-stoichiometrically functionalised peptides. The shorter retention times for sub-stoichiometrically functionalised peptides are an expected result, as the functional groups of glutaryl and PEG would increase the HPLC retention time. This is caused by increased hydrophobicity, resulting from an increased carbon chain length when compared with the native lysine residue. Additionally, there was observed a greater difference in retention times between sub- and perPEGylated peptides compared with the corresponding sub- and perglutarylated peptides. This is consistent with the difference in chain length between a glutaryl and PEGyl moiety.

The data were examined to determine whether any particular lysine was preferentially missed in the synthetic acylation process. On examination of the TIC obtained for the HPLC-ESI-Q-O-TOF MS/MS of AspN digested PEGylated ubiquitin, it was observed that two of the peaks (2+ and 3+ charge states) corresponding to the AspN peptide 21-31 (m/z 738.9 and m/z 492.9) are amongst the major peaks in the combined spectrum (see Figure 5.3). m/z values for the remaining peptides missing a PEG group are visible only upon closer inspection of the spectrum.

The predominant sub-functionalisation of K-29 can be explained by the increased steric hindrance at this site. K-29 is in close proximity to two other lysine residues: K-27 and K-33. It is postulated that, if the particular order of functionalisation during the reaction process for any given ubiquitin molecule means that K-27 and K-33 are functionalised, steric effects would hinder the reaction at the K-29 residue. If the order of functionalisation is different, such as K-27, followed by K-29, followed by K-33, then the steric effects limiting the accessibility at K-29 would be decreased. This explanation is in agreement with the approximate ratio of perfunctionalised AspN peptide 21-31 to sub-functionalised AspN peptide 21-31 as shown in Figure 5.3. This is approximately 1:3, with the statistical chances of K-29 being the last of the three sites to be functionalised also being 1:3. (It should be noted, however, that when comparing peaks in an electrospray mass spectrum, peptides containing a native lysine residue may preferentially ionise more efficiently than a peracylated peptide during the electrospray process. This occurs because they contain an additional free primary amine group. An over estimation of the proportion of protein containing an unmodified lysine residue may result.)

Further support for the sub-stoichiometric functionalisation of ubiquitin being due to steric effects is given by the data showing that unmodified lysines were only detected for glutarylated and PEGylated ubiquitin, which are of a longer chain length compared with acetyl-, benzoyl- or iodoacetylated groups. Additional evidence from the MALDI-TOF MS data (Figure 5.2) shows a lower degree of sub-stoichiometric addition of glutaryl compared with PEG; glutaryl (with a five carbon chain length) being a smaller moiety compared to PEG (with a ten carbon chain length).

5.2.1.4 Identification of sites containing truncated PEG in PEGylated ubiquitin.

According to the MALDI-TOF-MS of the PEGylated ubiquitin sample (Figure 5.2 part a)) there appears to be a large proportion of a truncated form of PEG. Two possible sites containing this truncation were identified in the HPLC-ESI-Q-o-TOF MS data from the AspN digestion of PEGylated ubiquitin (see Table 5.6). No MS/MS data was obtained for these peptides. No m/z values corresponding to peptides containing a truncated PEG were detected in the HPLC-ESI-Q-o-TOF MS data from the GluC digestion of PEGylated ubiquitin. The clear presence of a truncation in the intact spectrum, with little evidence in the analysis of the digested peptides, suggests that any truncations of the PEG moiety in PEGylated ubiquitin are spread throughout the eight possible sites of functionalisation, rather than at specific locations in ubiquitin.

The retention time of the PEGylated ubiquitin AspN peptide 58-76 proposed to contain a truncated PEG moiety, is shorter than that of the corresponding peptide containing the intact PEG moiety. This is in concordance with a truncated PEG moiety having a decreased degree of interaction with the LC medium, due to a shorter carbon chain length when compared to the intact PEG moiety. The mass difference between a full-length PEG moiety and the shorter moiety was calculated to be 116.13 u from the data for the AspN PEGylated ubiquitin peptide 58-76. The mass accuracy of the data obtained, both from MALDI-TOF MS and ESI-Q-o-TOF MS does not allow a confident determination of the molecular formula of the shorter form of the PEG moiety. A suggested formula is $\text{HOCH}_2\text{CH}_2\text{OCH}_2\text{C(O)-}$ (ethylene glycol acetyl; delta mass 102.032 u). The source of this impurity is most likely from the reagents used in the synthetic PEGylation of ubiquitin. This is supported by the absence of any observed

truncated PEG addition in the PEGylation of bovine carbonic anhydrase II (see section 5.2.2 below).

Table 5.6 Identification of possible sites containing truncated PEG addition to ubiquitin.

<i>m/z</i>	<i>z</i>	Experimental [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	Elution time (min)	Assignment	Calculated [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	Error (ppm)
AspN digestion						
772.081	3	2314.227	34.0	58-76 truncated PEG at K-63	2314.251	-10
981.859	3	2943.561	55.2	52-76 truncated PEG at K-63	2943.565	-1

5.2.2 Analysis of Peracylated Bovine Carbonic Anhydrase II by Mass Spectrometry

Bovine carbonic anhydrase II (BCA) is a medium-sized globular protein (*ca.*29 kDa see Scheme 5.4.). BCA was selected for study as is it readily available and is often used as a model protein.³²¹⁻³²³ BCA consists of a single polypeptide chain with no disulphide bonds; in its native form zinc and acetate are bound at its active site. BCA contains 259 residues, including 18 lysine residues. An initial methionine is removed post-translationally and the acetylated N-terminal serine residue is acetylated post-translationally.³²⁴

(ACET)SHHWGYG K HN	GPEHWH K DFF	IANGERQSPV	DIDT K AVVQD	PAL K PLALVY
GEATSRRMVN	NGHSFNVEYD	DSQD K AVL K D	GPLTGTYRLV	QFHFHWGSSD
DQGSEHTVDR	K KYAAELHLV	HWNT K YGDFG	TAAQQPDGLA	VVGVF L K VGD
ANPALQ K VLD	ALDSI K T K G K	STDFPNFDPG	SLLPNVLDYW	TYPGSLTTPP
LLESVTWIVL	K EPISVSSQQ	ML K FRTLNFN	AEGEPELLML	ANWRPAQPL K
NRQVRGFP K				

Scheme 5.4 Amino acid sequence for bovine carbonic anhydrase. (Lysine residues are indicated in bold.)

5.2.2.1 HPLC-ESI-Q-o-TOF MS/MS Analysis of Enzymatically digested Peracylated Bovine Carbonic Anhydrase II (BCA)

Peracetylated, pertrifluoroacetylated, perPEGylated and native bovine carbonic anhydrase II (BCA) were supplied by collaborators Katie Gudiksen *et. al.* and Jerry Yang *et. al.*^{17, 317} These modified proteins were enzymatically digested using AspN. PerPEGylated BCA was also digested using GluC. Standard protein enzymatic digestion conditions were employed as described in Chapter 2. Subsequent to digestion, samples were diluted to an appropriate concentration using Buffer A and analysed by HPLC-ESI-Q-o-TOF MS/MS, as described in Chapter 2. Results were obtained using Intelligent Data AcquisitionTM (IDATM; Applied Biosystems), with peptides of 2+, 3+ or 4+ charge state in the range m/z 350 – 1500 being selected for analysis. The resultant total ion current (TIC) trace for AspN digested perPEGylated BCA is shown in Figure 5.14. The majority of peptides resulting from the AspN enzymatic cleavage elute between 22 and 40 minutes. A shorter peptide (7-mer) eluted at *ca.* 16 minutes and a group of 12 to 20-mers at *ca.* 46 minutes. Similar TIC traces were observed for the remaining acylated BCA proteins and also for the GluC digested PEGylated BCA. The peptides eluted according to the type of amino acid sequence (length and nature) and

acylation (degree and type). These factors are discussed for each group of peptides in the sections that follow.

The HPLC-ESI-Q-o-TOF MS/MS data obtained for the enzymatic digestion of BCA are given in Appendices F (AspN digestions) and G (GluC digestion). Theoretical mass calculations were determined using Protein Prospector as described in Chapter 2.6.¹⁸¹

The majority of peptides identified were of a 2+ and 3+ charge state, with some smaller peptides being 1+ and some larger peptides 4+. These results are summarised in the sequence coverage maps shown in Scheme 5.5 and Scheme 5.6. The sequence coverage maps for the AspN digestion of BCA show a large length of the C-terminal protein being undetected. This is due to the absence of any aspartic acid residues in this portion of the sequence. Digestion of perPEGylated BCA with GluC allowed for the analysis of this portion of the sequence. In addition to the presence of the acylations located on lysine residues, there were a number of additional features observed, including the deamidation of asparagines residues and the sub-stoichiometric addition of PEG to the BCA protein. The tandem mass spectrometric analyses of the resultant peptides are discussed in the sections below.

5.2.2.2 Comparison of Acylating Agents

The ESI-Q-o-TOF MS/MS spectra for BCA AspN enzymatic peptide 150-159 in its unmodified form and with three different acylating agents (acetyl, trifluoroacetyl and PEG) are shown in Figure 5.15 - Figure 5.18. For clarity each spectrum is spread over three lines. The spectra show that the fragmentation observed is very similar for the unmodified peptide and the three acylated peptides examined. A complete b-ion series

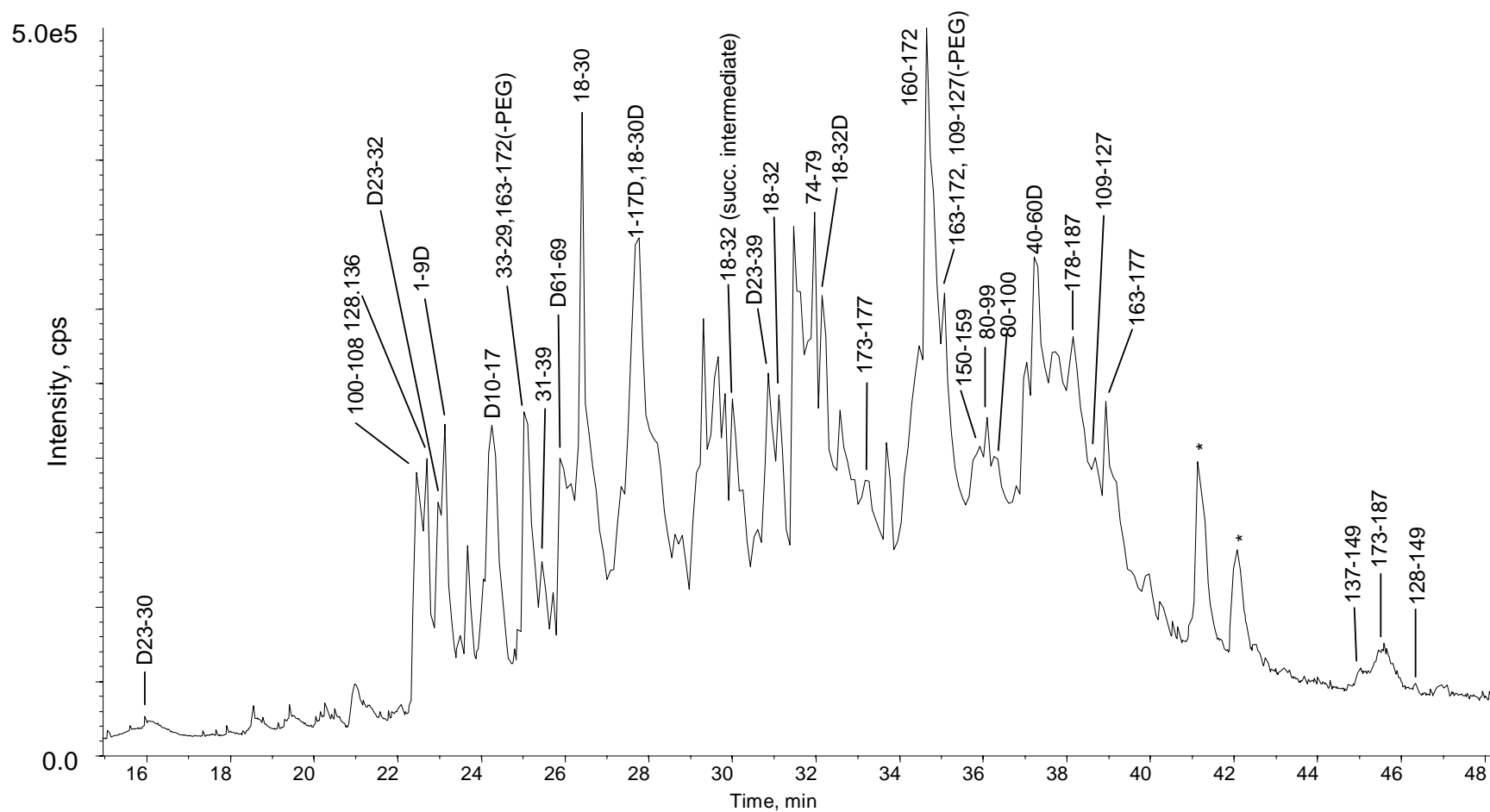
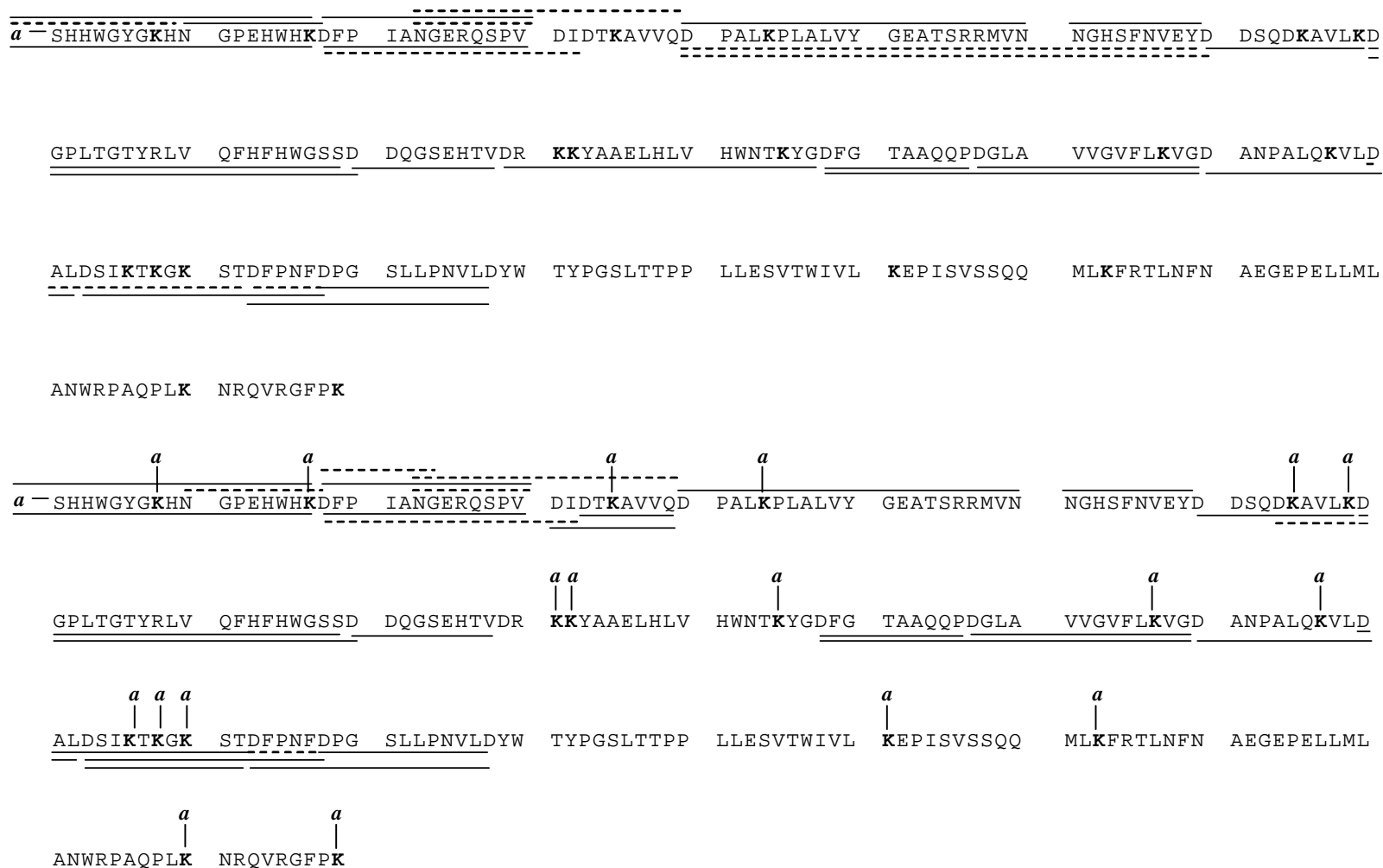
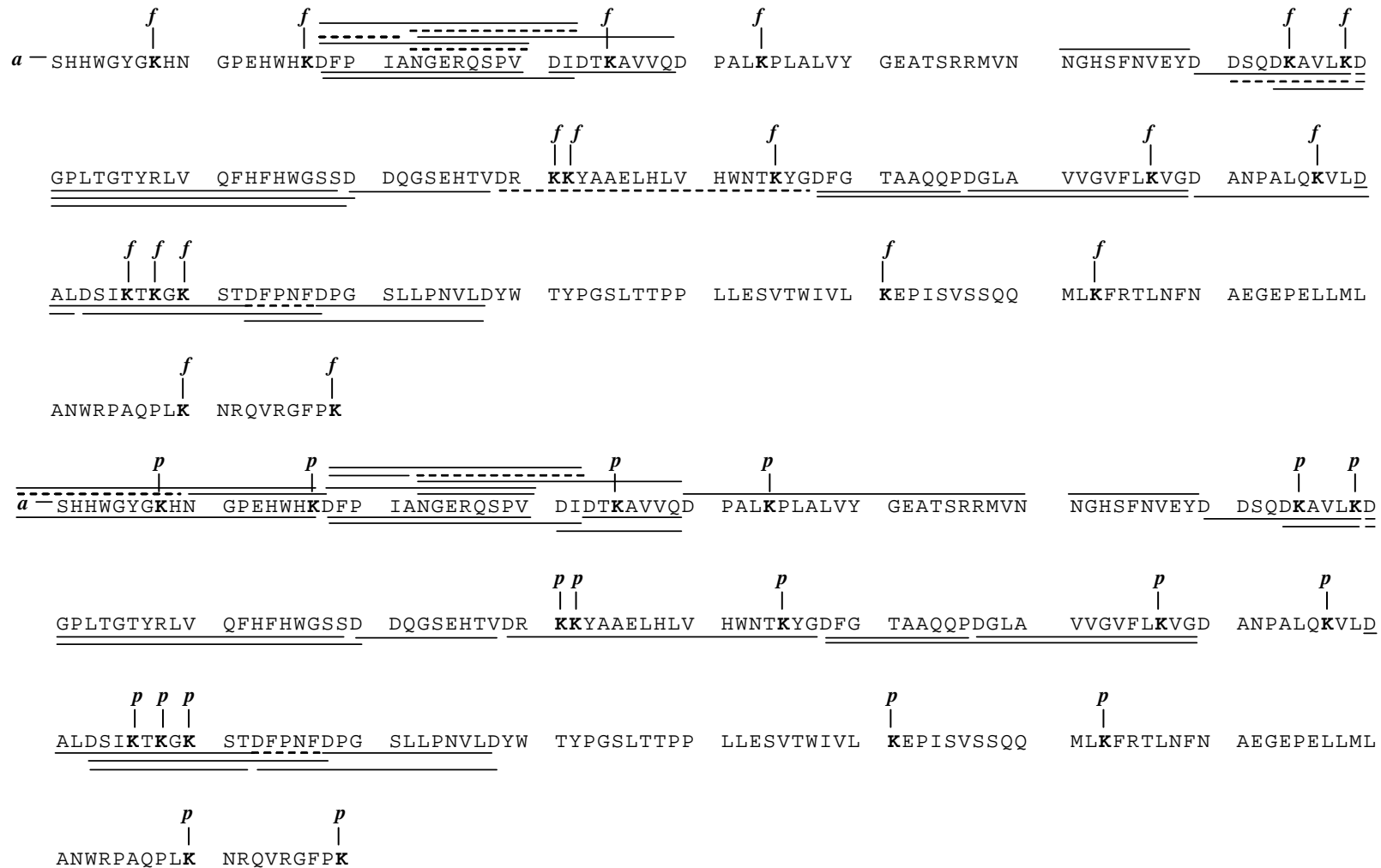


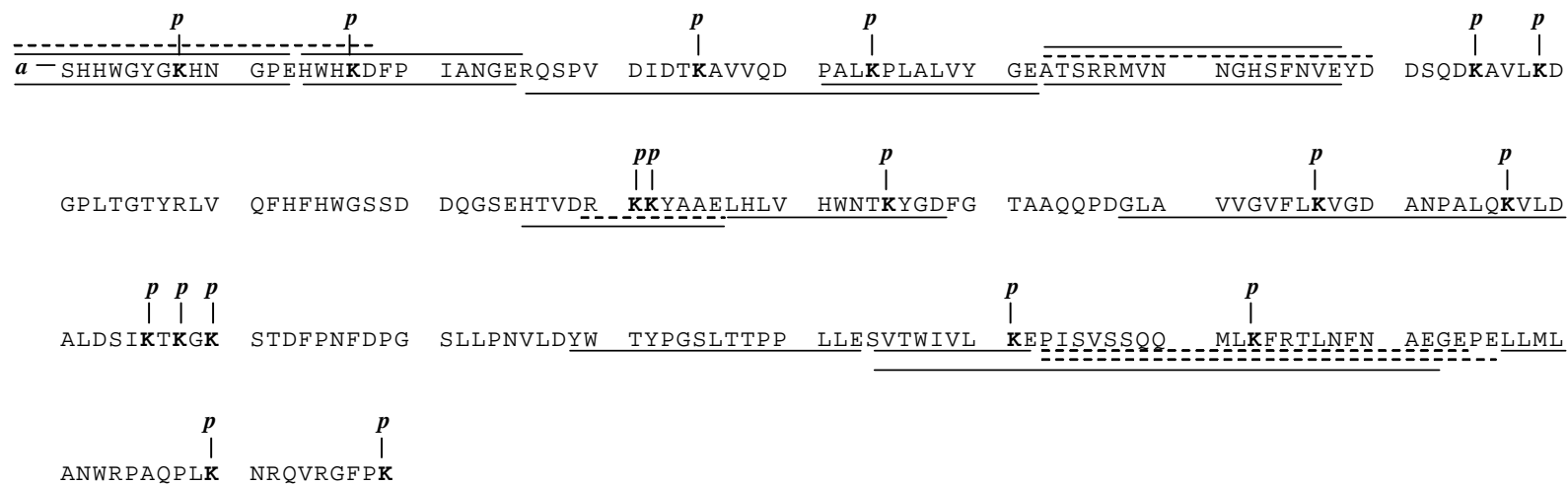
Figure 5.14 TIC for the HPLC-ESI-Q-o-TOF MS/MS analysis of AspN digested perPEGylated BCA. “D” indicates additional AspN enzymatic peptides containing a deamidated asparagine residue, see section 5.2.2.3 for details. (* indicates unidentified sequences.)



Scheme 5.5 Sequence coverage maps for the AspN digestion of perfunctionalised bovine carbonic anhydrase II. Where **a** indicates acetate addition. Lines below the sequence indicate sequence as written; lines above the sequence indicate peptides including a deamidated asparagine. Peptides analysed by MS/MS are indicated by an unbroken line; peptides detected by MS only are indicated by a dashed line. (Continued on next page.)



Scheme 5.5 cont. Sequence coverage maps for the AspN digestion of perfunctionalised bovine carbonic anhydrase II. Where *f* indicates trifluoroacetate and *p* indicates PEG addition. Lines below the sequence indicate sequence as written; lines above the sequence indicate peptides including a deamidated asparagine. Peptides analysed by MS/MS are indicated by an unbroken line; peptides detected by MS only are indicated by a dashed line.



Scheme 5.6 Sequence coverage maps for the GluC digestion of perfunctionalised bovine carbonic anhydrase II. Where *p* indicates PEG addition. Lines below the sequence indicate sequence as written; lines above the sequence indicate peptides including a deamidated asparagine. Peptides analysed by MS/MS are indicated by an unbroken line; peptides detected by MS only are indicated by a dashed line.

of b_{2-9} is observed, and an almost complete y sequence ion series of y_{1-8} is observed. These sequence ions confirm the presence of a modified lysine residue at lys-157. An amino-acylium ion for the unmodified lysine of m/z 129.10 is observed in the spectrum of the unmodified peptide, whilst acylated carbenium ions are observed in the spectra of the acylated peptides; these being an acetylated lysine carbenium ion at m/z 126.9, a trifluoroacetylated lysine carbenium ion at m/z 194.09, and a PEGylated lysine carbenium ion at m/z 302.20. These ions are the same as, or analogous to, those observed in the ESI-Q-o-TOF MS/MS analysis of acylated ubiquitin, in section 5.2.1 of this chapter. The contrast in the type of ions observed for native lysine (an amino-acylium ion, comprised of the entire amino acid residue) and acylated lysine (an acylated lysine carbenium ion) may be explained by the acyl moiety providing a stabilising influence, due to the increased mass of the residue, allowing for the formation of the carbenium ion containing less of the lysine residue.

The unmodified, acetylated, trifluoroacetylated and PEGylated BCA AspN enzymatic peptides 150-159 eluted at 28.3, 32.0, 36.0 and 35.8 minutes respectively. These elution times are consistent with a 9-mer peptide containing one modified lysine residue. The elution order is the same as that observed for acetylated and PEGylated ubiquitin enzymatic peptides. The trifluoroacetylated peptide elutes at a very similar time to that of the PEGylated peptide.

Analogous spectra were produced in the ESI-Q-o-TOF MS/MS of other enzymatic peptides for acylated BCA (spectra not shown). The data, showing the elution times and intact enzymatic peptide masses are given in Appendices G and F. Representative spectra are shown in the sections that follow.

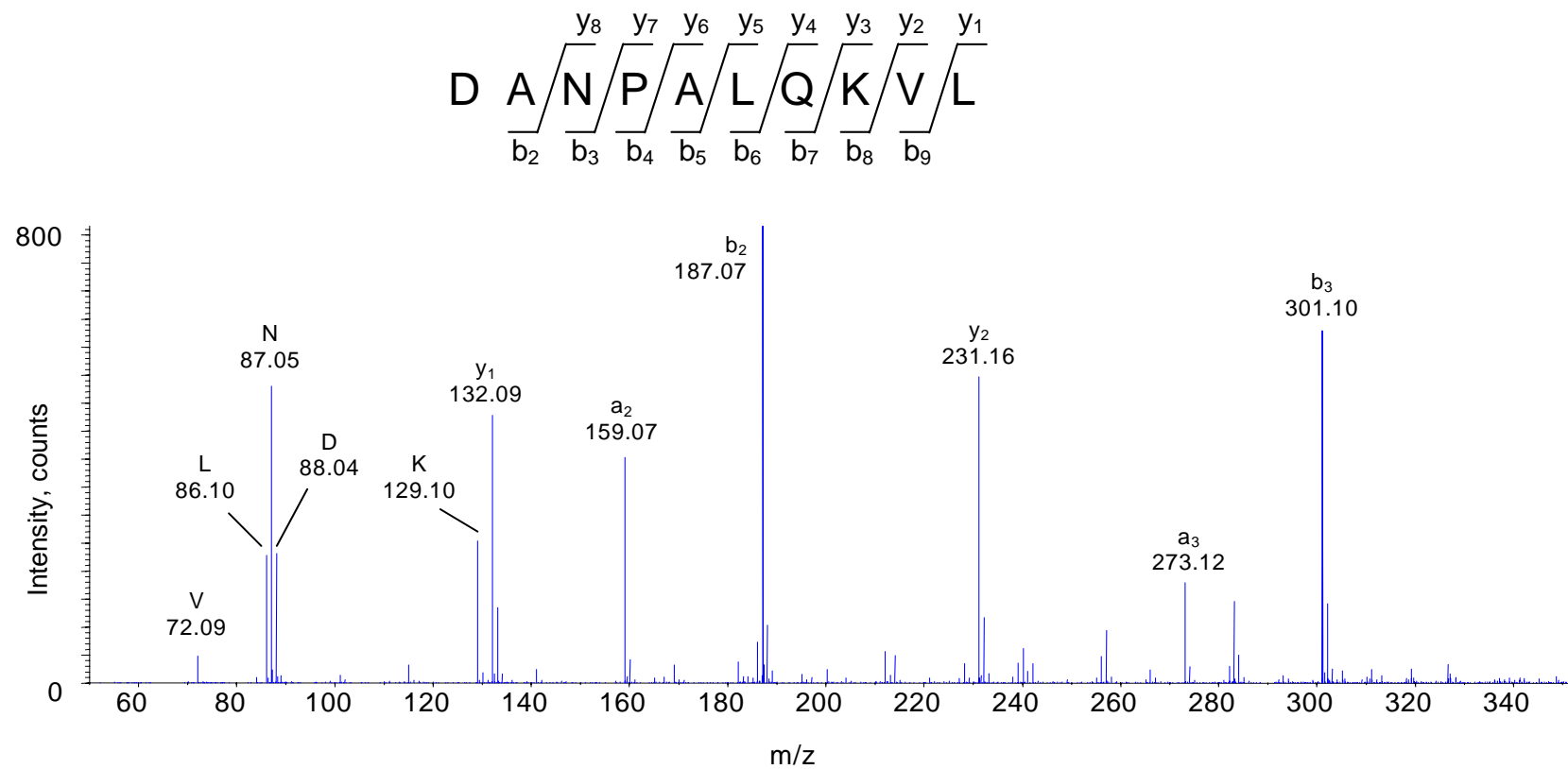


Figure 5.15 HPLC-ESI-Q-o-TOF MS/MS analysis of native BCA AspN enzymatic peptide 150-159. (Continued on next page.)

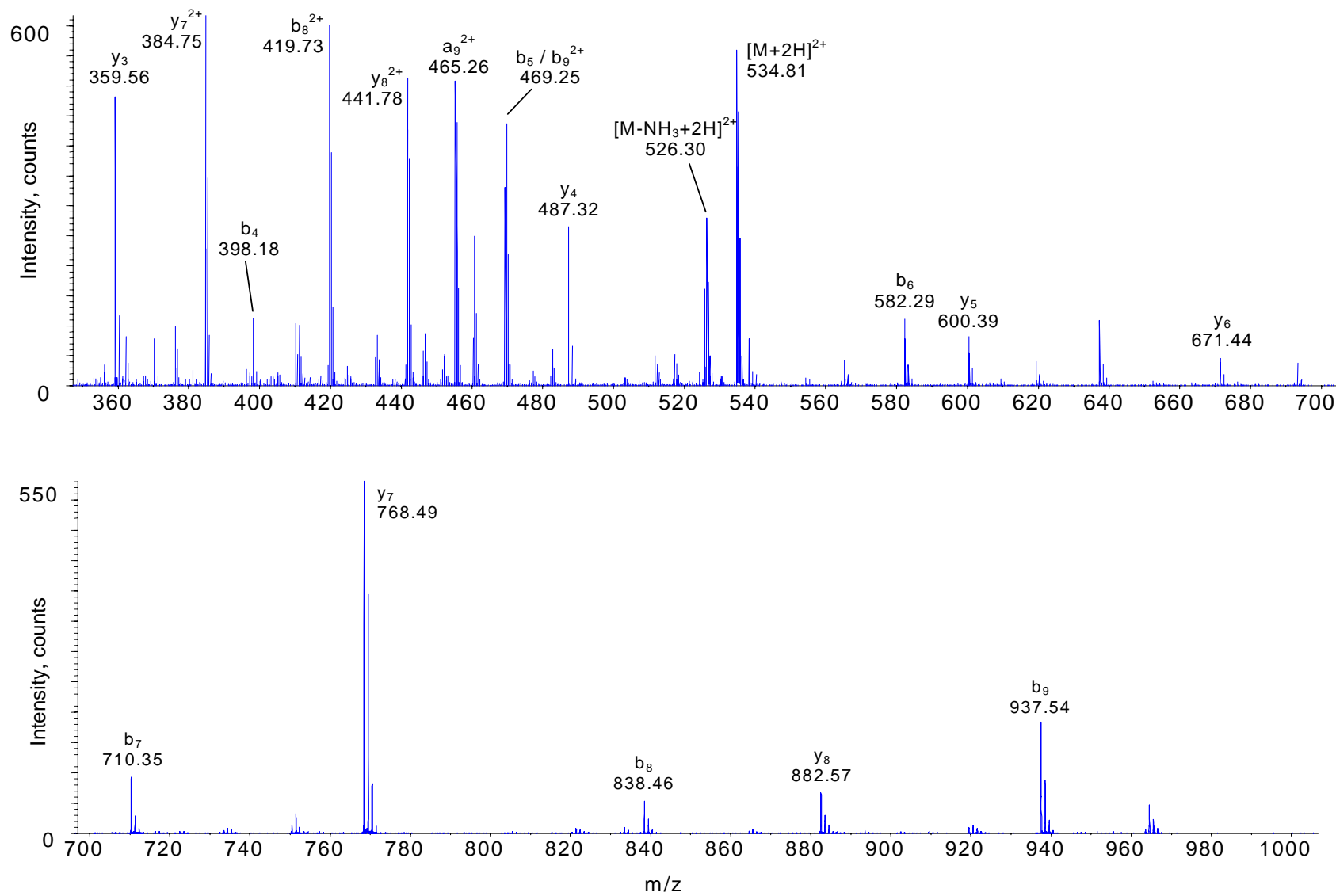


Figure 5.15 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of native BCA AspN enzymatic peptide 150-159.

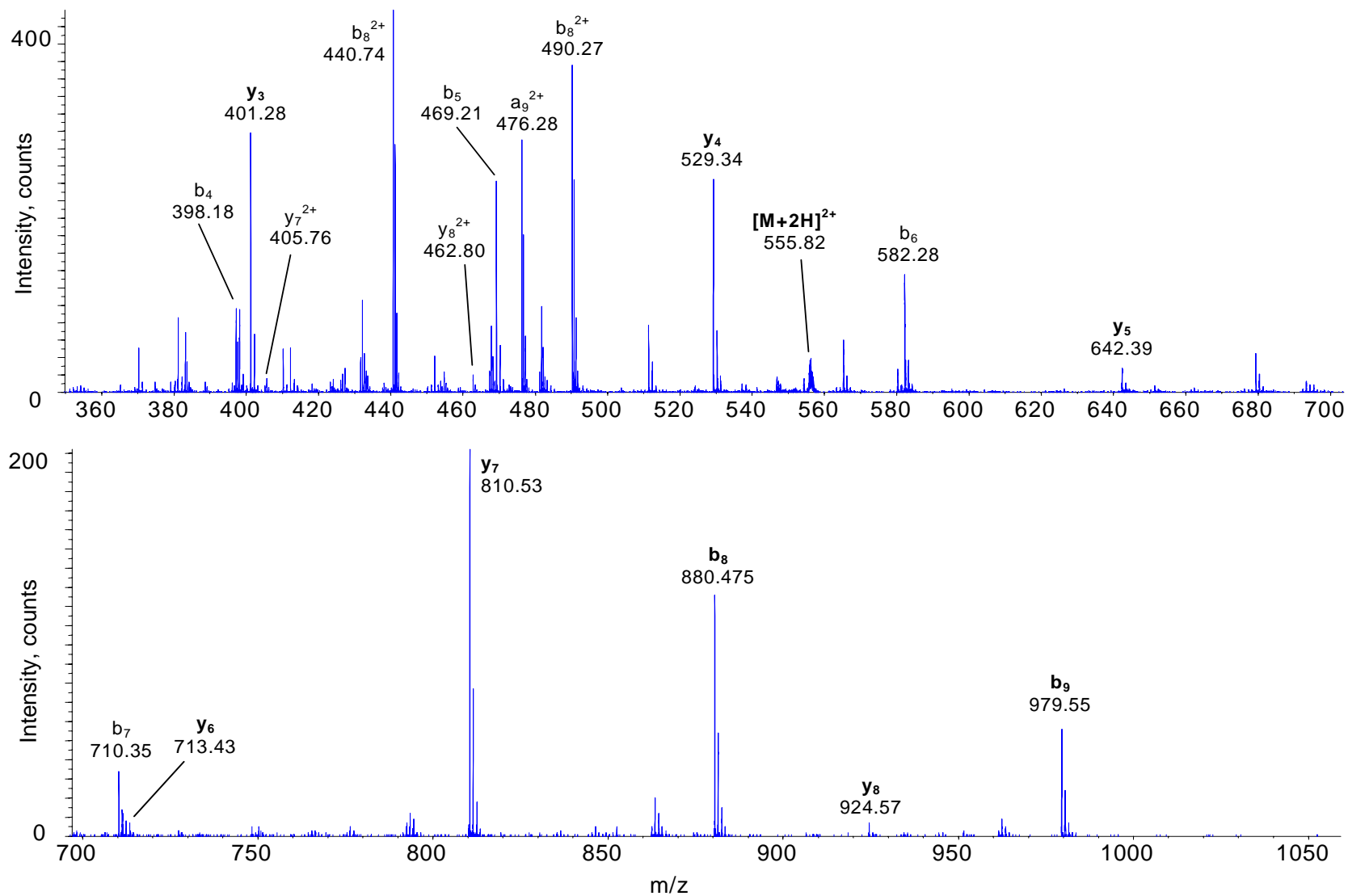


Figure 5.16 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of peracetylated BCA AspN enzymatic peptide 150-159. **a** indicates acetate; bold type indicates the presence of a product ion containing the acyl moiety.

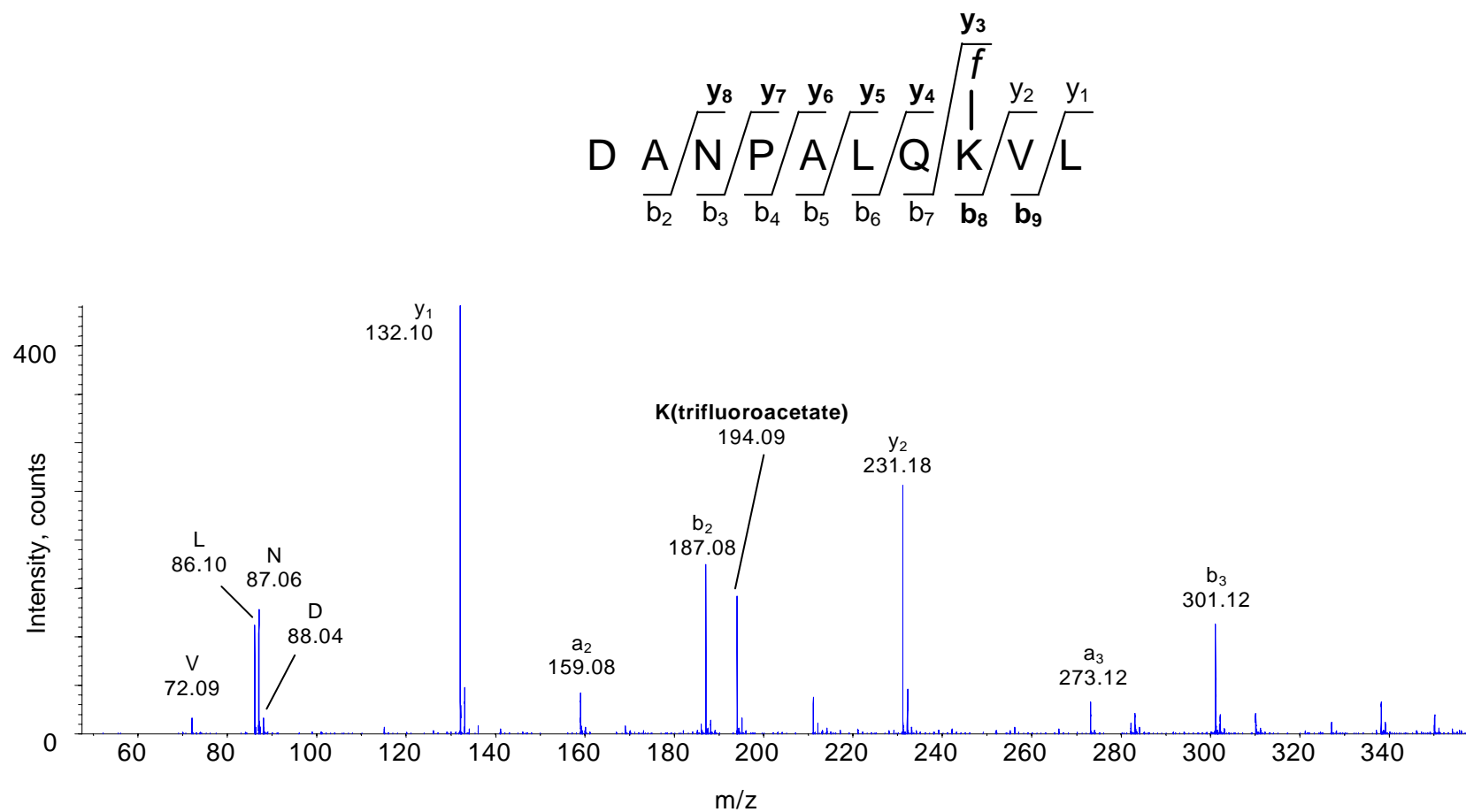


Figure 5.17 HPLC-ESI-Q-o-TOF MS/MS analysis of pertrifluoroacetylated BCA AspN enzymatic peptide 150-159. *f* indicates trifluoroacetate; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

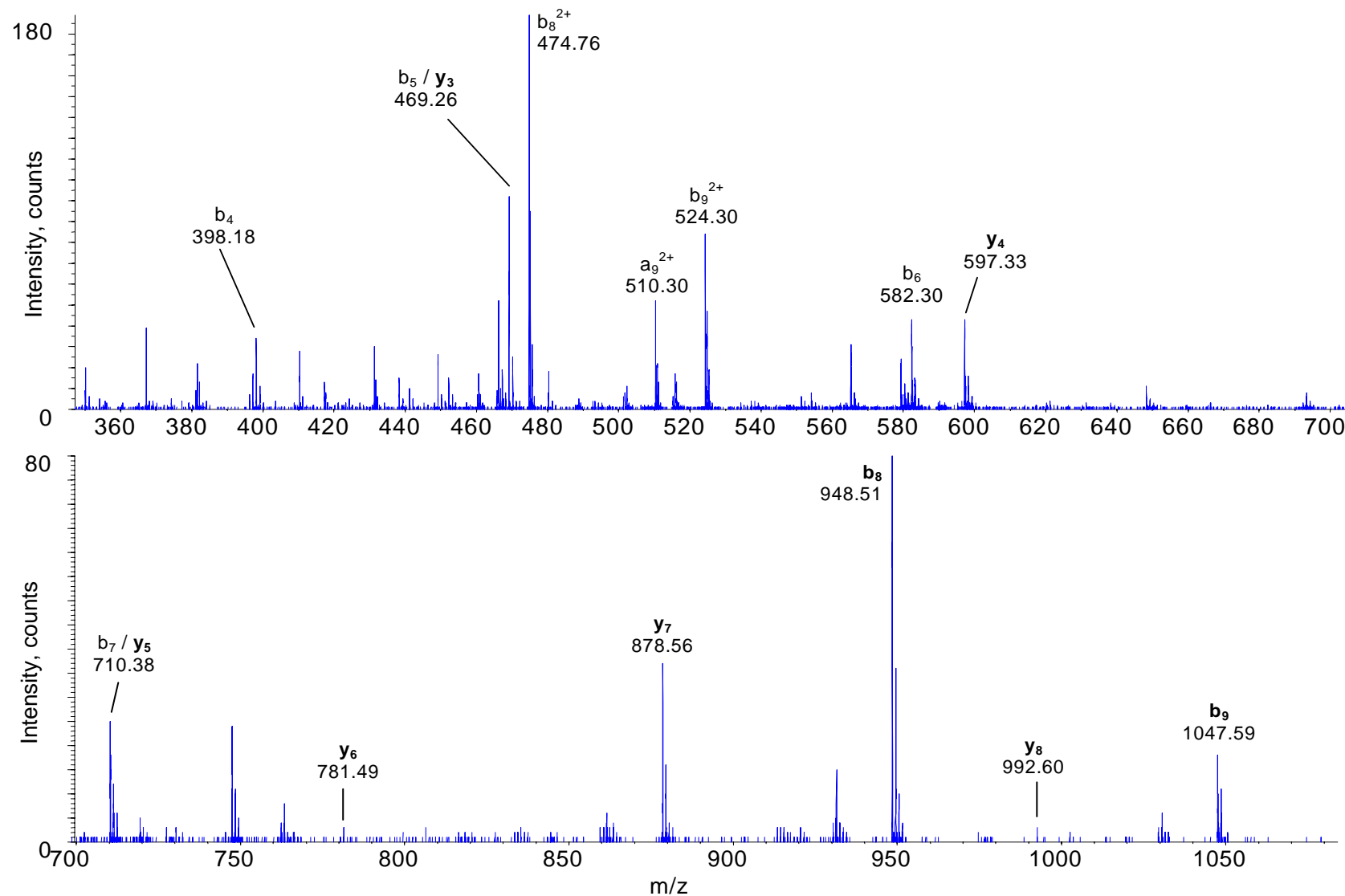


Figure 5.17 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of pertrifluoroacetylated BCA AspN enzymatic peptide 150-159. *f* indicates trifluoroacetate; bold type indicates the presence of a product ion containing the acyl moiety.

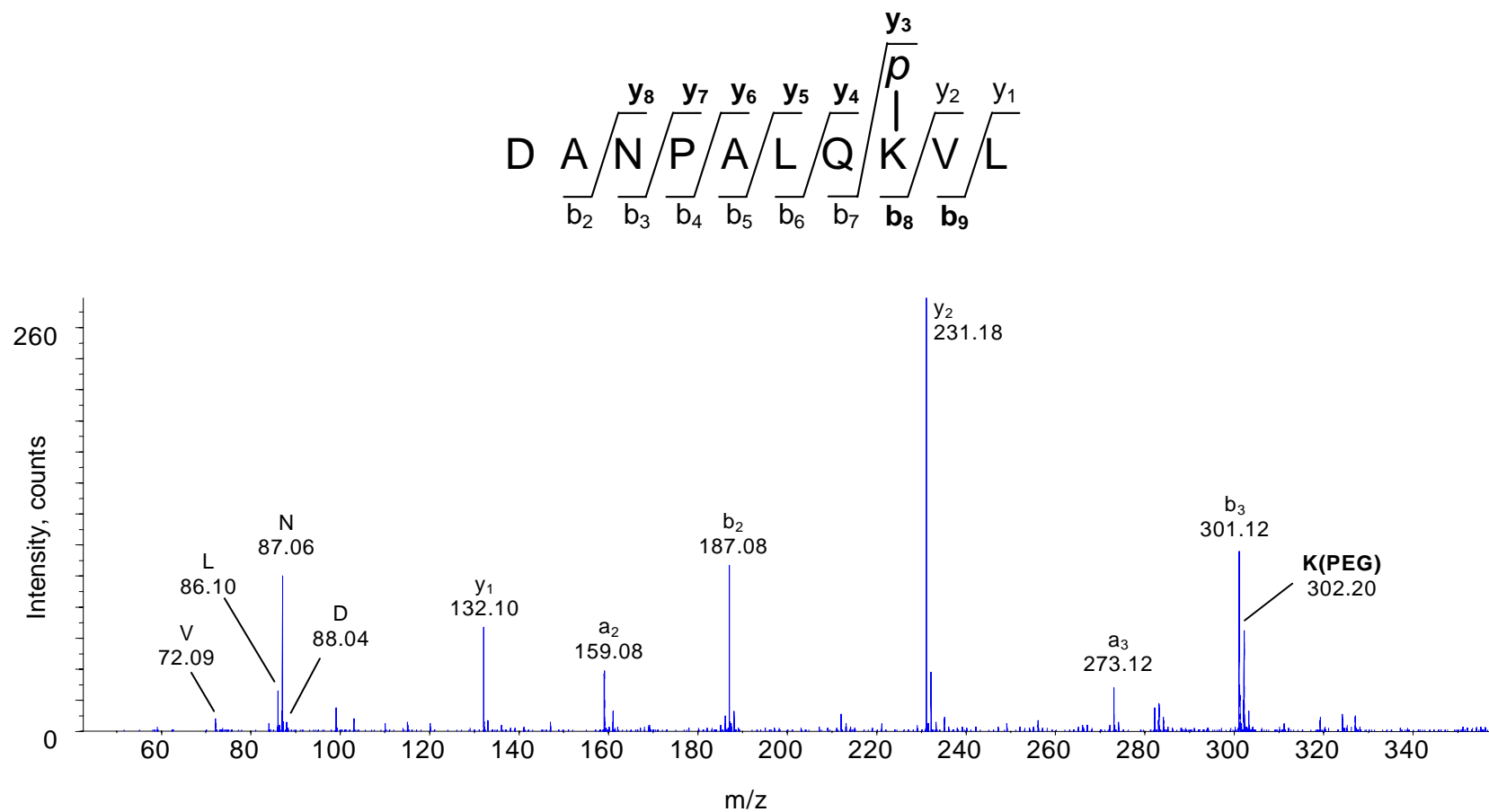


Figure 5.18 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 150-159. *p* indicates PEG; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

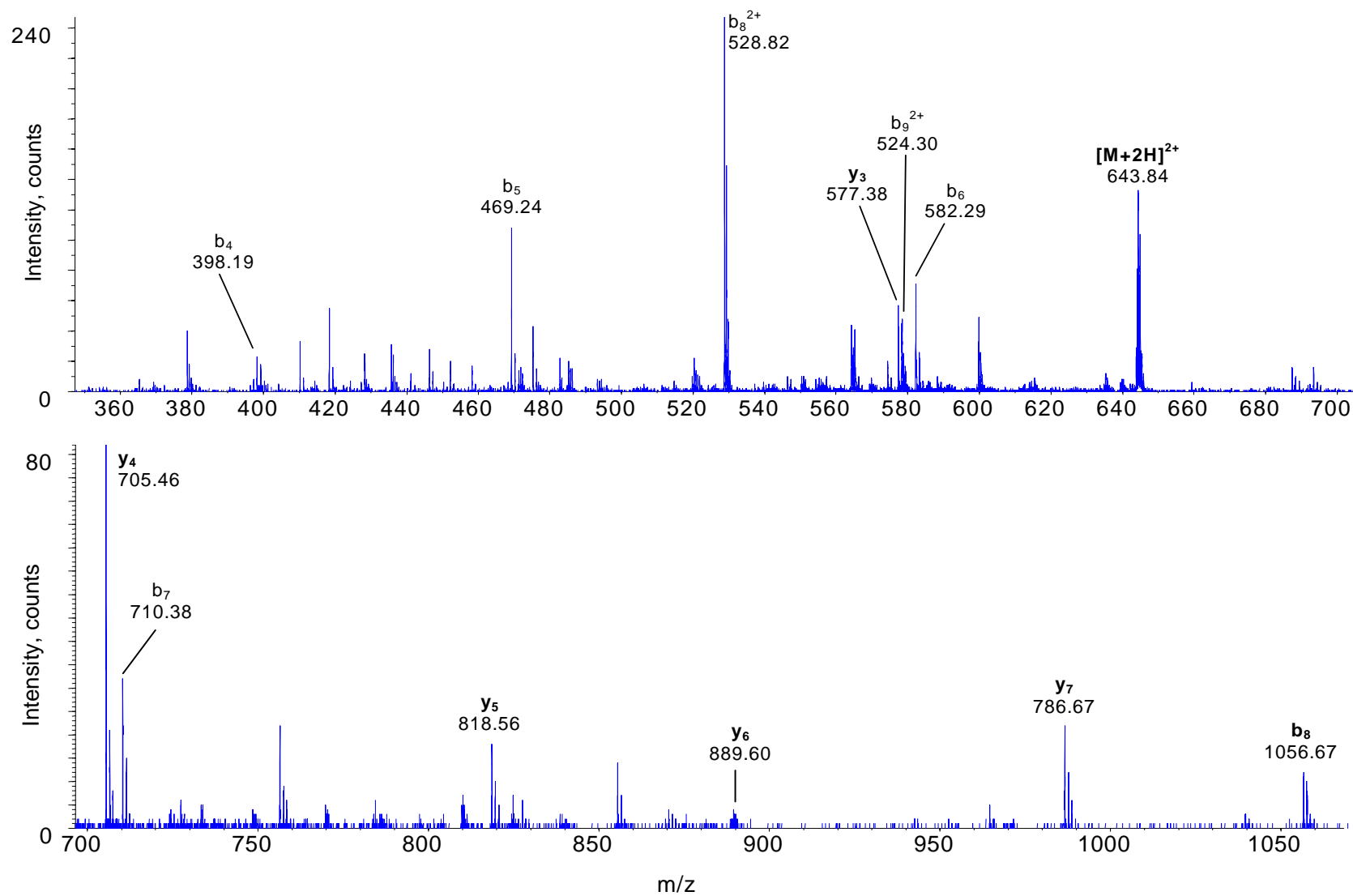
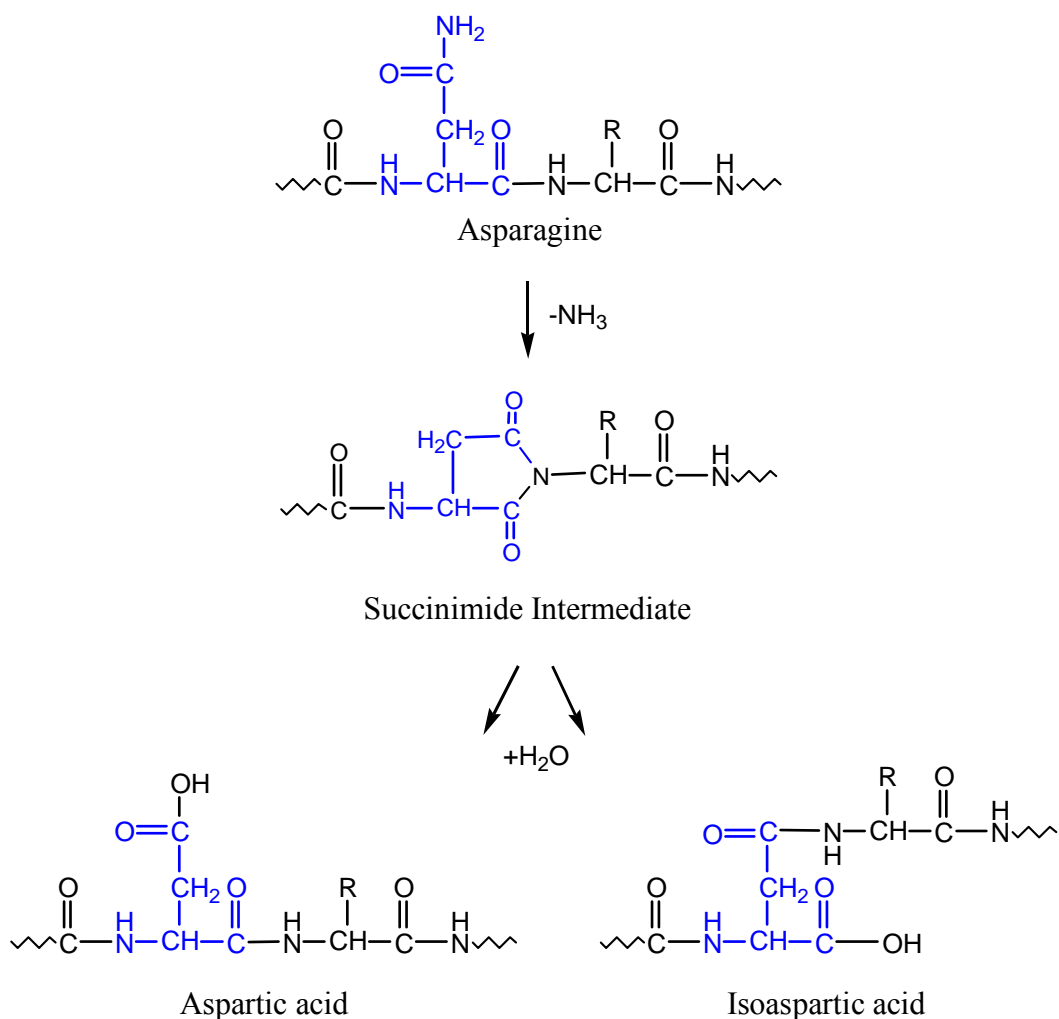


Figure 5.18 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 150-159. *p* indicates PEG; bold type indicates the presence of a product ion containing the acyl moiety.

5.2.2.3 Deamidation of Asparagine

The sequence coverage maps show a number of additional enzymatic cleavage products due to the deamidation of asparagine at Asn-10, Asn-23 and Asn-61. The presence of deamidated asparagine was noticed when analysing denatured and acylated BCA by capillary electrophoresis.³¹⁷ Capillary electrophoresis was being used to distinguish between native (folded) protein, and an aggregate of denatured protein and SDS based on mobility. In using capillary electrophoresis to check re-folded BCA (subsequent to being denatured), an additional charge peak was observed (data not shown). It was postulated that this peak may be due to a deamidated species, as it had one additional negative charge, relative to the native form of the protein, as determined from the electropherogram. Deamidation can occur at asparagine residues, which are converted to aspartic acid or isoaspartic acid via a cyclic succinimide intermediate.³²⁵⁻³²⁷ This process is shown in Scheme 5.7 below. Deamidation at asparagine occurs 30-50 times faster at asparagine residues followed by glycine, as is the case for asparagines 10, 23 and 61 in the sequence of BCA.^{328, 329} Deamidation is also likely to occur at an asparagine followed by a second asparagine.³³⁰ A succinimide intermediate can sometimes be observed during the deamidation process, resulting in a mass loss of 17 u. The final products in the process are aspartic acid and isoaspartic acid, resulting in a net mass gain of 1 u. Aspartic acid and isoaspartic acid cannot readily be distinguished by ESI-CID MS/MS, however, these isomers can be distinguished using electron capture dissociation, which allows for the cleavage of the peptide backbone via c-type ions.³³¹ The presence of an isoaspartic acid residue can also be detected using Edman degradation, where the additional methylene group in the peptide backbone prevents the continuation of the Edman sequencing procedure.³³² Isoaspartic acid can be chemically labelled using protein carboxyl methyltransferase which adds an additional methyl

group to the free alpha-carboxyl of the aspartic acid residue.³³³ (For ease of discussion and labelling, the two potential isometric products shall be referred to as aspartic acid, D, for the remainder of this chapter.)



Scheme 5.7 Deamidation of Asparagine, via a succinimide intermediate.

The presence of deamidated asparagines in a protein can be of critical importance in maintaining its biological activity,³³³ with the proximity of any deamidation to the active site being the most reliable indicator of whether the biological activity of a given protein will be affected.³³⁴ Deamidation has been associated with degenerative diseases such as Alzheimer's disease.³³⁵ If monodisperse polymers are to be utilised in a

therapeutic setting then they must be thoroughly characterised, and the impact of any modifications, such as deamidation, understood.

Figure 5.19 shows the perPEGylated BCA GluC enzymatic peptide 1-13. The ESI-Q-o-TOF MS/MS spectrum for the asparagine containing form of the peptide is shown in Figure 5.19 a). The ESI-Q-o-TOF MS/MS spectrum where asparagine (asn-10) has been deamidated to form aspartic acid (asp-10) is shown in Figure 5.19 b). The deamidation results in the presence of up to four additional peptides. One due to the new amino acid (aspartic acid), and two due to the new enzymatic cleavage products when digested using either GluC or AspN. A peptide can sometimes be observed for the succinimide intermediate, depending on its abundance in the sample. The ESI-Q-o-TOF MS/MS spectra for the new AspN peptides 1-9 and 10-17 are shown in Figure 5.20 below. An m/z value corresponding to the GluC enzymatic peptide 1-13 containing a succinimide intermediate was detected in the MS; however, no MS/MS data was obtained.

The ESI-Q-o-TOF MS/MS spectrum of BCA GluC enzymatic peptide 1-13 (Figure 5.19) shows sequence ions b_{1-12} , y_{1-4} , y_7 and y_{10-12} . An acetyl group is present at the N-terminus of the peptide which allows for the formation of a b_1 ion. The sequence ions and presence of an acylated lysine carbenium ion, K(PEG), at m/z 302.20 allow for the confirmation of an acylated lysine, at residue-8. The asparagine containing peptide eluted at 26.2 minutes, the deamidated peptide eluted at 26.9 minutes, and the ion indicative of a succinimide intermediate eluted at 29.7 minutes.

The ESI-Q-o-TOF MS/MS spectrum of PEGylated BCA AspN enzymatic peptide 1-9 shows almost complete sequence coverage, confirming both the presence of the acyl moiety at lys-8 and the new cleavage to the N-terminal side of an aspartic acid residue.

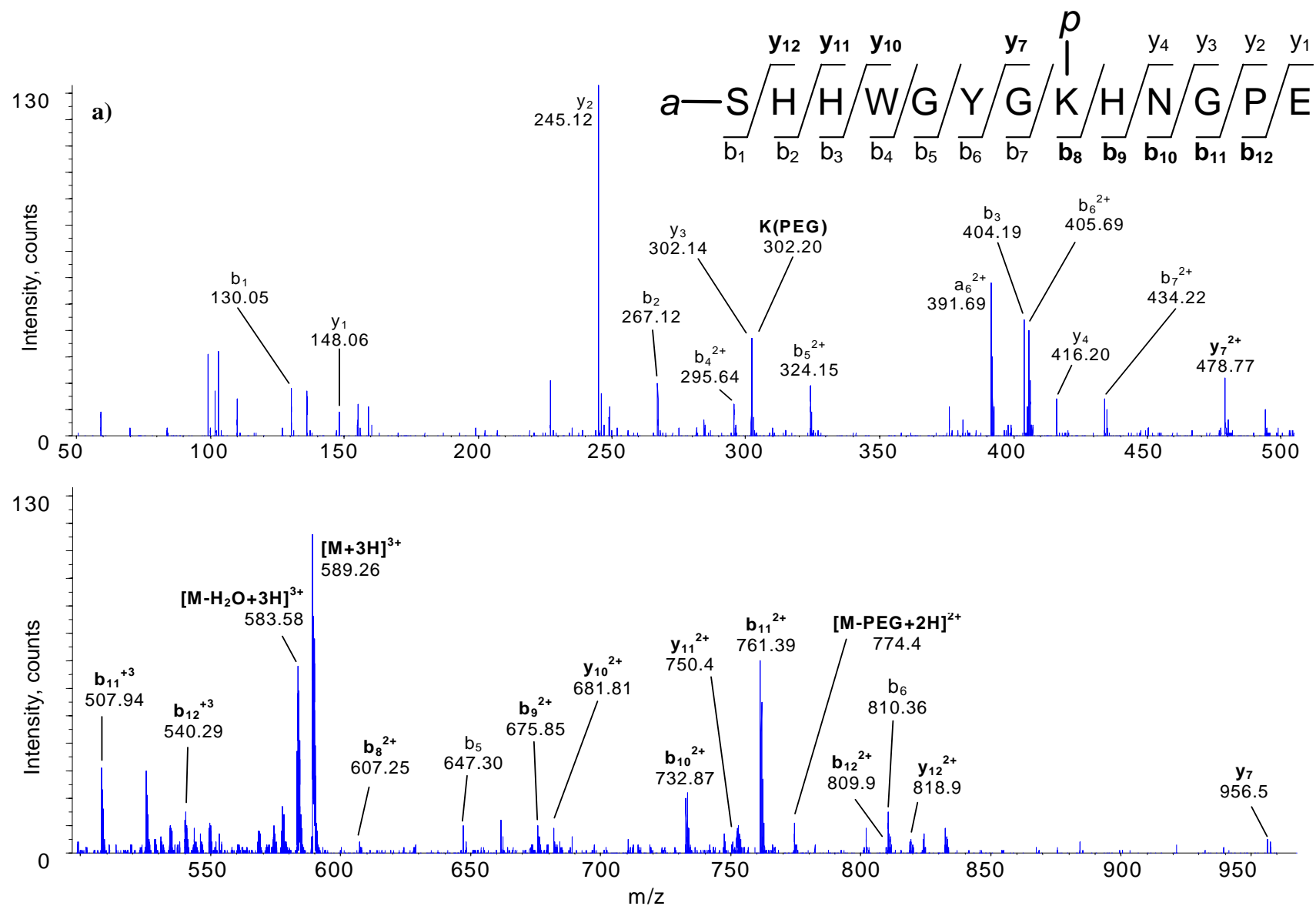


Figure 5.19 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA, GluC enzymatic peptide 1-13 a). *a* indicates an acetate group, *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety. (Continued on next page.)

The spectrum of AspN enzymatic peptide 10-17 shows good sequence coverage, allowing the identification of the aspartic acid residue, asp-10 (Figure 5.20 b)). The presence of PEGylated lys-17 is confirmed by a y_1 ion (m/z 365.25) and a PEGylated lysine carbenium ion (m/z 302.23). An AspN enzymatic peptide for the sequence 1-17, containing the native asparagine residue, asn-10, was also observed in the HPLC-ESI-Q-o-TOF MS/MS of the digested PEGylated BCA (spectrum not shown, data in Appendix F).

The second deamidated asparagine that was detected in the BCA sequence is at asn-23. Figure 5.21 shows the ESI-Q-o-TOF MS/MS of BCA AspN enzymatic peptide 18-30: in its native form; containing a succinimide intermediate, and containing asp-23. An additional peptide resulting from a new cleavage at asp-23 is shown in Figure 5.21 d).

The spectra for the native, succinimide containing and asp-23 containing peptides were found to be analogous to each other. The b sequence ions up to b_5 , and y sequence ions up to y_7 , were found to consist of equal m/z values for all three sequences. The b sequence ions, b_6 and beyond, and the y sequence ions, y_7 and beyond, were found to be -17 u and +1 u for the succinimide and asp-23 containing peptides respectively, relative to the native peptide. These data confirmed the deamidation of the peptide at asn-23. The native peptide eluted at 26.5 minutes, the deamidated peptide at 27.5 minutes and the succinimide intermediate at 29.0 minutes. The relationship between elution times is consistent with that observed for the peptides containing asn-10/asp-10. The ESI-Q-o-TOF MS/MS spectrum of the AspN enzymatic peptide 23-30 adds additional evidence for the presence of the deamidated residue at Asn-23/Asp-23.

The third deamidated asparagine found in acylated BCA was found to occur at asn-61. GluC enzymatic peptide 53-68 containing asn-61 or asp-61 were found to be present in

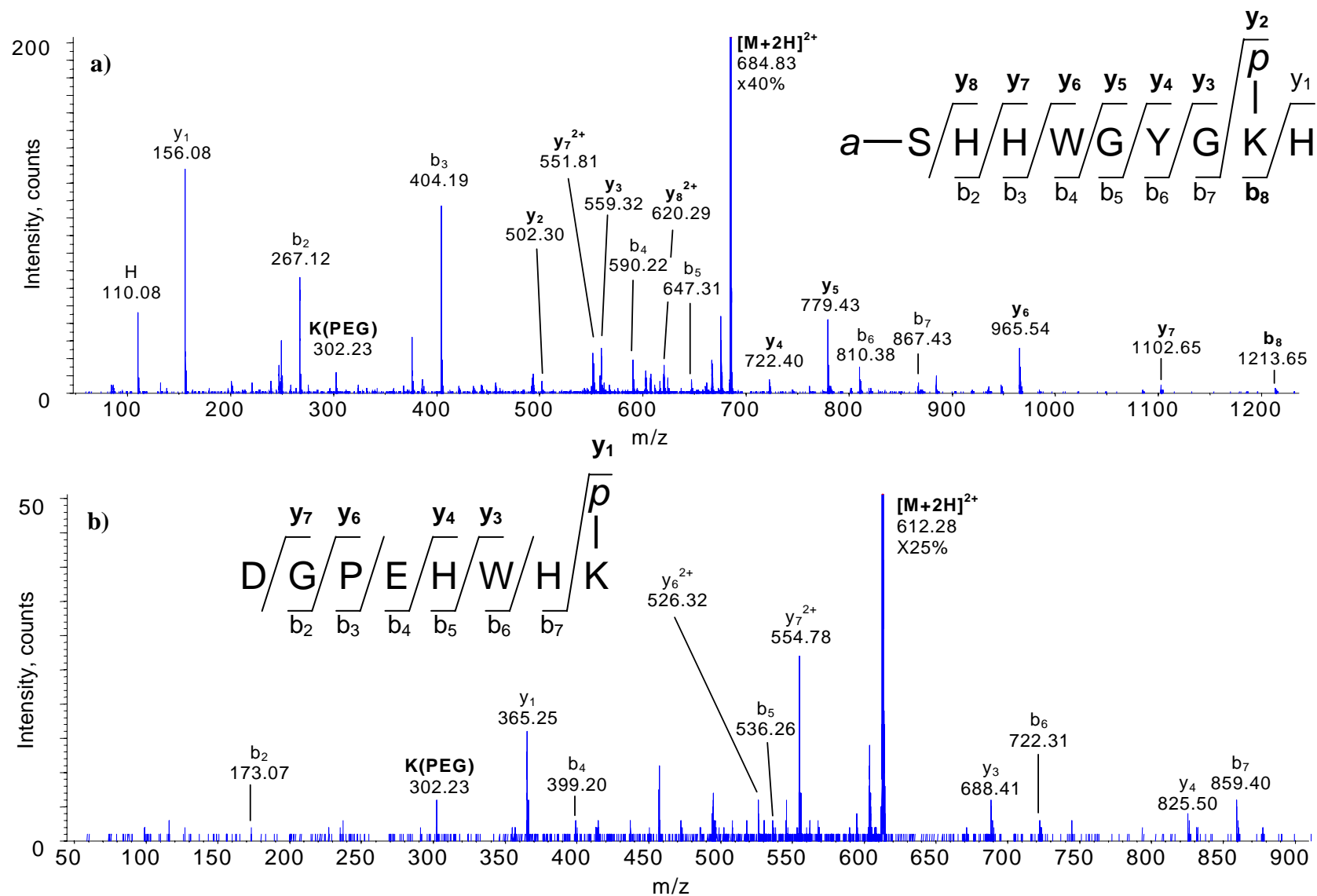


Figure 5.20 HPLC-ESI-Q-O-TOF MS/MS analysis of perPEGylated BCA, containing asp-10, AspN enzymatic peptide 1-9 a), containing asp-10, AspN enzymatic peptide 10-17 b). **a** indicates an acetate group, **p** indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

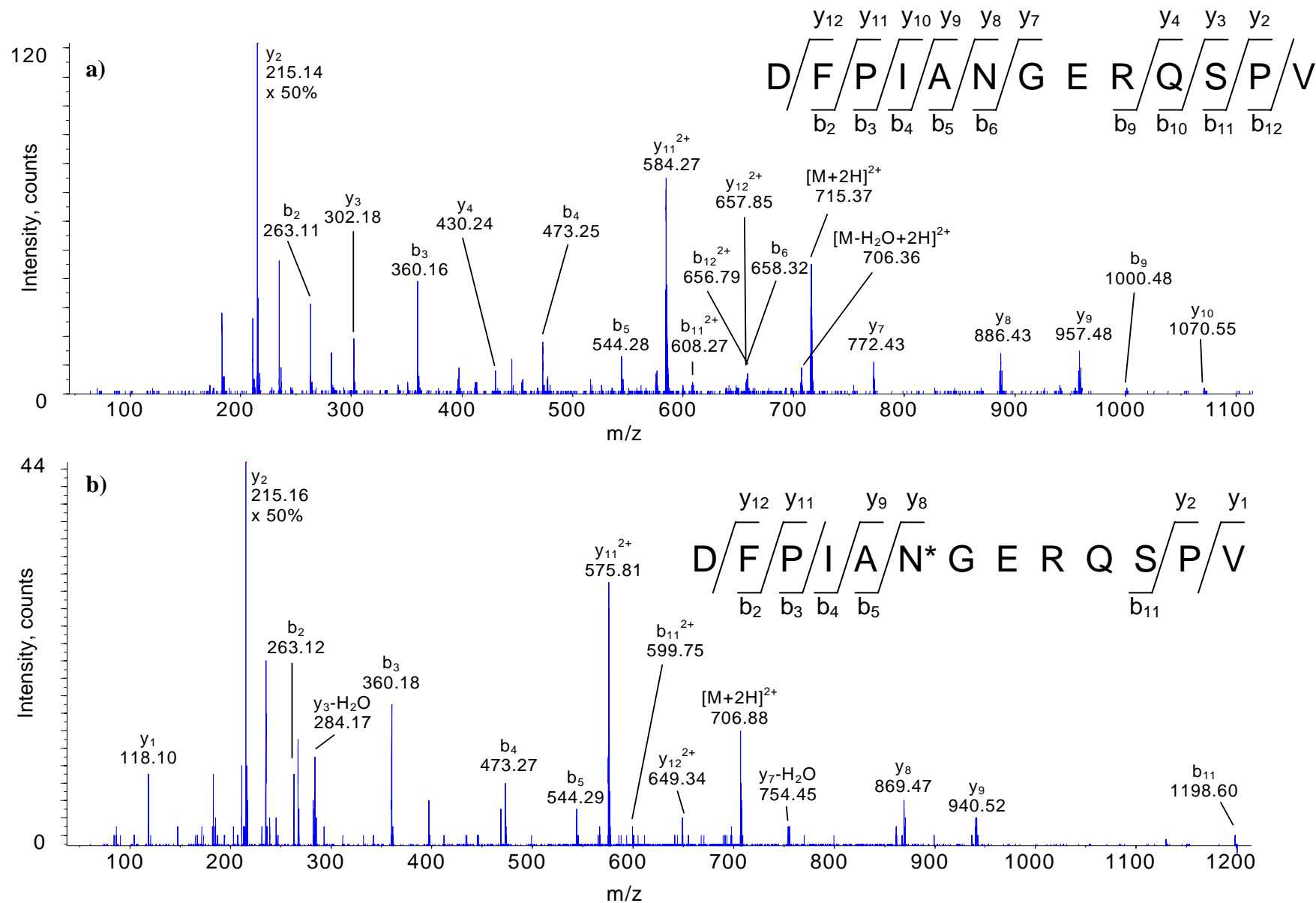


Figure 5.21 HPLC-ESI-Q-o-TOF MS/MS analysis of BCA AspN enzymatic peptide 18-30 a), BCA AspN enzymatic peptide 18-30 containing a succinimide intermediate at position-23 b). N* indicates the position of the succinimide intermediate. (Continued on next page.)

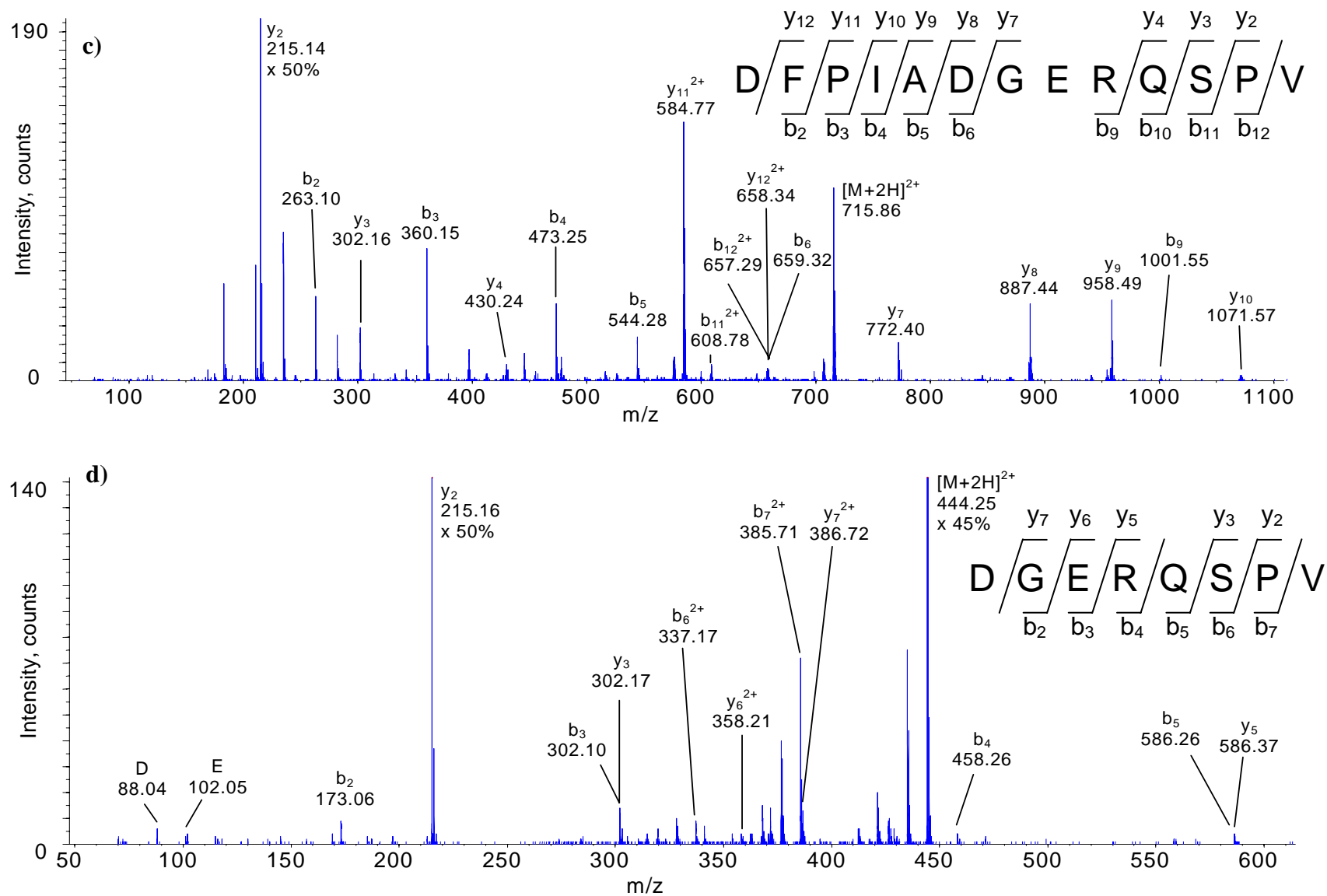


Figure 5.21 cont. BCA AspN enzymatic peptide 18-30 containing asp-23 **c)** BCA AspN enzymatic peptide 23-30 containing asp-23 **d)**

the digested perPEGylated BCA sample. Due to the close elution times of the asn-61 and asp-61 containing peptides, however, the ESI-Q-o-TOF MS/MS spectra of the asn-61 containing peptide (native) was contaminated with the asp-61 containing peptide (spectrum not shown). Examination of the spectra allowed for identification of the deamidated asparagine as being asn-61. An MS/MS spectrum was not obtained for the peptide containing a succinimide intermediate. The elution times were 23.4 minutes for the asn-61 containing peptide, 23.9 minutes for the asp-61 containing peptide and 24.0 minutes for the m/z consistent with the peptide containing a succinimide intermediate. This distribution of elution times is consistent with that observed for the peptides containing asn-10/asp-10 and asn-23/asp-23.

The AspN enzymatic peptide 61-69 resulting from a cleavage adjacent to an aspartic acid residue at residue-61 is shown in Figure 5.22. Whilst the sequence-ion coverage is incomplete, covering the C-terminal portion NVEY, the absence of any similar sequence pattern within the BCA molecule, in addition to an accurate precursor mass (m/z 534.23, 2+, see Appendix F), gives a good indication that this peptide is correctly identified. There was also MS evidence for the formation of BCA AspN enzymatic peptide 40-60 (m/z 1260.18, 2+, see Appendix F).

Asn-61 is preceded by an asparagine group and followed by a glycine group. According to Chelius *et. al.*, in addition to asparagines followed by a glycine residue, deamidation is also likely to occur at an asparagine followed by a second asparagine.³³⁰ In this particular sample there was no evidence for deamidation at asp-60, either through MS/MS of the deamidated peptide (AspN 53-68) or through additional enzymatic cleavage products. The new AspN cleavage product 61-69 confirms conversion of asn-61 to asp-61.

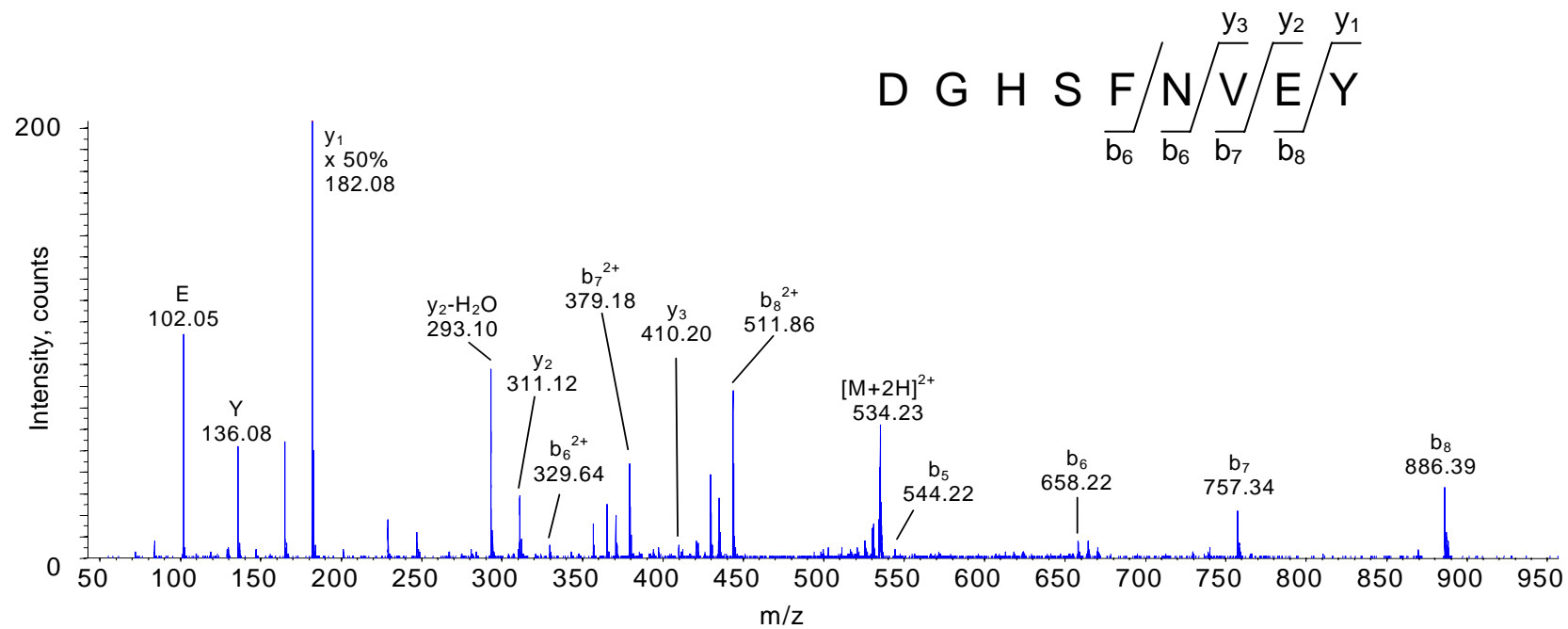


Figure 5.22 HPLC-ESI-Q-o-TOF MS/MS analysis of BCA AspN enzymatic peptide 61-69 containing asp-61.

It is not easy to distinguish between the aspartic acid and isoaspartic acid forms of the deamidated asparagine residue using collisionally induced dissociation (CID) MS/MS.³³¹ One means of detecting the presence of isoaspartic acid in a peptide using CID is by comparison of the b/y ion intensity ratio to the equivalent peptide containing an aspartic acid residue, with a lower ratio being detected for peptides containing an isoaspartic acid residue.³³⁶ The presence of enzymatic cleavage peptides as a result of incubation with AspN (which cleaves the N-terminal side of aspartic acid residues), however, indicates that a detectable proportion of the deamidation product is of the aspartic acid form.³³⁷ Chelius *et. al.* were able to separate isoaspartic acid from aspartic acid containing tryptic peptides of the conserved regions of human immunoglobulin gamma antibodies using reversed-phase HPLC. In their work a RP-HPLC gradient of *ca.* 0-60 % acetonitrile (0.1 % TFA) was run over a period of 195 minutes. This allowed for a separation of the isoaspartic acid and aspartic acid containing peptides by 1-3 minutes, assuming that the isoaspartic acid form elutes sooner than the aspartic acid form.³³⁸ In the work presented here, a gradient of *ca.* 3-60 % acetonitrile (0.1 % formic) was run over a period of 60 minutes, which was insufficient to allow for the separation of the two isoforms. Separation of the native peptide, succinimide intermediate containing peptides and a broad peak attributed to peptides containing both isoforms of aspartic acid were observed in the TIC of the enzymatically digested acylated BCA samples (see Figure 5.14).

Subsequent to the completion of this research, three dissociation techniques (using FTICR or ion-trap mass analysers) have been utilised for the facile detection of aspartic acid and isoaspartic acid residues. Using the techniques of electron capture dissociation (ECD),³³¹ electron ionisation dissociation (EID)³³⁷ and electron transfer dissociation

(ETD)³³⁹ the fragment ions c^{+58} and z^{*-57} were only found to be present in the spectra of the peptides containing isoaspartyl residues.

5.2.2.4 Sub-stoichiometric PEG addition

Two PEGylated BCA lysine residues were found to be sub-stoichiometrically modified, being lys-75 and lys-166. Figure 5.23 shows the ESI-Q-o-TOF MS/MS spectra for the perPEGylated and sub-stoichiometrically PEGylated BCA AspN enzymatic peptide 74-79. The spectrum for the perPEGylated peptide shows complete b and y product ion series, clearly showing the addition of two PEG moieties on lys-75 and lys-79. The peptide containing a sub-stoichiometric PEG addition was present in a small percentage of the protein molecules. The MS/MS obtained for this peptide was not of a high quality. Sufficient data were obtained, however, to enable the native lysine residue to be identified as lys-75. A small y_1 ion (containing PEGylated lys-79) can be seen in the spectrum. There was also a very small peak at m/z 527.3 (not shown) corresponding to the b_5 ion with an unmodified lys-75. The two peptides are well resolved chromatographically: the perPEGylated peptide had an elution time of 31.9 minutes, while the monoPEGylated peptide had an elution time of 23.2 minutes (see Appendix F). This is consistent with the results obtained for sub-stoichiometrically acylated ubiquitin, as discussed in section 5.2.1.3 above. There was no evidence that remaining acyl groups used to perfunctionalise BCA lysine residues (acetyl and trifluoroacetyl) resulted in a sub-stoichiometric functionalisation at lys-75.

Figure 5.24 shows the ESI-Q-o-TOF MS/MS spectra for the perPEGylated and sub-stoichiometrically PEGylated BCA AspN enzymatic peptide 163-172. A large number of sequence product ions are present in the spectra of both the per- and sub-stoichiometrically PEGylated peptides. The ESI-Q-o-TOF MS/MS spectrum for the

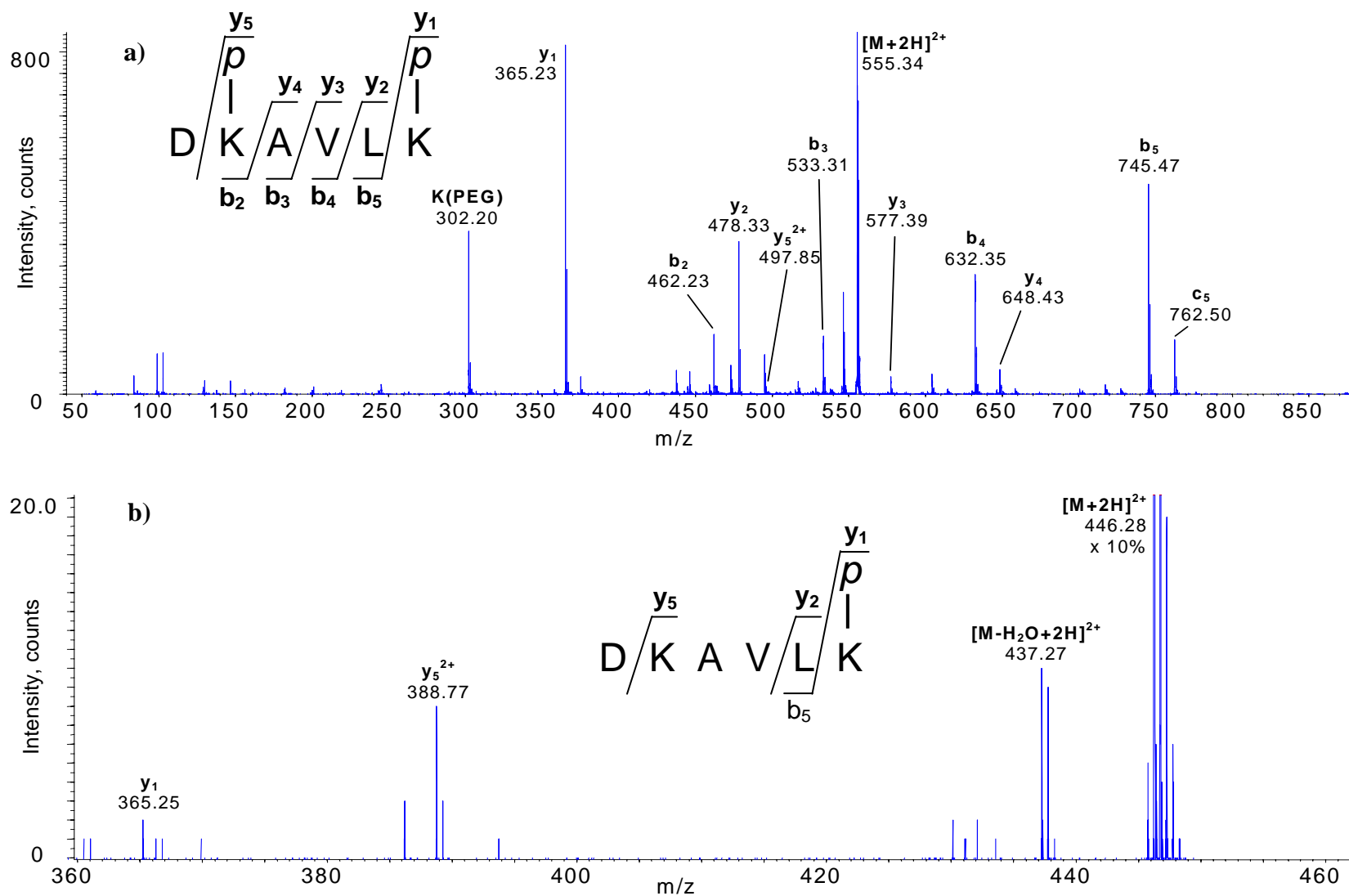


Figure 5.23 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 74-79 a) and sub-stoichiometrically PEGylated BCA Asp N enzymatic peptide 74-79 b). *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

perPEGylated peptide shows the sequence ions b_{2-9} and y_{1-7} , while the ESI-Q-o-TOF MS/MS spectrum for the sub-stoichiometrically modified peptide shows the sequence ions b_{2-6} , b_{8-9} and y_{1-9} . In addition, the MS/MS spectrum for the sub-stoichiometrically modified peptide shows a lysine amino-acylium ion at m/z 129.08. This lysine amino-acylium ion, at m/z 129, was not observed in the MS/MS spectra of any of the peracylated peptides studied in this work. Both of the peptides contain the PEGylated lysine carbenium ion at m/z 302.30. This portion of the BCA sequence contains three lysine residues: lys-166, lys-168 and lys-170. All three lysine residues are shown to be modified in the perPEGylated peptide. The lysine found to be unmodified in the diacylated peptide was lys-166: **DSIKTK(PEG)GK(PEG)ST**. The position of this unfunctionalised lysine is in contrast to the lysine that was found to be unmodified in the case of ubiquitin (section 5.2.1.3 above), where, in the PEGylated ubiquitin GluC enzymatic peptide 25-34, lys-29 was found to be the residue most likely to be unmodified: **NVK(PEG)AKIQDK(PEG)E**. It was postulated that this may have been due to steric hindrance if residues lys-27 and lys-33 reacted prior to lys-29. In the case of PEGylated BCA AspN enzymatic peptide 163-172, which is part of a much larger protein, there may have been local steric effects due to the conformation of the protein, prior to reaction with the acylating agent, causing the sub-stoichiometric reaction at this site. The perPEGylated peptide eluted at 35.1 minutes, in comparison to the diPEGylated peptide which eluted at 25.5 minutes. This difference in elution time can be attributed to the difference in hydrophobicity contributed to by the presence of the PEG moiety. This result is consistent with that observed for the per- and mono-PEGylated BCA AspN enzymatic peptide 75-79. There was no evidence that remaining acyl groups used to perfunctionalise BCA lysine residues (acetyl and trifluoroacetyl) resulted in a sub-stoichiometric functionalisation at lys-75.

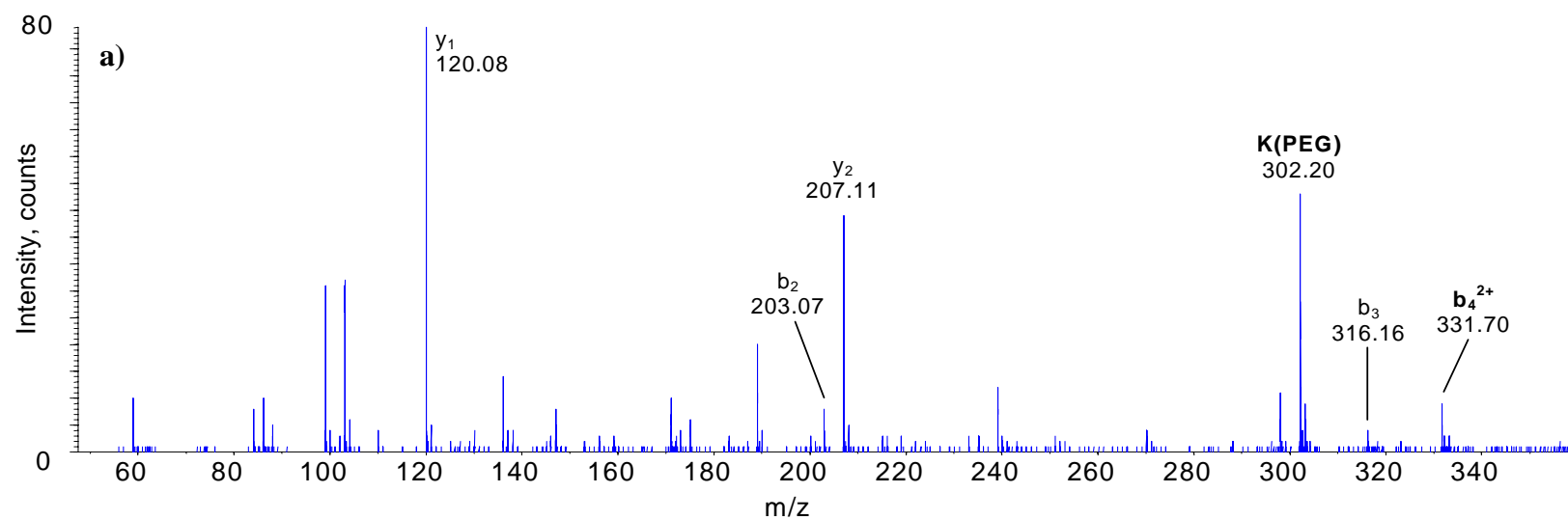
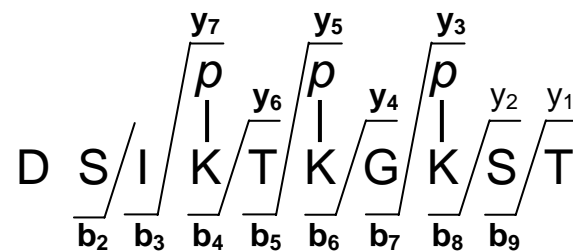


Figure 5.24 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 163-172 a). (Spread over three lines for clarity; Continued on next page.)

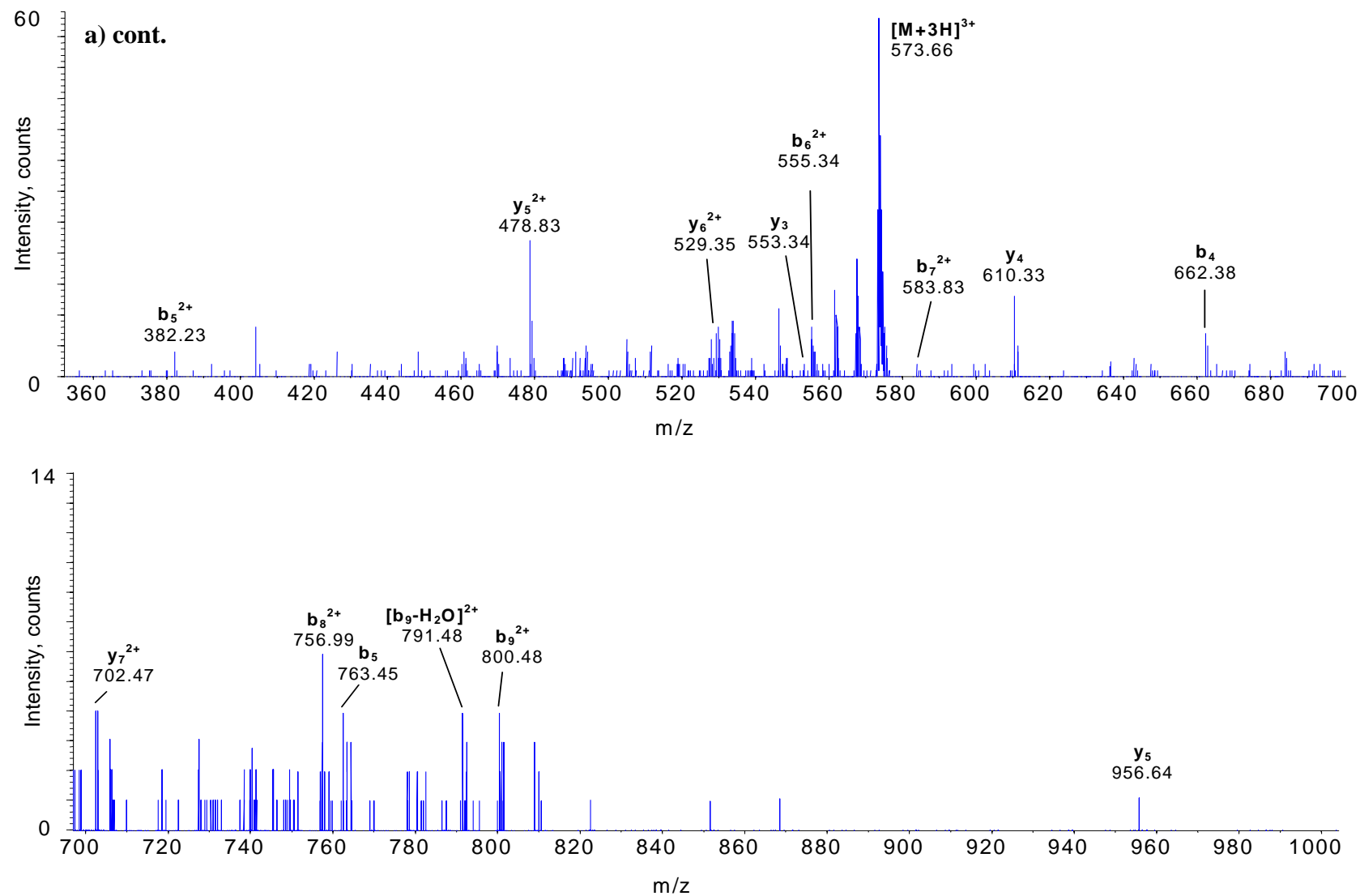


Figure 5.24 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 163-172 a). (Continued on next page.)

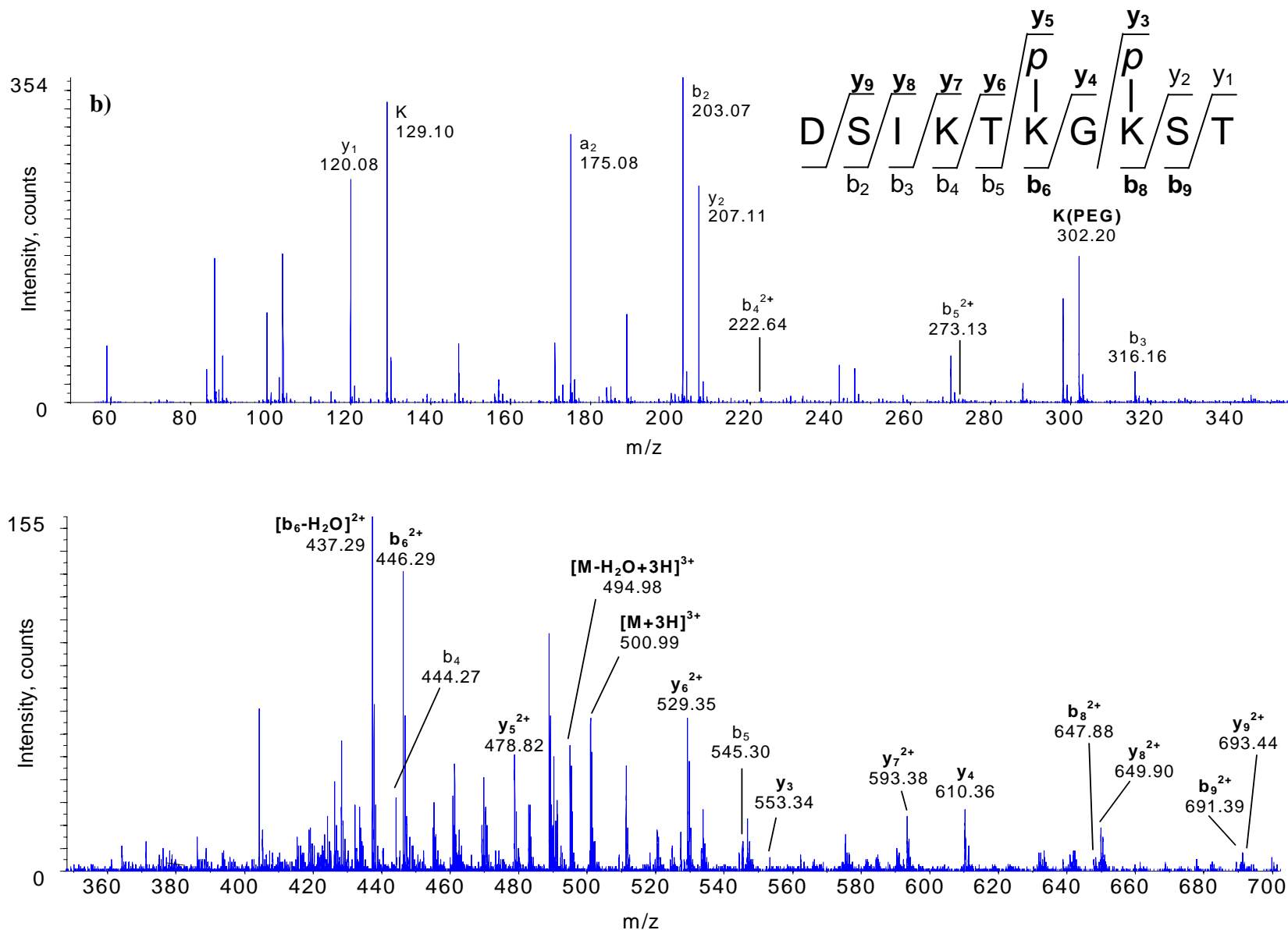


Figure 5.24 cont. HPLC-ESI-Q-o-TOF MS/MS analysis of substoichiometrically PEGylated BCA Asp N enzymatic peptide 163-172 b). *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety. (Spread over two lines for clarity.)

5.2.2.5 Lysine acylations: lys-35; lys-44; lys-111; lys-112; lys-125; lys-147; lys-211; lys-223; lys-250 and lys-259

The enzymatic peptides discussed below contain acyl additions, without extraneous changes to the protein. The examples shown are for AspN and GluC enzymatic peptides of PEGylated BCA. AspN enzymatic peptide 31-39 of BCA contains one acylated lysine residue (lys-35). The ESI-Q-o-TOF MS/MS of the perPEGylated BCA peptide is shown in Figure 5.25. The spectrum shows a complete b-ion series, b_{1-8} , and a partial y-ion series, y_{1-5} . A PEGylated lysine carbenium ion is also present at m/z 302.20. These sequences and carbenium ion indicate the presence of a PEGylated lysine residue at lys-35. The peptide elutes at 25.5 minutes which is towards the beginning of the elution profile (see Figure 5.14). This elution time is consistent with the length of this enzymatic peptide of BCA, being a 9-mer and containing one PEG moiety.

GluC enzymatic peptide 41-52 of BCA also contains one acylated residue (lys-44). The ESI-Q-o-TOF MS/MS of the perPEGylated BCA peptide is shown in Figure 5.26. The spectrum shows sequence ions b_{7-11} and y_{1-4} along with a PEGylated lysine carbenium ion at m/z 302.20. These data support the presence of a PEGylated lys-44. This peptide eluted at 41.5 minutes. This is approximately at the mid-point in the elution profile (TIC not shown for the GluC digestion of BCA) and is consistent with this enzymatic peptide being a 12-mer and containing one PEG moiety.

The acylated BCA sequence between residues 106 and 149 contains four acylated lysine residues. The ESI-Q-o-TOF MS/MS spectra for the enzymatic peptides from region of perPEGylated BCA are shown in Figure 5.27 - Figure 5.29 below. The spectrum for perPEGylated BCA GluC enzymatic peptide 106-116 (Figure 5.27) contains b_4 - b_{10} and y_1 - y_{10} ions, showing the presence of two PEGylated lysine residues: lys-111 and lys-

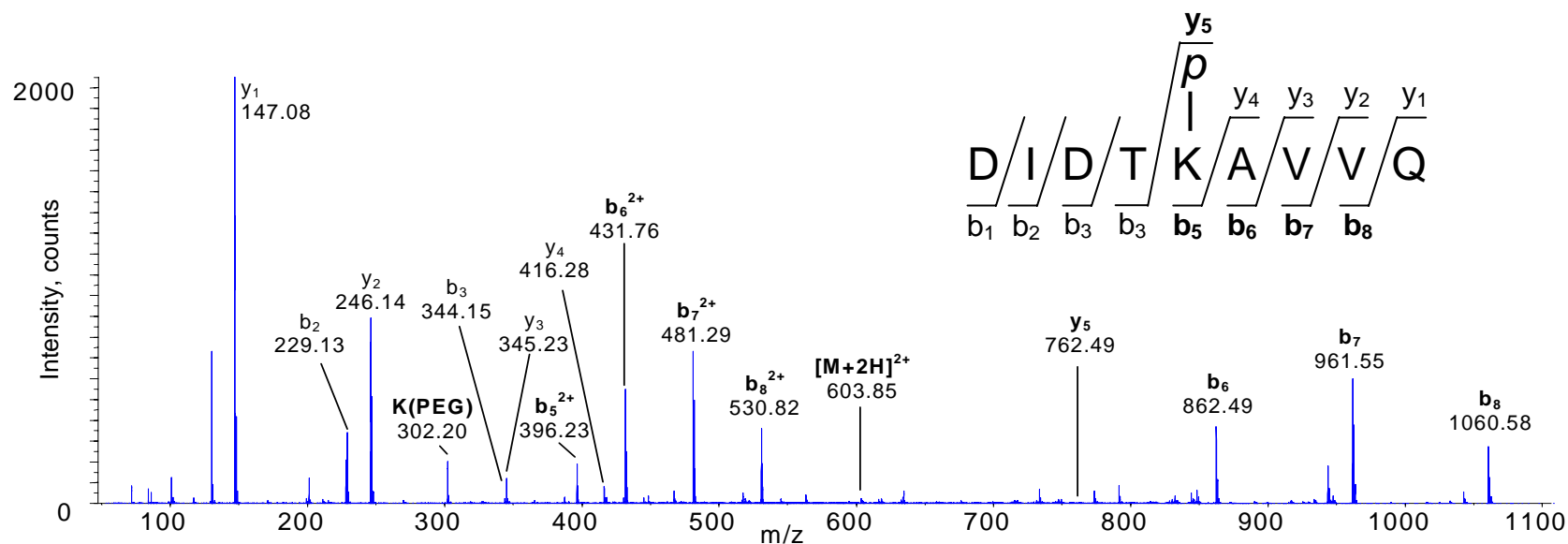


Figure 5.25 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 31-39. p indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

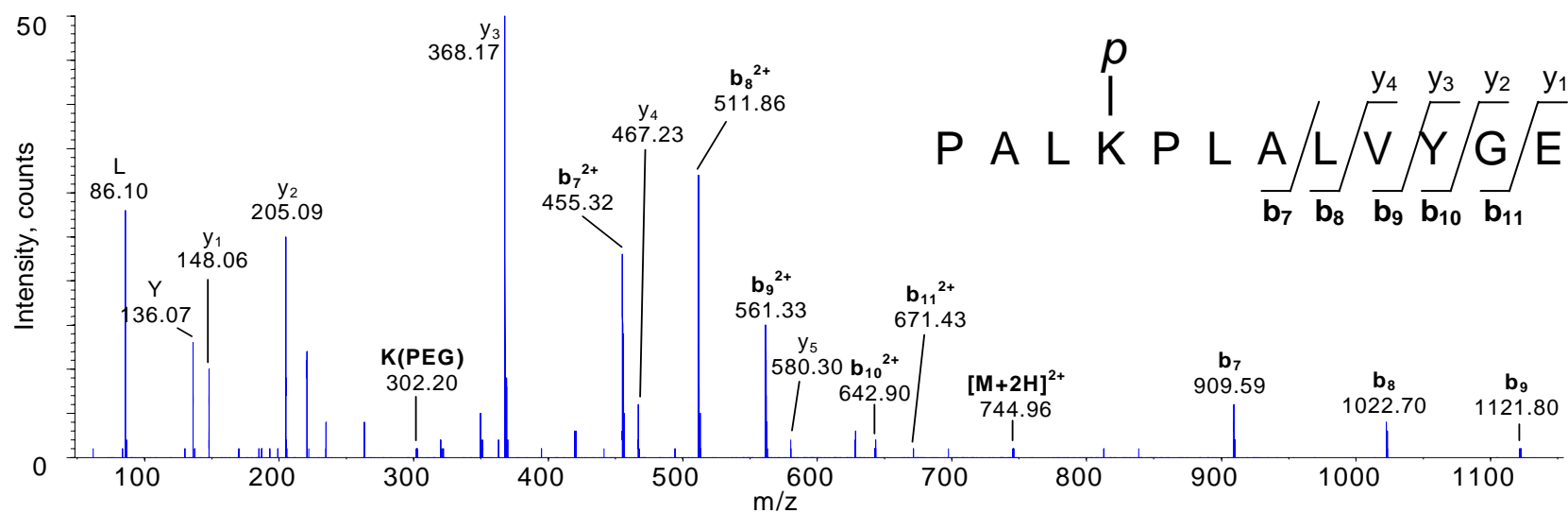


Figure 5.26 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 41-52. p indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

112. The b_4 and y_4 ions are indistinguishable at the resolution of the ESI-Q-o-TOF MS/MS spectrum shown here, with an observed m/z value of 453.20 corresponding to a theoretical value of m/z 453.21 for b_4 and m/z 453.20 for y_4 . There was no evidence for an unmodified lysine at either lys-111 or lys-112. This enzymatic peptide eluted at 26.4 minutes, which is relatively early in the elution profile for a peptide containing two PEG moieties and eleven amino acids, and reflects presence of several hydrophilic amino acids in this peptide. According to the hydropathy reference scales created by Wolfenden *et.al.* and Kyte and Doolittle as reported and discussed on the Prowl website, the sequence HTVDRKKYAAE, contains six hydrophilic residues (histidine, aspartic acid, arginine, two lysines and glutamic acid); three hydrophobic (valine and two alanines) and two residues of relatively neutral hydropathicity (threonine and tyrosine).³⁴⁰⁻³⁴²

Figure 5.28 shows the ESI-Q-o-TOF MS/MS spectrum of perPEGylated BCA GluC enzymatic peptide 117-128. This spectrum shows the product ions b_2 , b_6 , b_{9-11} , y_{1-3} and y_{10} along with the PEGylated lysine carbenium ion at m/z 302.21. Together these data confirm the identity of the peptide and the presence of a PEG moiety on the lysine residue, being lys-125. This GluC enzymatic peptide eluted at 34.1 minutes which is consistent for a 12-mer containing one PEG moiety.

ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA AspN enzymatic peptide 137-149 is shown in Figure 5.29. Product ions b_{3-11} , y_{2-5} and y_{7-8} are present in the spectrum, along with a PEGylated lysine carbenium ion. The sequence ions confirm the position of the PEGylated lysine at lys-147. The elution time of 45.1 minutes for this AspN enzymatic peptide of BCA is towards the end of the elution profile, and is consistent with 13-mer containing one PEG moiety (see Figure 5.14).

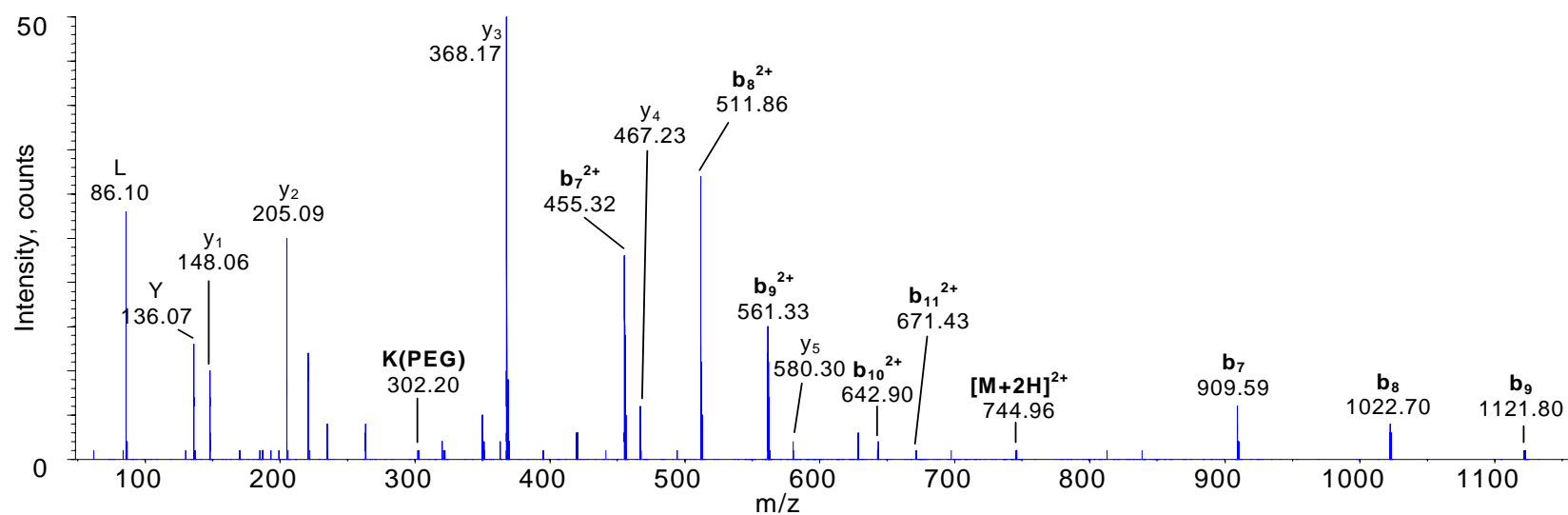
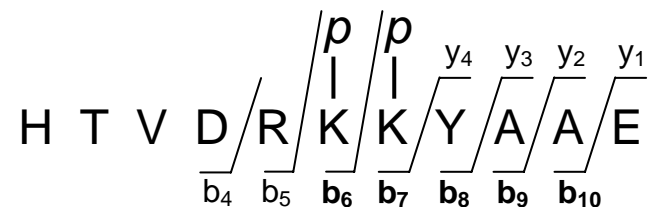


Figure 5.27 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 106-116. *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.



Figure 5.28 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 117-128. *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

The acylated BCA sequence between residues 204 and 259 contains four acylated lysine residues. (The region 173 – 204 is not shown as there are no sites of acylation, or other features of interest in this portion of the BCA sequence.) The ESI-Q-o-TOF MS/MS spectra for the enzymatic peptides of this region of perPEGylated BCA are shown in Figure 5.30 - Figure 5.32 below. The ESI-Q-o-TOF MS/MS spectrum for perPEGylated BCA GluC enzymatic peptide 204-212 is shown in Figure 5.30. The spectrum contains an almost complete series of sequence product ions, with b_{2-8} and y_{1-6} . The PEGylated lysine carbenium ion is also present at m/z 302.20. These data confirm the position of the PEGylated lysine at lys-211. The spectrum contains a number of relatively abundant $b-H_2O$ ions. The presence of these ions can be attributed to the presence of the serine and threonine residues at the N-terminal end of the GluC enzymatic peptide.¹⁶¹ The elution time for this peptide was 40.2 minutes which is towards the middle of the elution profile for the GluC enzymatic peptides of BCA. This elution time is later than would otherwise be expected for a 9-mer containing one PEG moiety. It is consistent, however, with the amino acid sequence (SVTWIVLKE) which contains a number of longer-chain and/or hydrophobic residues, including valine, tryptophan, isoleucine and leucine.

Figure 5.31 shows the ESI-Q-o-TOF MS/MS spectrum of perPEGylated BCA GluC enzymatic peptide 204-232. The MS/MS spectrum produced for this peptide (a 29-mer) does not give adequate data to confirm the two acylation sites. The sequence ions present in the spectrum include y_{1-3} , b_2 and b_4 . And a PEGylated lysine carbenium ion is also present in the spectrum. These data, along with the spectrum for perPEGylated BCA GluC enzymatic peptide 204-212, indicate that lys-223 is likely to also be PEGylated, in addition to lys-211. The elution time of 46.1 minutes for BCA GluC

enzymatic peptide 204-232 is towards the end of the elution profile for the GluC digestion of BCA and is consistent with a 29-mer containing two PEG moieties.

The ESI-Q-o-TOF MS/MS spectrum for perPEGylated BCA GluC enzymatic peptide 237-259 is shown in Figure 5.32. This peptide shows mostly N-terminal product sequence ions: b_{2-11} and y_{15-22} with an additional y_8 . The peptide (a 23-mer) contains two potentially PEGylated lysine residues (lys-205 and lys-259). Concomitant with the mostly N-terminal sequence ions is the absence of an m/z value corresponding to a PEGylated lysine carbenium ion, which has been present in the majority of other PEGylated peptides studied in this chapter. The presence of a y_8 ion allows for the isolation of the PEGyl moieties to two separate regions of the C-terminal end of the peptide. This peptide is a 23-mer, containing two PEG moieties, however it eluted prior to the BCA GluC enzymatic peptide 41-52, a 12-mer containing one PEG moiety. The relatively early elution time of GluC peptide 237-259 can be explained by the relatively hydrophilic nature of the amino acids contained in this section of the BCA sequence. The GluC enzymatic peptide 237-259 (LLMLANWRPA QPLKNRQVRG FPK) contains eight hydrophilic residues, comprising approximately 35 % of the sequence (being: two asparagine, three arginine, one glutamine and two lysine residues) The sequence contains three arginine residues, which is a particularly hydrophilic amino acid. This is in comparison to the GluC enzymatic peptide 41-52 (PALKPLALVYGE) that contains only two hydrophilic residues (being lysine and glutamic acid).³⁴⁰⁻³⁴²

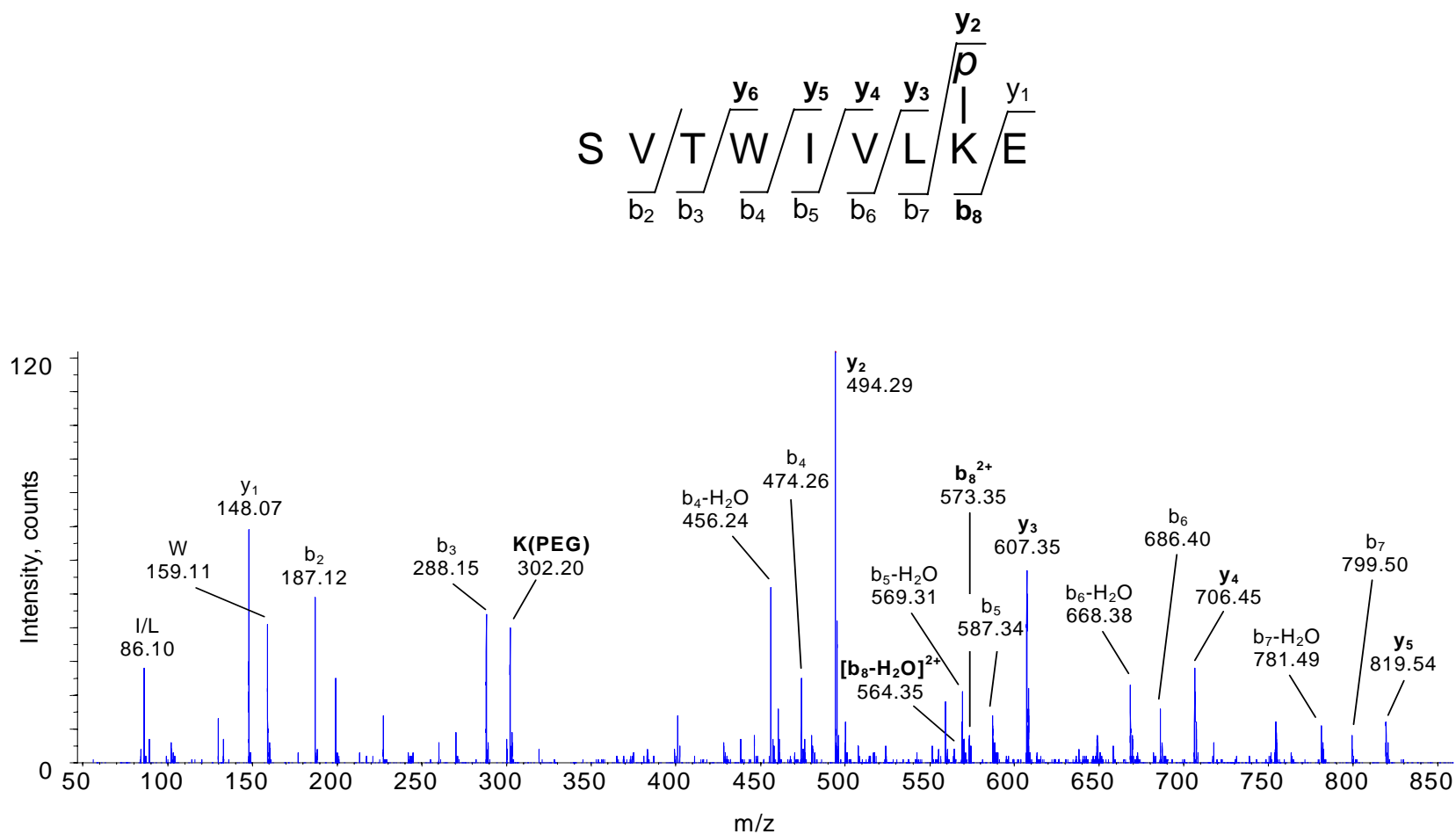


Figure 5.30 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 204-212. *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

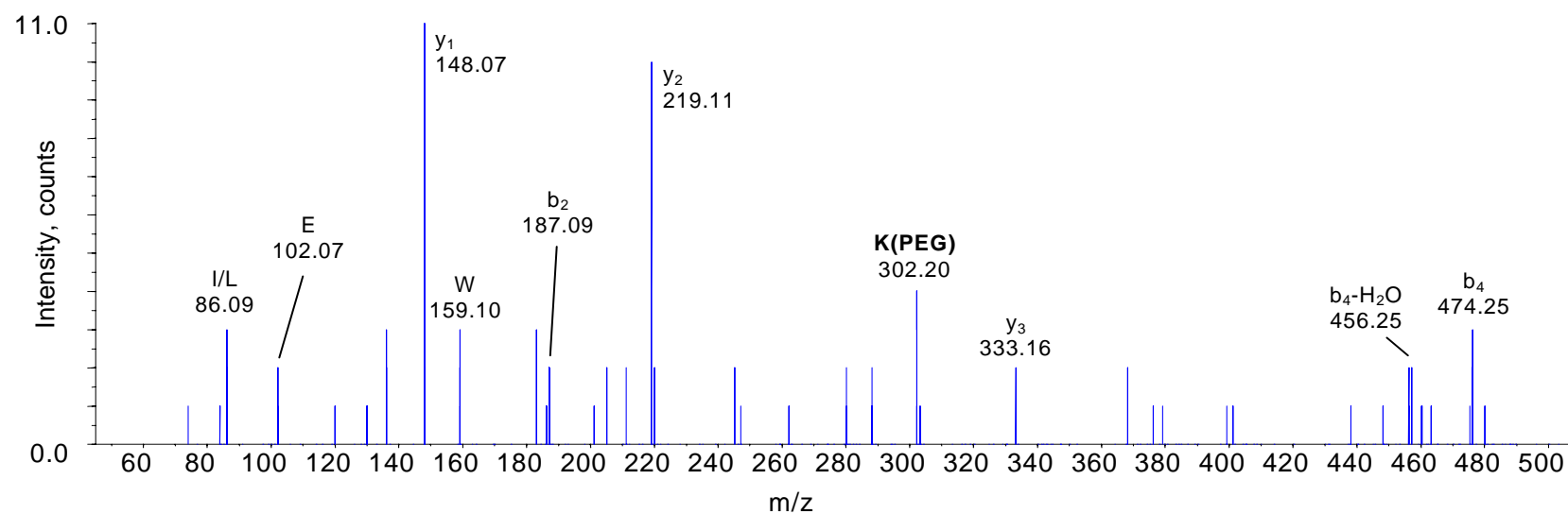


Figure 5.31 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 204-232. *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

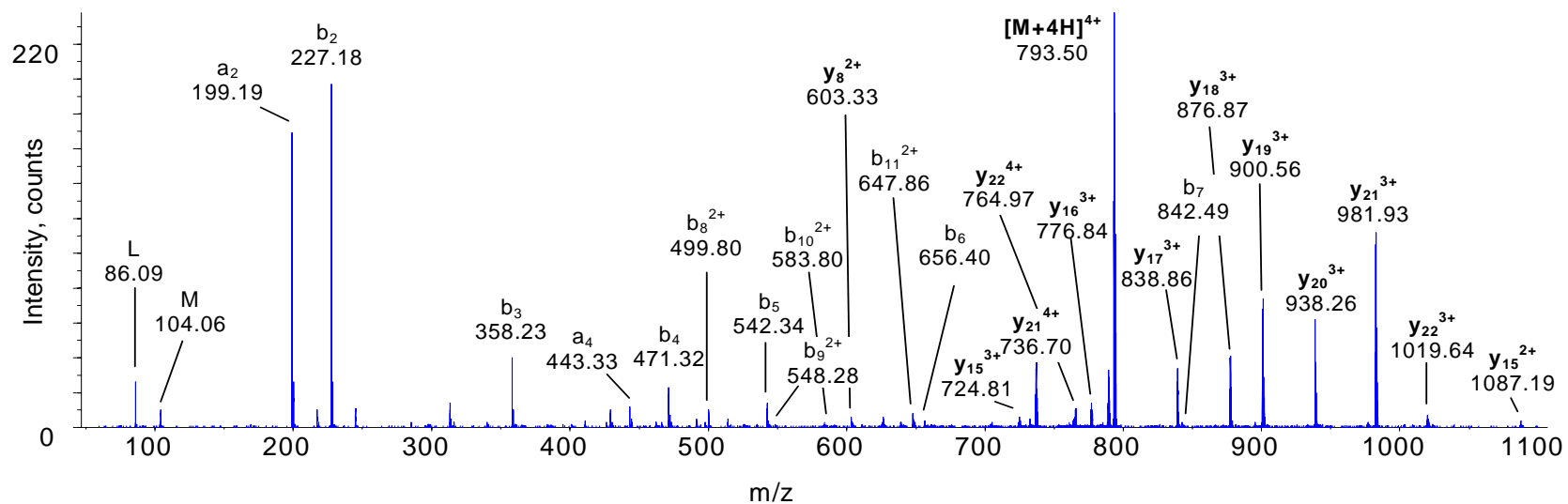


Figure 5.32 HPLC-ESI-Q-o-TOF MS/MS analysis of perPEGylated BCA GluC enzymatic peptide 237-259. *p* indicates a PEG group; bold type indicates the presence of a product ion containing the acyl moiety.

5.2.3 General Discussion

A dominant feature of the analysis of peracylated ubiquitin and BCA was the presence of an acylated lysine carbenium ion in the tandem mass spectra. The acylated lysine carbenium ion observed is equivalent to the acylated immonium ion with the loss of NH_3 as a neutral. This carbenium ion was consistently observed irrespective of the acylating agent. This marker ion for acylated lysine is in contrast to the acylated immonium ions observed for acylated serine, tyrosine and cysteine, and the additional acylated acylium ion observed for cysteine (as discussed in Chapter 3). The presence of the acylated lysine carbenium ion is in agreement, however, with the observations of Kim *et. al.*, who, when studying acetylated proteins, found the ion at m/z 126 to be more abundant (*ca.* nine times) and a more reliable marker for the presence of acetylated lysine, when compared to the acetylated lysine immonium ion.³¹⁹ This marker ion is particularly useful for identifying smaller acylations on lysine residues. For larger acylations, such as PEG, however, there is the possibility of this marker ion having a similar mass to sequence or internal product ions. This was observed for perPEGylated BCA GluC enzymatic peptide 1-13, where the sequence ion y_3 (m/z 302.14) has a similar m/z to the K(PEG) ion at m/z 302.20. As mass analyser technology improves, the resolution required to easily differentiate these small differences in m/z will become available.

A number of sites in both ubiquitin and BCA were identified as being sub-stoichiometrically acylated when acylated with glutarate or PEG. These sites were found in parts of the proteins' sequences where lysine residues were in close proximity to each other, suggesting that steric interactions played a role in the sub-stoichiometric acylation at these sites.

5.3 Conclusions

The use of chromatographic separation (HPLC), in conjunction with mass spectrometry (ESI-Q-o-TOF MS/MS), has allowed the characterisation of enzymatically digested monodisperse polymers based on the proteins ubiquitin and bovine carbonic anhydrase II (BCA), peracylated at the amine containing side-chain of their lysine residues, and, if available, N-terminal amino group. Using collisionally induced dissociated (CID), a reliable marker ion, being an acylated lysine carbenium ion, was identified. This marker ion was observed for all acylations studied: lys(acetate), m/z 126.1; lys(benzoate), m/z 188.1; lys(trifluoropropionate), m/z 194.1; lys(glutarate), m/z 198.1; lys(iodoacetate), m/z 252.0 and lys(PEG), m/z 302.2. In the absence of adequate product ion sequence coverage, the presence of an acylated lysine carbenium ion can be used to confirm the presence of an acylated residue within a peptide.

The identity of the acyl moiety was found to have a marked affect on the chromatographic elution time of the various enzymatic peptides. It was found that the acylated peptides eluted according to the order: acetyl, glutaryl, PEGyl, trifluoroacetyl, iodoacetyl, followed by benzoyl acylations. Generally the greater the number of acyl moieties contained within a peptide, the later the elution time, given a similar amino acid length and composition.

The use of mass spectrometry allowed the detection of a number of characteristics of the molecules being studied. An additional acylation to the ubiquitin molecule, thought to be on an available hydroxyl group of a serine or tyrosine residue, was found. The detection of this additional, unwanted acylation, lead to a modification to the synthesis of the monodisperse polymers, with the addition of a hydrolysis step eliminating any hydroxyl-type acylations. The use of mass spectrometry also enabled the detection of

deamidated asparagine residues, truncated PEG modifications and also sub-stoichiometric acylation.

In conclusion, the successful synthesis of a monodisperse polymer was found to depend on a number of specific properties of the acyl moiety and protein backbone. Both ubiquitin and BCA, when acylated with glutaryl or methoxypolyethylene glycol propionyl at lysine residues, were found to contain sub-stoichiometric acylation. This was attributed to both the size of the glutaryl and PEGyl moieties relative to the other acyl moieties studied, and to the proximity of lysine residues within the protein sequences. The protein BCA was not found to be a suitable backbone for the production of monodisperse polymers due to its propensity to undergo deamidation at three of the asparagine residues contained within its sequence (asn-10, asn-23 and asn-61), thus limiting the purity of the final product.

Monodisperse polymers, which have many potential uses, will require stringent characterisation prior to use as standards or pharmaceuticals. Mass spectrometry provides a useful means for the detailed characterisation of these types of acylated molecules.

6 General Conclusions

The aim of this thesis was to gain an understanding of the tandem mass spectrometry of acylated peptides and then apply this knowledge to the analysis of a range of novel acylated peptides and proteins. A number of characteristic fragments were observed in the MS/MS of acylated peptides, which were utilised in the sequencing of a pair of acylated 12-mer peptides. The tandem mass spectrometry fragmentation patterns of peptides derived from monodisperse polymers (proteins acylated with novel acylating agents) were found to be generally similar to those of the smaller acylated peptides studied.

A range of acylated standard peptides were synthesised and subjected to MS/MS analysis. Standard peptides suitable for the study of *O*-, *S*- and N-terminal *N*-acylation were used, with acyl chain lengths of 8- to 16-carbons. Fragmentation patterns observed were generally found to be independent of the presence and length of the acyl chain. The intensities of the various fragment ions observed, were, however, found to be dependent on the collision energy selected. The general trend observed was that the longer the acyl chain length, the higher the collision energy required to achieve the same degree of fragmentation. The most informative ions observed when fragmenting acylated peptides using collisionally induced dissociation, were the presence of acylated immonium and related ions, and the presence of a b_1 ion in the case of the N-terminally acylated peptide. Additional useful ions included ions resulting from the neutral loss of the acyl moiety from the precursor or fragment ions, as either an acyl group or carboxylic acid, and an acyl carbenium ion. The presence of an acylated immonium ion and the b_1 ion assist in identifying the acyl chain length and the identity of the modified

amino acid residue. The observation of neutral losses and the presence of an acyl carbenium ion assist in confirming the identity of the acyl moiety.

The knowledge gained from the tandem mass spectrometry study of acylated peptide standards assisted greatly in the characterisation of a pair of unidentified 12-mer peptides. The peptides were synthesised as an inadvertent side-reaction to the preparation of an HIV-antigenic *N*-acetylated 6-mer peptide. These 12-mer peptides were found to be acylated at their N-termini with stearoyl (18:0) and oleoyl (18:1) respectively, with identical amino acid sequences. The sequence was determined to be (ste/ole)GGKWSK-O-kynSKWSK. The presence of a pair of b_1 ions being, m/z 266 and m/z 264 higher than that calculated for an unmodified glycine b_1 ion, gave the first indication of the presence of an *N*-acylated peptide. The second indication of the presence of an acylation was the loss of 266 u and 264 u respectively from the peptide b-series sequence ions. The tandem mass spectral analysis of the intact 12-mer peptides, and of shorter peptides resulting from tryptic digestion, allowed the complete sequence of the peptides to be determined. An unusual ether-linked kynurenine residue was identified as amino acid-7 in the peptide sequence. The sequencing of this pair of acylated peptides led to the subsequent synthesis of a stearoyl modified 12-mer peptide analogue, containing a tryptophan at amino acid-7, and further study of the immunogenic properties of these types of acylated peptides. The 12-mer peptide has also been patented as a possible compound for use in immune modulation, or as a reagent in evaluating the activity of T cells in a patient.³¹⁰

A range of acylated proteins (termed monodisperse polymers) were analysed by mass spectrometry. Ubiquitin and BCA II were modified at lysine side-chain amine groups with a range of novel modifying agents. The novel acyl moieties included: acetyl; benzoyl; trifluoropropionyl; glutaryl; iodoactyl and polyethylene glycol (PEG). The N-

terminal methionine of ubiquitin was also *N*-acylated, with the N-terminal serine of BCA II already being acetylated. The modified proteins were digested with the enzymes Glu-C and Asp-N prior to analysis by LC/MS. The fragmentation patterns observed for novel lysine acylations were similar to that observed for simple carbon-chain acylations. Analogous fragmentation patterns were observed for identical peptide sequences regardless of the identity, and indeed presence (with the exception of b_1 ions), of the acyl moiety. As previously observed, b_1 ions were present in the tandem mass spectrum where the N-terminal residue of a peptide was acylated. Acylated lysine carbenium ions were observed in the tandem mass spectra of peptides containing an acylated lysine residue, in comparison to the immonium ions observed for acylated serine, tyrosine and cysteine acylations.

The modification of the majority of individual lysine residues was confirmed by the sequence analysis of the MS/MS spectra of individual enzymatic peptides. In a few cases, the manner in which the peptides fragmented in MS/MS prohibited the specific localisation of the modification. In these cases, the presence of the acyl moiety could be confirmed by the presence of an appropriate marker ion and an appropriate delta mass for the intact peptide. A benefit of mass spectrometry was highlighted when an additional acylation, not observed in a capillary electrophoresis analysis by our collaborators, was observed in the MS of acylated ubiquitin. This additional acylation was removed by treatment with base, suggesting an addition at tyrosine or serine. A number of sites susceptible to sub-stoichiometric acylation were identified, underlining the importance of the careful selection of protein backbone sequences in the synthesis of monodisperse polymers.

The successful demonstration of the utility of tandem mass spectrometry for the analysis of acylated peptides opens the way for future studies of biological samples. Knowledge of the characteristic ions formed in the fragmentation of acylated peptides combined with HPLC separation, allows for the identification of hitherto unknown acylated peptides and proteins, from complex biological mixtures.

APPENDICES

Appendix A. Assignment of MALDI-Q-o-TOF MS/MS results for (pam)Nef₁₋₆.

Experimental [m+H] ⁺ /u	Relative area (%)	Assignment [†]	Calculated [m+H] ⁺ /u	Error (ppm)
900.61	27.6	[M+H] ⁺	900.59	18
882.59	2.1	[M-H ₂ O] ⁺	882.58	11
771.45	1.6	c ₅	771.51	-83
754.52	2.5	b ₅	754.49	37
736.49	0.8	b ₅ -H ₂ O	736.48	19
726.49	0.8	a ₅	726.49	6
667.47	10.6	b ₄	667.45	33
649.46	6.8	b ₄ -H ₂ O	649.44	34
639.48	4.6	a ₄	639.46	24
605.37	2.0	y ₅	605.34	48
548.34	1.3	y ₄	548.32	29
530.32	1.1	y ₄ -H ₂ O	530.31	12
481.40	3.3	b ₃	481.38	33
463.38	4.8	b ₃ -H ₂ O	463.37	20
459.24	1.5	GKWS	459.24	11
441.25	0.8	GKWS-H ₂ O	441.23	41
420.23	15.0	y ₃	420.22	19
403.21	2.3	y ₃ -NH ₃	403.20	22
402.22	2.1	y ₃ -H ₂ O KWS	402.21	21
385.21	1.9	KWS-NH ₃	385.19	55
372.22	2.3	GKW	372.20	41
367.20	1.2	KWS-H ₂ O-NH ₃	367.18	65
353.27	2.2	b ₂	353.28	-36
344.21	1.7	GKW-28	344.21	5
296.25	6.1	b ₁	296.26	-17
274.13	3.4	WS	274.12	28
246.13	2.0	WS-28	246.12	52
243.14	2.4	b ₃ -pam	243.15	-24
239.24	1.1	palmitoyl	239.24	23
234.15	9.4	y ₂	234.15	-19
225.13	7.3	b ₃ -pam-H ₂ O	225.14	-14
216.13	3.2	y ₂ -H ₂ O	216.13	-3
199.11	1.2	y ₂ -H ₂ O-NH ₃	199.11	-10
198.13	1.0	a ₃ -NH ₃ -pam	198.12	44
186.12	67.9	GK	186.12	27
169.10	5.9	GK-NH ₃	169.10	-13
159.09	13.4	W	159.09	19
147.11	8.0	y ₁	147.11	11
132.08	6.3	c ₂ -pam	132.08	-16
129.10	100.0	K	129.10	2
101.11	0.8	K	101.11	-25
84.08	18.4	K	84.08	-19

[†] pam: Palmitoyl.

Appendix B. Assignment of ESI-Q-o-TOF MS/MS results for (pam)Nef₁₋₆.

Experimental [m+H] ^{††} /u	Relative area (%)	Assignment ^{††}	Calculated mass [m+H] [†] /u	Error (ppm)
900.64	100.00	[M+H] [†]	900.59	49
882.64	5.20	[M-H ₂ O+H] [†]	882.58	71
772.53	0.02	b ₅ +H ₂ O	772.50	34
754.53	0.04	b ₅	754.49	49
662.37	22.42	M-pam+H	662.36	15
605.35	13.23	y ₅	605.34	18
589.30	1.06	z ₅	589.32	-26
587.33	2.24	y ₅ -H ₂ O	587.33	5
553.35	0.10	y ₅ -H ₂ O-2NH ₃	553.28	135
548.38	1.06	y ₄	548.32	109
531.32	2.52	y ₄ -NH ₃	531.29	54
530.35	0.43	y ₄ -H ₂ O	530.31	71
481.39	1.38	b ₃	481.38	21
459.26	1.77	GKWS	459.24	41
429.25	6.97	b ₄ - pam	429.23	47
420.24	15.45	y ₃	420.22	51
401.26	6.22	a ₄ - pam	401.23	69
385.22	1.38	y ₃ -H ₂ O-NH ₃ KWS-NH ₃	385.19	85
372.22	3.81	GKW	372.20	37
367.21	0.64	KWS-H ₂ O-NH ₃	367.18	89
344.22	0.92	GKW-28	344.21	39
315.18	0.95	KW	315.18	2
296.27	10.25	b ₁	296.26	38
274.14	3.75	WS	274.12	56
246.15	1.96	WS-28	246.12	103
243.15	5.81	b ₃ - pam	243.15	21
239.24	14.28	palmitoyl	239.24	7
234.15	19.00	y ₂	234.15	13
216.14	7.86	y ₂ -H ₂ O	216.13	60
199.11	3.37	y ₂ -H ₂ O-NH ₃	199.11	2
186.13	22.86	GK	186.12	53
159.09	15.48	W	159.09	18
147.12	19.22	y ₁	147.11	99
132.08	2.70	c ₂ - pam	132.08	35
129.10	99.46	K	129.10	26
101.11	4.05	K	101.11	-42
84.09	59.70	K	84.08	92

[†] m values are determined automatically by Analyst QSTM with the result combining multiple charge states. The combined m values were imported into EXCEL[™] and converted to [m+H][†] values (or in the case of a, b or c ions, m⁺ values).

^{††} pam: palmitoyl.

Appendix C Assignment of MALDI-Q-o-TOF MS/MS results for Nex 1 and Nex 2.

Experimental [m+H] ⁺ /u	Relative area (%)	Assignment [†]	Calculated [m+H] ⁺ /u	Error (ppm)
1736.04	100.0	[Nex2+H] ⁺	1736.02	14
1734.01	45.0	[Nex1+H] ⁺	1734.00	5
1718.00	21.0	[Nex2-H ₂ O+H] ⁺	1718.00	-2
1716.00	10.1	[Nex1-H ₂ O+H] ⁺	1715.99	9
1700.02	1.5	[Nex2-2H ₂ O+H] ⁺	1699.99	15
1698.04	0.7	[Nex1-2H ₂ O+H] ⁺	1697.98	36
1607.95	23.9	b ₁₁ +H ₂ O	1607.92	20
1605.90	12.7	b ₁₁ +H ₂ O*	1605.90	1
1589.94	11.1	b ₁₁	1589.91	19
1587.89	5.5	b ₁₁ *	1587.89	-2
1571.95	3.5	b ₁₁ -H ₂ O	1571.90	30
1569.90	1.3	b ₁₁ -H ₂ O*	1569.88	9
1520.89	1.2	b ₁₀ +H ₂ O	1520.89	4
1518.93	0.6	b ₁₀ +H ₂ O*	1518.87	38
1502.93	7.0	b ₁₀	1502.88	33
1500.87	4.3	b ₁₀ *	1500.86	6
1484.91	6.0	b ₁₀ -H ₂ O	1484.87	31
1482.85	2.9	b ₁₀ -H ₂ O*	1482.85	-3
1412.75	5.3	y ₁₁	1412.73	9
1316.83	8.5	b ₉	1316.80	22
1314.79	5.7	b ₉ *	1314.78	6
1298.81	1.4	b ₉ -H ₂ O	1298.79	20
1227.64	14.3	y ₉	1227.62	22
1209.66	4.1	y ₉ -H ₂ O KWSKXSKWS	1209.60	51
1188.72	2.1	b ₈	1188.70	15
1186.76	1.8	b ₈ *	1186.69	65
1099.59	2.8	WSKXSKWS+H ₂ O	1099.52	66
1081.58	3.6	WSKXSKWS	1081.51	59
994.53	2.4	WSKXSKW	994.48	56
977.48	1.2	WSKXSKW-NH ₃	977.45	33
975.55	1.1	GKWSKXSK-H ₂ O	975.51	43
937.52	2.6	y ₇ -NH ₃	937.48	41
928.64	5.1	c ₆	928.62	22
847.43	1.8	GKWSKXS-H ₂ O	847.41	21
808.41	14.4	KWSKXS SKXSKW WSKXSK KXSKWS	808.40	12
761.39	1.5	GKWSKX-NH ₃	761.36	34

680.31	3.2	WSKXS	680.30	10
	0.0	XSKWS	680.30	
662.32	0.0	c ₆ -acyl	662.36	-63
662.32	4.601	XSKWS-H ₂ O WSKXS-H ₂ O	662.294	41
644.33	1.3	XSKWS-2H ₂ O	644.28	74
644.33	0.0	WSKXS-2H ₂ O		
635.36	1.5	y ₅	635.35	21
617.30	1.9	y ₅ -H ₂ O	617.34	-61
605.36	2.5	GKWSK+H ₂ O	605.34	34
587.30	1.9	SKXSK-H ₂ O-NH ₃	587.28	29
587.30	0.0	GKWSK	587.33	-52
586.30	0.6	SKXSK-2H ₂ O	586.30	2
575.27	2.172	WSKX-H ₂ O XSKW-H ₂ O	575.262	19
557.27	2.746	WSKX-2H ₂ O XSKW-2H ₂ O	557.251	26
547.33	12.771	WSKX-H ₂ O-28 XSKW-H ₂ O-28	547.262	116
530.26	8.5	KWSK	530.31	-95
518.26	3.7	KXSK-NH ₃	518.26	4
489.41	1.0	SKWS	489.25	335
436.23	2.4	SKWS-2H ₂ O-NH ₃	436.20	64
420.24	4.303	y ₃ WSK+H ₂ O SKW+H ₂ O KWS+H ₂ O	420.225	43
389.20	2.8	SKX-H ₂ O KXS-H ₂ O XSK-H ₂ O	389.18	56
274.12	0.0	WS	274.12	3
261.10	0.0	XS-H ₂ O	261.09	48
234.15	1.6	y ₂	234.15	21
186.14	8.0	GK	186.12	89
129.11	10.6	K	129.10	44

[†] X denotes ester linked kynurine; Nex 1 refers to the Nex peptide with an oleoyl (18:1(9)) residue at the N-terminal glycine, Nex 2 refers to the Nex peptide with a stearoyl (18:0) residue at the N-terminal glycine; * indicates the oleoyl (18:1(9)) form of the product ion; “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from Nex.

Appendix D Assignment of ESI-Q-o-TOF MS/MS results for Nex 1 and Nex 2.

Experimental [m+H] ⁺ /u	Relative area (%)	Assignment ^{††}	Calculated [m+H] ⁺ /u	Error (ppm)
1735.98	100.0	[Nex2+H] ⁺	1736.02	-21
1733.97	32.3	[Nex1+H] ⁺	1734.00	-16
1717.98	10.5	[Nex2-H ₂ O+H] ⁺	1718.00	-15
1715.99	3.0	[Nex1-H ₂ O+H] ⁺	1715.99	0
1608.03	3.9	b ₁₁ +H ₂ O	1607.92	70
1605.94	0.7	b ₁₁ +H ₂ O*	1605.90	19
1590.03	1.2	b ₁₁	1589.91	77
1588.01	0.5	b ₁₁ *	1587.89	71
1502.96	0.7	b ₁₀	1502.88	56
1501.02	0.3	b ₁₀ *	1500.86	108
1469.84	9.3	[Nex-acyl+H] ⁺	1469.75	60
1451.81	1.5	[Nex-H ₂ O-acyl+H] ⁺	1451.74	43
1412.80	5.6	y ₁₁	1412.73	49
1394.80	1.0	y ₁₁ -H ₂ O	1394.72	53
1355.85	0.5	y ₁₀	1355.71	100
1323.68	1.3	b ₁₁ -acyl	1323.65	21
1316.80	0.8	b ₉	1316.80	4
1314.81	0.4	b ₉ *	1314.78	18
1266.62	0.6	GKWSKXSKWS	1266.63	-10
1236.59	1.3	b ₁₀ -acyl	1236.62	-18
1227.61	2.1	y ₉	1227.62	-4
1209.64	1.7	y ₉ -H ₂ O or KWSKXSKWS	1209.61	30
1179.61	0.5	GKWSKXSKW	1179.60	10
966.64	1.6	WSKXSKW-28	966.48	160
960.60	1.3	WSKXSKW-2NH ₃	960.43	183
948.54	0.9	WSKXSKW-28-H ₂ O	948.47	73
937.51	1.4	y ₇ -NH ₃	937.48	32
928.67	3.0	c ₆	928.62	48
926.65	1.2	c ₆ *	926.61	45
923.50	0.7	GKWSKXSK-2H ₂ O-2NH ₃	923.44	64
910.69	0.3	b ₆	910.61	84
900.63	1.9	c ₆ -C ₂ H ₄	900.59	44
865.47	0.2	GKWSKXS	865.42	53
808.46	31.1	KWSKXS SKXSKW WSKXSK KXSKWS	808.3994	69

790.45	2.2	KWSKXS-H2O SKXSKW-H2O WSKXSK-H2O KXSKWS-H2O	790.39	72
780.49	0.6	WSKXSK-28 SKXSKW-28 KWSKXS-28 KXSKWS-28	780.40	108
773.40	0.7	KWSKXS-H2O-NH3 SKXSKW-H2O-NH3 KWSKXS-H2O-NH3 KXSKWS-H2O-NH3	773.36	49
548.35	1.4	y ₄	548.32	46
420.24	5.6	y ₃	420.22	39
403.22	0.8	y ₃ -NH ₃	403.20	55
324.30	1.6	b1	324.29	34
322.31	1.5	b1*	322.27	119
296.26	2.3	b1-C ₂ H ₄	296.26	7
267.28	1.9	stearoyl (18:0)	267.27	47
265.27	3.8	oleoyl (18:1)	265.25	56
234.15	17.8	y ₂	234.15	-7
216.14	6.7	y ₂ -H ₂ O or SK	216.13	63
	0.0		216.13	
199.11	2.0	SK-NH ₃	199.11	-8
186.13	12.0	GK	186.12	28
169.12	1.5	GK-NH ₃	169.10	142
159.10	6.2	W	159.09	83
147.12	10.5	y ₁	147.11	90
129.11	50.0	K	129.10	40
101.11	2.3	K	101.11	-37
84.08	24.5	K	84.08	-29

[†] m values are determined automatically by Analyst QSTM with the result combining multiple charge states. The combined m values were imported into EXCELTM and converted to [m+H]⁺ values (or in the case of a, b or c ions, m⁺ values).

^{††} X denotes ester linked kynurine; Nex 1 refers to the Nex peptide with an oleoyl (18:1(9)) residue at the N-terminal glycine, Nex 2 refers to the Nex peptide with a stearoyl (18:0) residue at the N-terminal glycine; * indicates the oleoyl (18:1(9)) form of the product ion; “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from Nex.

Appendix E Assignment of ESI-FTICR MS/MS results for Nex 1 and Nex 2.

<i>m/z</i>	<i>z</i>	experimental [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	relative intensity (%)	assignment [†]	theoretical [<i>m</i> + <i>H</i>] ⁺ / <i>u</i>	error (ppm)
579.342	3	1736.011	33.4	[Nex2+H] ⁺	1736.015	-2.2
868.507	2	1736.007	6.5	[Nex2+H] ⁺	1736.015	-4.8
578.671	3	1733.998	14.2	[Nex1+H] ⁺	1733.999	-0.6
867.501	2	1733.993	5.8	[Nex1+H] ⁺	1733.999	-3.5
573.339	3	1718.003	11.0	[Nex2-H ₂ O+H] ⁺	1718.005	-1.1
859.505	2	1718.001	2.7	[Nex2-H ₂ O+H] ⁺	1718.005	-1.9
572.669	3	1715.990	4.1	[Nex1-H ₂ O+H] ⁺	1715.989	0.9
858.496	2	1715.985	1.1	[Nex1-H ₂ O+H] ⁺	1715.989	-2.5
854.494	2	1707.980	1.0	[Nex2-C ₂ H ₄ +H] ⁺	1707.984	-2.3
804.462	2	1607.916	9.5	b ₁₁ +H ₂ O	1607.920	-2.8
803.455	2	1605.903	7.1	b ₁₁ +H ₂ O*	1605.904	-0.9
795.457	2	1589.907	7.3	b ₁₁	1589.910	-1.7
794.451	2	1587.894	2.7	b ₁₁ *	1587.894	0.1
786.453	2	1571.898	1.0	b ₁₁ -H ₂ O	1571.899	-0.6
751.941	2	1502.873	4.6	b ₁₀	1502.878	-2.9
750.934	2	1500.861	1.9	b ₁₀ *	1500.862	-0.8
742.937	2	1484.867	2.5	b ₁₀ -H ₂ O	1484.867	0.1
741.930	2	1482.852	1.0	b ₁₀ -H ₂ O*	1482.851	0.6
735.380	2	1469.753	22.1	[Nex-acyl+H] ⁺	1469.754	-0.6
490.591	3	1469.757	1.4	[Nex-acyl+H] ⁺	1469.754	2.2
726.375	2	1451.742	8.3	[Nex-H ₂ O-acyl+H] ⁺	1451.744	-1.2
717.370	2	1433.732	1.1	[Nex-2H ₂ O-acyl+H] ⁺	1433.733	-0.9
706.870	2	1412.732	12.7	y ₁₁	1412.733	-0.6
697.865	2	1394.723	5.7	y ₁₁ -H ₂ O	1394.722	0.6
688.860	2	1376.712	0.9	y ₁₁ -2H ₂ O	1376.712	0.0
678.360	2	1355.711	0.9	y ₁₀	1355.711	0.1
671.333	2	1341.659	1.7	b ₁₁ -acyl+H ₂ O	1341.659	-0.1
669.354	2	1337.701	0.7	y ₁₀ -H ₂ O	1337.701	0.3
662.329	2	1323.649	3.2	b ₁₁ -acyl	1323.649	0.6
1316.803	1	1316.803	3.24	b ₉	1316.798	3.5
658.903	2	1316.798	3.9	b ₉	1316.798	-0.2
1314.787	1	1314.787	1.12	b ₉ *	1314.783	3.7
657.896	2	1314.784	1.3	b ₉ *	1314.783	0.9
653.323	2	1305.638	0.8	b ₁₁ -acyl-H ₂ O	1305.638	-0.4
633.817	2	1266.626	2.1	GKWSKXSKWS	1266.627	-1.1
618.813	2	1236.617	2.3	b ₁₀ -acyl	1236.617	0.6
1236.618	1	1236.618	1.4	b ₁₀ -acyl	1236.617	1.2
614.311	2	1227.615	6.1	y ₉	1227.616	-1.0
1227.620	1	1227.620	1.9	y ₉	1227.616	2.8
1218.610	1	1218.610	1.1	b ₁₀ -acyl-H ₂ O	1218.606	3.0

609.807	2	1218.607	1.4	b10-acyl-H ₂ O	1218.606	0.5
605.306	2	1209.605	2.2	y ₉ -H ₂ O or KWSKXSKWS	1209.606	-0.7
1209.607	1	1209.607	1.4	y ₉ -H ₂ O or KWSKXSKWS	1209.606	1.1
1205.728	1	1205.728	2.2	c ₈	1205.730	-1.6
1203.711	1	1203.711	0.7	c ₈ *	1203.714	-2.4
1188.689	1	1188.689	2.7	b ₈	1188.703	-12.0
1186.675	1	1186.675	0.7	b ₈ *	1186.688	-10.5
590.302	2	1179.597	1.1	GKWSKXSKW	1179.595	1.2
1179.586	1	1179.586	1.2	GKWSKXSKW	1179.595	-7.5
1170.691	1	1170.691	1.1	b ₈ -H ₂ O	1170.693	-1.1
1161.587	1	1161.587	0.9	GKWSKXSKW-H ₂ O	1161.585	1.8
1099.519	1	1099.519	6.5	WSKXSKWS+H ₂ O	1099.521	-2.4
1081.509	1	1081.509	9.9	WSKXSKWS	1081.511	-1.8
1063.497	1	1063.497	2.5	WSKXSKWS-H ₂ O	1063.500	-3.0
1050.534	1	1050.534	7.6	b ₉ -acyl	1050.537	-2.8
1032.522	1	1032.522	0.7	b ₉ -acyl-H ₂ O	1032.527	-4.5
994.478	1	994.478	3.9	WSKXSKW	994.479	-0.9
993.515	1	993.515	3.6	GKWSKXSK	993.516	-0.4
976.465	1	976.465	1.7	WSKXSKW-H ₂ O	976.468	-3.2
475.234	2	949.459	1.0	a ₁₀ Z ₉	949.457	2.5
949.457	1	949.457	20.1	a ₁₀ Z ₉	949.457	-0.2
928.621	1	928.621	100.0	c ₆	928.624	-2.9
926.605	1	926.605	37.3	c ₆ *	926.608	-2.8
922.440	1	922.440	1.6	b ₈ -acyl	922.442	-2.8
910.611	1	910.611	6.7	b ₆	910.613	-2.7
908.594	1	908.594	1.6	b ₆ *	908.597	-3.2
904.431	1	904.431	0.7	b ₈ -acyl-H ₂ O	904.432	-0.7
900.592	1	900.592	15.4	c ₆ -C ₂ H ₄	900.592	0.0
882.579	1	882.579	0.8	c ₆ -H ₂ O-C ₂ H ₄	882.581	-2.1
865.420	1	865.420	2.6	GKWSKXS	865.421	-1.3
847.409	1	847.409	1.2	GKWSKXS-H ₂ O	847.410	-1.7
404.704	2	808.401	2.5	y ₆ -H ₂ O or KXSKWS KWSKXS SKXSKW WSKXSK	808.399	1.6
808.398	1	808.398	9.8	y ₆ -H ₂ O or KXSKWS KWSKXS SKXSKW WSKXSK	808.399	-1.2

790.388	1	790.388	1.1	y_6-2H_2O or KXSKWS-H ₂ O KWSKXS-H ₂ O SKXSKW-H ₂ O WSKXSK-H ₂ O	790.389	-1.3
782.517	1	782.517	2.3	b ₅	782.518	-1.3
780.501	1	780.501	0.7	b ₅ *	780.502	-2.1
695.485	1	695.485	3.1	b ₄	695.486	-1.4
693.471	1	693.471	1.0	b ₄ *	693.470	0.3
680.304	1	680.304	6.6	WSKXS or XSKWS	680.304	-0.1
677.475	1	677.475	1.7	b ₄ -H ₂ O	677.475	-0.8
675.460	1	675.460	0.6	b ₄ -H ₂ O*	675.460	0.1
662.363	1	662.363	1.6	b ₆ -acyl+H ₂ O	662.363	0.9
662.294	1	662.294	5.2	WSKXS-H ₂ O or XSKWS-H ₂ O	662.294	0.6
644.283	1	644.283	2.4	WSKXS-2H ₂ O or XSKWS-2H ₂ O	644.283	-0.7
575.262	1	575.262	4.1	WSKX-H ₂ O or XSKW-H ₂ O	575.262	0.9
547.267	1	547.267	6.1	WSKX-H ₂ O-28 or XSKW-H ₂ O-28	547.267	0.7
530.241	1	530.241	3.6	a ₇ z ₉ -H ₂ O	530.241	-0.8
491.396	1	491.396	1.0	b ₃ -H ₂ O	491.396	-0.4
420.225	1	420.225	1.7	y ₃ or WSK+H ₂ O SKW+H ₂ O KWS+H ₂ O	420.225	1.6

[†]X denotes an ester linked kynurine; Nex 1 refers to the Nex peptide with an oleoyl (18:1(9)) residue at the N-terminal glycine, Nex 2 refers to the Nex peptide with a stearoyl (18:0) residue at the N-terminal glycine; * indicates the oleoyl (18:1) form of the product ion, unmarked refers to ions derived from Nex 2 or either form of precursor ion; “-acyl” indicates the loss of either the oleoyl or stearoyl moiety from the product ion. The spectra was internally calibrated based on ions $[M+3H]^{3+}$, y₉, c₆ and their isotopes.

Appendix F. Peptide Assignments for the AspN Enzymatic Digest of Peracetylated Bovine Carbonic Anhydrase II.

m/z ^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment ^b	Calculated [m+H] ⁺ /u	Error (ppm)
native BCA						
639.272	1	639.272	33.2	173-177	639.277	-8
320.142	2	639.276	33.2	173-177	639.277	-2
444.218	2	887.428	15.3	23-30 N-D	887.422	7
934.425	1	934.425	21.8	128-136	934.427	-2
467.723	2	934.438	21.8	128-136	934.427	12
312.150	3	934.434	21.8	128-136	934.427	8
494.176	2	987.344	21.8	100-108	987.401	-58
329.792	3	987.360	21.8	100-108	987.401	-42
503.233	2	1005.458	15.0	10-17 N-D	1005.454	4
335.827	3	1005.465	15.0	10-17 N-D	1005.454	12
1024.557	1	1024.557	37.7	178-187	1024.567	-10
512.789	2	1024.570	37.7	178-187	1024.567	3
1067.447	1	1067.447	25.0	61-69 N-D	1067.443	4
534.227	2	1067.446	25.0	61-69 N-D	1067.443	3
1068.605	1	1068.605	28.3	150-159	1068.605	0
534.796	2	1068.584	28.3	150-159	1068.605	-19
384.182	3	1150.530	15.0	1-9 N-D	1150.518	11
1273.753	1	1273.753	39.4	137-149	1273.751	1
637.383	2	1273.758	39.4	137-149	1273.751	5
455.258	3	1363.758	39.4	160-172	1363.743	11
684.386	2	1367.764	33.3	150-162	1367.753	8
715.362	2	1429.716	27.8	18-30	1429.707	6
477.246	3	1429.722	27.8	18-30	1429.707	11
1430.7	1	1430.700	28.2	18-30 N-D	1430.691	6
715.853	2	1430.698	28.2	18-30 N-D	1430.691	5
477.577	3	1430.715	28.2	18-30 N-D	1430.691	17
822.923	2	1644.838	45.4	173-187	1644.827	7
829.413	2	1657.818	30.7	18-32	1657.818	0
553.273	3	1657.803	30.7	18-32	1657.818	-9
619.657	3	1856.955	24.2	23-39 N-D	1856.935	11
706.992	3	2118.960	16.3	1-17 succinimide intermediate	2118.939	10
712.669	3	2135.991	15.3	1-17.	2135.970	10
713.003	3	2136.993	16.5	1-17 N-D	2136.954	19
1095.097	2	2189.186	42.0	128-149	2189.160	12
730.396	3	2189.172	42.0	128-149	2189.160	6
767.759	3	2301.261	32.5	40-60 N-D	2301.238	10
1153.07	2	2305.132	35.6	80-99	2305.115	7
769.052	3	2305.140	35.6	80-99	2305.115	11
777.084	3	2329.236	26.0	109-127	2329.220	7

1210.586	2	2420.164	35.7	80-100	2420.142	9
807.396	3	2420.172	35.7	80-100	2420.142	13
1116.909	3	3348.711	34.0	40-69	3348.679	10
1155.262	3	3463.770	42.0	40-70	3463.706	18
Acetate						
562.283	1	562.283	30.0	18-22 N-D	562.287	-7
639.266	1	639.266	33.1	173-177	639.277	-18
320.141	2	639.274	33.1	173-177	639.277	-5
757.447	1	757.447	23.1	74-79	757.445	2
379.226	2	757.444	23.1	74-79	757.445	-2
802.43	1	802.430	18.9	33-39	802.431	-1
401.721	2	802.434	18.9	33-39	802.431	5
444.218	2	887.428	15.3	23-30 N-D	887.422	7
934.429	1	934.429	21.6	128-136	934.427	3
467.719	2	934.430	21.6	128-136	934.427	4
329.791	3	987.357	21.6	100-108	987.401	-45
1024.543	1	1024.543	37.7	178-187	1024.567	-24
512.771	2	1024.534	37.7	178-187	1024.567	-32
1030.537	1	1030.537	25.0	31-39	1030.542	-4
515.777	2	1030.546	25.0	31-39	1030.542	5
524.241	2	1047.474	18.7	10-17 N-D	1047.464	9
349.830	3	1047.474	18.7	10-17 N-D	1047.464	10
1067.441	1	1067.441	26.0	61-69 N-D	1067.443	-2
534.226	2	1067.444	26.0	61-69 N-D	1067.443	1
1110.613	1	1110.613	32.0	150-159	1110.615	-2
555.791	2	1110.574	32.0	150-159	1110.615	-37
370.877	3	1110.615	32.0	150-159	1110.615	0
574.817	2	1148.626	32.0	163-172 -acetate	1148.616	9
383.526	3	1148.562	32.0	163-172 -acetate	1148.616	-46
1190.625	1	1190.625	22.6	163-172	1190.626	-1
595.823	2	1190.638	22.6	163-172	1190.626	10
397.552	3	1190.640	22.6	163-172	1190.626	12
1315.782	1	1315.782	44.0	137-149	1315.762	15
658.39	2	1315.772	44.0	137-149	1315.762	8
439.263	3	1315.773	44.0	137-149	1315.762	9
1409.714	1	1409.714	36.4	150-152	1409.764	-35
705.377	2	1409.746	36.4	150-162	1409.764	-12
470.597	3	1409.775	36.4	150-162	1409.764	8
1429.74	1	1429.740	27.5	18-30	1429.707	23
715.362	2	1429.716	27.5	18-30	1429.707	6
477.245	3	1429.719	27.5	18-30	1429.707	9
1430.691	1	1430.691	28.0	18-30 N-D	1430.691	0
715.84	2	1430.672	28.0	18-30 N-D	1430.691	-13
477.576	3	1430.712	28.0	18-30 N-D	1430.691	15
1489.766	1	1489.766	27.7	160-172	1489.774	-6
745.396	2	1489.784	27.7	160-172	1489.774	7
497.262	3	1489.770	27.7	160-172	1489.774	-3
530.917	3	1590.735	29.5	10-22 N-D	1590.734	1

822.923	2	1644.838	44.8	173-187	1644.827	7
829.414	2	1657.820	30.8	18-32	1657.818	1
553.280	3	1657.824	30.8	18-32	1657.818	4
829.913	2	1658.818	31.1	18-32 N-D	1658.802	10
553.619	3	1658.841	31.1	18-32 N-D	1658.802	24
905.957	2	1810.906	35.2	163-177	1810.886	11
604.308	3	1810.908	35.2	163-177	1810.886	12
949.976	2	1898.944	26.3	23-39 N-D	1898.945	-1
633.657	3	1898.955	26.3	23-39 N-D	1898.945	5
740.682	3	2220.030	21.9	1-17.	2219.991	18
1110.998	2	2220.988	23.2	1-17 N-D	2220.975	6
741.006	3	2221.002	23.2	1-17 N-D	2220.975	12
1116.093	2	2231.178	46.0	128-149	2231.171	3
744.404	3	2231.196	46.0	128-149	2231.171	11
769.048	3	2305.128	35.9	80-99	2305.115	6
1172.117	2	2343.226	36.2	40-60 N-D	2343.249	-10
781.761	3	2343.267	36.2	40-60 N-D	2343.249	8
1204.188	2	2407.368	47.0	137-159	2407.360	4
803.131	3	2407.377	47.0	137-159	2407.360	7
807.402	3	2420.190	36.0	80-100	2420.142	20
1015.199	3	3043.581	38.1	74-99	3043.543	13
Trifluoroacetate						
562.283	1	562.283	27.5	18-22 N-D	562.287	-8
639.275	1	639.275	33.1	173-177	639.277	-3
320.148	2	639.287		173-177	639.277	16
887.426	1	887.426	30.3	23-30 N-D	887.422	5
447.222	2	893.436	33.5	74-79	893.420	17
934.456	1	934.456	21.3	128-136	934.427	31
467.721	2	934.434		128-136	934.427	8
494.182	2	987.356	21.2	100-108	987.401	-46
329.795	3	987.370		100-108	987.401	-32
1024.541	1	1024.541	37.7	178-187	1024.567	-26
512.767	2	1024.526		178-187	1024.567	-40
534.231	2	1067.454	24.9	61-69 N-D	1067.443	10
549.770	2	1098.533	25.5	31-39	1098.529	3
366.846	3	1098.524	25.5	31-39	1098.529	-5
558.276	2	1115.544	21.8	23-32 N-D	1115.533	10
372.497	3	1115.476	21.8	23-32 N-D	1115.533	-51
1178.603	1	1178.603	36.0	150-159	1178.603	0
589.804	2	1178.601	36.0	150-159	1178.603	-1
393.542	3	1178.611	36.0	150-159	1178.603	7
408.537	3	1223.594	23.0	71-79	1223.538	46
692.385	2	1383.762	47.3	137-149	1383.749	9
461.919	3	1383.742	47.3	137-149	1383.749	-5

706.883	2	1412.758	29.0	18-30 succinimide intermediate	1412.676	58
715.359	2	1429.710	26.5	18-30	1429.707	2
477.232	3	1429.682	26.5	18-30	1429.707	-18
715.855	2	1430.702	27.5	18-30 N-D	1430.691	7
477.579	3	1430.723	27.5	18-30 N-D	1430.691	22
739.381	2	1477.753	39.6	150-162	1477.751	2
547.605	3	1640.800	30.8	18-32 succinimide intermediate	1640.787	8
1644.814	1	1644.814	45.2	173-187	1644.827	-8
822.925	2	1644.842	45.2	173-187	1644.827	9
829.412	2	1657.815	30.6	18-32	1657.818	-2
553.275	3	1657.809	30.6	18-32	1657.818	-5
829.913	2	1658.817	31.6	18-32 N-D	1658.802	9
553.619	3	1658.841	31.6	18-32 N-D	1658.802	24
983.940	2	1966.871	38.2	23-39 N-D	1966.934	-32
656.305	3	1966.901	38.2	23-39 N-D	1966.934	-17
1150.083	2	2299.158	49.0	128-149	2299.158	0
767.070	3	2299.196	49.0	128-149	2299.158	16
1060.535	3	3179.590	37.8	74-99	3179.517	23
PEG						
562.279	1	562.279	15.4	18-22 N-D	562.287	-14
639.232	1	639.232	33.0	173-177	639.277	-71
444.221	2	887.433	15.9	23-30 N-D	887.422	13
446.247	2	891.487	23.2	74-79 -PEG	891.540	-59
934.404	1	934.404	22.6	128-136	934.427	-24
467.725	2	934.441	22.6	128-136	934.427	16
312.145	3	934.420	22.6	128-136	934.427	-7
978.505	1	978.505	25.5	33-39	978.535	-31
489.752	2	978.495	25.5	33-39	978.535	-41
494.241	2	987.475	22.6	100-108	987.401	74
1024.542	1	1024.542	38.0	178-187	1024.567	-25
512.796	2	1024.583	38.0	178-187	1024.567	16
1067.444	1	1067.444	26.0	61-69 N-D	1067.443	1
534.216	2	1067.423	26.0	61-69 N-D	1067.443	-18
1109.640	1	1109.640	31.9	74-79	1109.655	-13
555.334	2	1109.660	31.9	74-79	1109.655	4
1115.526	1	1115.526	23.0	23-32 N-D	1115.533	-6
1206.639	1	1206.639	25.7	31-39	1206.646	-6
603.831	2	1206.653	24.0	31-39	1206.646	6
1223.560	1	1223.560	24.0	10-17 N-D	1223.569	-7
612.290	2	1223.573	24.0	10-17 N-D	1223.569	3

408.525	3	1223.560	24.0	10-17 N-D	1223.569	-8
1286.703	1	1286.703	35.8	150-159	1286.720	-14
643.852	2	1286.696	35.8	150-159	1286.720	-19
1368.665	1	1368.665	23.1	1-9 N-D	1368.633	23
684.820	2	1368.632	23.1	1-9 N-D	1368.633	-1
456.877	3	1368.615	23.1	1-9 N-D	1368.633	-13
715.364	2	1429.720	26.8	18-30	1429.707	9
477.256	3	1429.753	26.8	18-30	1429.707	32
1430.682	1	1430.682	27.8	18-30 N-D	1430.691	-6
715.830	2	1430.651	27.8	18-30 N-D	1430.691	-28
477.578	3	1430.719	27.8	18-30 N-D	1430.691	19
746.447	2	1491.887	45.1	137-149	1491.867	13
497.969	3	1491.892	45.1	137-149	1491.867	17
500.955	3	1500.850	25.5	163-172 -PEG	1500.825	16
820.899	2	1640.790	30.1	18-32 succinimide intermediate	1640.787	2
547.608	3	1640.808	30.1	18-32 succinimide intermediate	1640.787	13
1644.823	1	1644.823	45.5	173-187	1644.827	-2
822.926	2	1644.844	45.5	173-187	1644.827	11
829.414	2	1657.820	31.1	18-32	1657.818	1
1658.817	1	1658.817	32.2	18-32 N-D	1658.802	9
829.897	2	1658.786	32.2	18-32 N-D	1658.802	-10
553.615	3	1658.829	32.2	18-32 N-D	1658.802	16
1718.911	1	1718.911	35.1	163-172	1718.941	-18
573.654	3	1718.946	35.1	163-172	1718.941	3
673.357	3	2018.054	35.0	160-172	2018.089	-17
1038.030	2	2075.051	30.6	23-39 N-D	2075.050	1
692.359	3	2075.060	30.6	23-39 N-D	2075.050	5
769.054	3	2305.147	35.9	80-99	2305.115	14
780.416	3	2339.232	39.2	163-177	2339.200	13
803.117	3	2407.335	46.9	128-149	2407.276	25
807.398	3	2420.177	36.0	80-100	2420.142	15
1260.180	2	2519.352	37.3	40-60 N-D	2519.354	-1
840.454	3	2519.347	37.3	40-60 N-D	2519.354	-3
1287.096	2	2573.183	27.7	1-17 N-D	2573.184	0
858.405	3	2573.201	27.7	1-17 N-D	2573.184	6
922.469	3	2765.390	35.1	109-127 -PEG	2765.451	-22
995.203	3	2983.595	38.0	109-127	2983.566	9

^a bold-face type indicates peptides for which MS/MS data was obtained.

^b N-D indicates deamidation of asparagine residue.

Appendix G. Peptide Assignments for the GluC Enzymatic Digest of Perpegylated Bovine Carbonic Anhydrase II.

m/z ^a	z	Experimental [m+H] ⁺ /u	Elution time (min)	Assignment ^b	Calculated [m+H] ⁺ /u	Error (ppm)
PEG						
646.877	2	1292.745	40.2	204-212	1292.735	8
431.592	3	1292.761	40.2	204-212	1292.735	20
651.370	2	1301.733	27	110-116	1301.720	10
434.591	3	1301.757	27	110-116	1301.720	29
744.935	2	1488.861	41.5	41-52	1488.856	4
496.970	3	1488.895	41.5	41-52	1488.856	26
834.900	2	1668.792	33.4	14-25	1668.802	-6
556.944	3	1668.817	33.4	14-25	1668.802	9
835.404	2	1669.800	34	14-25 N-D	1669.786	9
557.277	3	1669.816	34	14-25 N-D	1669.786	18
850.934	2	1700.861	34.1	117-128	1700.864	-2
567.636	3	1700.892	34.1	117-128	1700.864	16
425.980	4	1700.897	34.1	117-128	1700.864	20
1737.842	1	1737.842	34.1	189-203	1737.873	-18
869.438	2	1737.868	34.1	189-203	1737.873	-3
579.970	3	1737.893	34.1	189-203	1737.873	11
874.888	2	1748.768	29.7	1-13 succinimide intermediate	1748.790	-13
583.599	3	1748.783	29.7	1-13 succinimide intermediate	1748.790	-4
437.958	4	1748.808	29.7	1-13 succinimide intermediate	1748.790	10
877.436	2	1753.864	26.4	106-116	1753.922	-33
585.301	3	1753.889	26.4	106-116	1753.922	-19
883.400	2	1765.793	26.2	1-13.	1765.793	0
589.272	3	1765.801	26.2	1-13.	1765.793	5
883.896	2	1766.785	26.9	1-13 N-D	1766.777	5
589.595	3	1766.770	26.9	1-13 N-D	1766.777	-4
601.284	3	1801.836	24	53-68 succinimide intermediate	1801.864	-15
451.223	4	1801.870	24	53-68 succinimide intermediate	1801.864	4
606.967	3	1818.885	23.4	53-68	1818.866	10
455.478	4	1818.890	23.4	53-68	1818.866	13
910.449	2	1819.890	23.9	53-68 N-D	1819.850	22
607.297	3	1819.875	23.9	53-68 N-D	1819.850	13
455.718	4	1819.850	23.9	53-68 N-D	1819.850	0
525.259	4	2098.013	23.9	53-70 N-D	2097.941	35

662.309	4	2646.213	33.7	1-18 N-D	2646.201	5
679.344	4	2714.351	38.5	213-234	2714.371	-7
920.534	3	2759.587	49.5	138-160	2759.569	7
690.658	4	2759.610	49.5	138-160	2759.569	15
980.827	3	2940.464	38.1	213-236	2940.466	-1
735.889	4	2940.532	38.1	213-236	2940.466	23
1057.273	3	3169.804	40.6	237-259	3169.792	4
793.209	4	3169.813	40.6	237-259	3169.792	7
1120.283	3	3358.835	43	26-52	3358.824	3
840.471	4	3358.860	43	26-52	3358.824	11
951.257	4	3802.006	46.1	204-232	3802.024	-5

^a bold-face type indicates peptides for which MS/MS data was obtained.

^b N-D indicates deamidation of asparagine residue.

REFERENCES

1. McCarty, M. Discovering Genes Are Made of DNA. *Journal of Experimental Medicine*, **1946**, 83, 89.
2. Avery, O.; MacLeod, C. and McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. *Journal of Experimental Medicine*, **1944**, 79, 297.
3. Pauling, L. and Corey, R. Atomic Coordinates and Structure Factors for Two Helical Configurations of Polypeptide Chains. *Proceedings of the National Academy of Sciences of the United States of America*, **1951**, 37, 235.
4. Pauling, L. and Corey, R. The Pleated Sheet, a New Layer Configuration of Polypeptide Chains. *Proceedings of the National Academy of Sciences of the United States of America*, **1951**, 37, 251.
5. Chothia, C.; Levitt, M. and Richardson, D. Structure of Proteins: Packing of α -Helices and Pleated Sheets. *Proceedings of the National Academy of Sciences USA*, **1977**, 74, 4130.
6. Epstein, C.; Goldberger, R. and Anfinsen, C. The Genetic Control of Tertiary Protein Structure: Studies with Model Systems in *Cold Spring Harbor Symposia on Quantitative Biology*. 1963: Cold Spring Harbor Laboratory Press.
7. Givol, D.; De Lorenzo, F.; Goldberger, R. and Anfinsen, C. Disulfide Interchange and the Three-Dimensional Structure of Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **1965**, 53, 676.
8. Klotz, I.; Langebman, N. and Dahnall, D. Quaternary Structure of Proteins. *Annual Review of Biochemistry*, **1970**, 39, 25.
9. Svedberg, T.; Linderstrom-Lang, K.; Adair, G.; Pedersen, K.; Philpot, F.; Philpot, J.; Dodwell, E.; Small, P.; Wrinch, D. and Neuberger, A. A Discussion on the Protein Molecule. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **1939**, 127, 1.
10. Strader, C.; Fong, T.; Tota, M.; Underwood, D. and Dixon, R. Structure and Function of G Protein-Coupled Receptors. *Annual Review of Biochemistry*, **1994**, 63, 101.
11. Ophardt, Amino Acids, Site: *Virtual Chembook*. Date Accessed: 15/01/2010. URL: <http://www.elmhurst.edu/~chm/vchembook/562review.html>
12. Farazi, T.A.; Waksman, G. and Gordon, J.I. The Biology and Enzymology of Protein N-Myristoylation. *The Journal of Biological Chemistry*, **2001**, 276, 39501.
13. Mann, M. and Jensen, O.N. Proteomic Analysis of Post-Translational Modifications. *Nature Biotechnology*, **2003**, 21, 255.
14. Stadtman, E.R. Protein Modification in Aging. *The Journal of Gerontology*, **1988**, 43, B112.
15. Parsons, R.B. and Austen, B.M. Protein Lipidation of BACE. *Biochemical Society Transactions*, **2005**, 33, 1091.
16. Veronese, F.M. Peptide and Protein Pegylation: A Review of Problems and Solutions. *Biomaterials*, **2001**, 22, 405.

17. Yang, J.; Gitlin, I.; Krishnamurthy, V.M.; Vazquez, J.A.; Costello, C.E. and Whitesides, G.M. Synthesis of Monodisperse Polymers from Proteins. *Journal of the American Chemical Society*, **2003**, *125*, 12392.
18. Xie, B.; Luo, X.; Zhao, C.; Priest, C.M.; Chan, S.-Y.; O'Connor, P.B.; Kirschner, D.A. and Costello, C.E. Molecular Characterization of Myelin Protein Zero in *Xenopus Laevis* Peripheral Nerve: Equilibrium between Non-Covalently Associated Dimer and Monomer. *International Journal of Mass Spectrometry*, **2007**, *298*, 304.
19. Cohen, P. Review Lecture: Protein Phosphorylation and Hormone Action. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **1988**, *234*, 115.
20. Liu, C.; Li, Y.; Semenov, M.; Han, C.; Baeg, G.; Tan, Y.; Zhang, Z.; Lin, X. and He, X. Control of β -Catenin Phosphorylation/Degradation by a Dual-Kinase Mechanism. *Cell*, **2002**, *108*, 837.
21. Cohen, P. The Role of Protein Phosphorylation in Human Health and Disease. *European Journal of Biochemistry*, **2001**, *268*, 5001.
22. Kouzarides, T. Acetylation: A Regulatory Modification to Rival Phosphorylation? *The EMBO Journal*, **2000**, *19*, 1176.
23. Polevoda, B. and Sherman, F. The Diversity of Acetylated Proteins. *Genome Biology*, **2002**, *3*, 1.
24. Megee, P.; Morgan, B.; Mittman, B. and Smith, M. Genetic Analysis of Histone H4: Essential Role of Lysines Subject to Reversible Acetylation. *Science*, **1990**, *247*, 841.
25. Glozak, M.; Sengupta, N.; Zhang, X. and Seto, E. Acetylation and Deacetylation of Non-Histone Proteins. *Gene*, **2005**, *363*, 15.
26. Sadoul, K.; Boyault, C.; Pabion, M. and Khochbin, S. Regulation of Protein Turnover by Acetyltransferases and Deacetylases. *Biochimie*, **2008**, *90*, 306.
27. Grunstein, M. Histone Acetylation in Chromatin Structure and Transcription. *Nature*, **1997**, *389*, 349.
28. Dompierre, J.; Godin, J.; Charrin, B.; Cordelieres, F.; King, S.; Humbert, S. and Saudou, F. Histone Deacetylase-6 Inhibition Compensates for the Transport Deficit in Huntington's Disease by Increasing Tubulin Acetylation. *Journal of Neuroscience*, **2007**, *27*, 3571.
29. Barnes, P.; Adcock, I. and Ito, K. Histone Acetylation and Deacetylation: Importance in Inflammatory Lung Diseases. *European Respiratory Journal*, **2005**, *25*, 552.
30. Timmermann, S.; Lehrmann, H.; Polesskaya, A. and Harel-Bellan, A. Histone Acetylation and Disease. *Cellular and Molecular Life Sciences*, **2001**, *58*, 728.
31. Marshall, R.D. Glycoproteins. *Annual Review of Biochemistry*, **1972**, *41*, 673.
32. Parodi, A.J. Protein Glucosylation and Its Role in Protein Folding. *Annual Review of Biochemistry*, **2000**, *69*, 69.
33. Marquardt, T. and Denecke, J. Congenital Disorders of Glycosylation: Review of Their Molecular Bases, Clinical Presentations and Specific Therapies. *European Journal of Pediatrics*, **2003**, *162*, 359.
34. Jaeken, J. and Matthijs, G. Congenital Disorders of Glycosylation: A Rapidly Expanding Disease Family. *Annual Review of Genomics and Human Genetics*, **2007**, *8*, 261.

35. Brownlee, M.; Vlassara, H. and Cerami, A. Nonenzymatic Glycosylation and the Pathogenesis of Diabetic Complications. *Annals of Internal Medicine*, **1984**, *101*, 527.
36. Brownlee, M., Michael Advanced Protein Glycosylation in Diabetes and Aging. *Annual Review of Medicine*, **1995**, *46*, 223.
37. Folch, J. and Lees, M. Proteolipids, a New Type of Tissue Lipoproteins. *The Journal of Biological Chemistry*, **1951**, *191*, 807.
38. Gagnon, J.; Finch, P.; Wood, D.D. and Moscarello, M.A. Isolation of a Highly Purified Myelin Protein. *Biochemistry*, **1971**, *10*, 4756.
39. Wirtz, J. and Kolter, T. Novel Tools for the Proteomic Identification of Acylated Proteins. *ChemBioChem*, **2007**, *8*, 1631.
40. Towler, D.A.; Gordon, J.I.; Adams, S.P. and Glaser, L. The Biology and Enzymology of Eukaryotic Protein Acylation. *Annual Review of Biochemistry*, **1988**, *57*, 69.
41. Zhang, F.L. and Casey, P.J. Protein Prenylation: Molecular Mechanisms and Functional Consequences. *Annual Review of Biochemistry*, **1996**, *65*, 241.
42. Porter, J.A.; Young, K.E. and Beachy, P.A. Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science*, **1996**, *274*, 255.
43. Mann, R.K. and Beachy, P.A. Cholesterol Modification of Proteins. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, **2000**, *1529*, 188.
44. Englund, P.T. The Structure and Biosynthesis of Glycosyl Phosphatidylinositol Protein Anchors. *Annual Review of Biochemistry*, **1993**, *62*, 121.
45. Casey, P.J. Protein Lipidation in Cell Signalling. *Science*, **1995**, *268*, 221.
46. Bross, P.; Corydon, T.J.; Andresen, B.S.; Jørgensen, M.M.; Bolund, L. and Gregersen, N. Protein Misfolding and Degradation in Genetic Diseases. *Human Mutation*, **1999**, *14*, 186.
47. Kayed, R.; Sokolov, Y.; Edmonds, B.; McIntire, T.M.; Milton, S.C.; Hall, J.E. and Glabe, C.G. Permeabilization of Lipid Bilayers Is a Common Conformation-Dependent Activity of Soluble Amyloid Oligomers in Protein Misfolding Diseases. *Journal of Biological Chemistry*, **2004**, *279*, 46363.
48. Marshall, C. Protein Prenylation: A Mediator of Protein-Protein Interactions. *Science*, **1993**, *259*, 1865.
49. Rak, A.; Pylypenko, O.; Niculae, A.; Pyatkov, K.; Goody, R.S. and Alexandrov, K. Structure of the Rab7:Rep-1 Complex: Insights into the Mechanism of Rab Prenylation and Choroideremia Disease. *Cell*, **2004**, *117*, 749.
50. Chistie, W.W., The Lipid Library, Site: *Proteolipids*. Date Accessed: 15/01/2010. URL: <http://lipidlibrary.aocs.org/Lipids/protlip/index.htm>
51. Yang, C.; Gu, Z.-W.; Yang, M.; Lin, S.-N.; Siuzdak, G. and Smith, C.V. Identification of Modified Tryptophan Residues in Apolipoprotein B-100 Derived from Copper Ion-Oxidized Low-Density Lipoprotein. *Biochemistry*, **1999**, *38*, 15903.
52. Moss, G.P., Nomenclature of Lipids. Date Accessed: 2004. URL: <http://www.chem.qmul.ac.uk/iupac/lipid/>
53. Dietze, E.C.; Grillo, M.P.; Kalhorn, T.; Nieslanik, B.S.; Jochheim, C.M. and Atkins, W.M. Thiol Ester Hydrolysis Catalyzed by Glutathione S-Transferase A1-1. *Biochemistry*, **1998**, *37*, 14948.

54. Delgado, C.; Francis, G. and Fisher, D. The Uses and Properties of Peg-Linked Proteins. *Critical Reviews in Therapeutic Drug Carrier Systems*, **1992**, 9, 249.
55. Wu, L. and Payne, G. Biofabrication: Using Biological Materials and Biocatalysts to Construct Nanostructured Assemblies. *Trends in Biotechnology*, **2004**, 22, 593.
56. Sachon, E.; Nielsen, P.F. and Jensen, O.N. Characterization of N-Palmitoylated Human Growth Hormone by in Situ Liquid-Liquid Extraction and MALDI Tandem Mass Spectrometry. *Journal of Mass Spectrometry*, **2007**, 42, 724.
57. Ben Mohamed, L.; Wechsler, S.L. and Nesburn, A.B. Lipopeptide Vaccines--Yesterday, Today, and Tomorrow. *Lancet Infectious Diseases*, **2002**, 2, 425.
58. Morishita, M.; Aikawa, R.; Katsuragi, S.; Yamamoto, Y. and Sugimoto, K., Suppository Preparation Having Excellent Absorption Property. **1989**, United States, 4,873,087.
59. Hederos, S.; Karkson, B.; Tegler, L. and Broo, K.S. Ligand-Directed Labeling of a Single Lysine Residue in Hg²⁺ A1-1 Mutants. *Bioconjugate Chemistry*, **2005**, 16, 1009.
60. Kay, R.; Harris, D. and Entenman, C. Quantification of the Ninhydrin Color Reaction as Applied to Paper Chromatography. *Archives of Biochemistry and Biophysics*, **1956**, 63, 14.
61. Bayer, E.; Grom, E.; Kaltenegger, B. and Uhmman, R. Separation of Amino Acids by High Performance Liquid Chromatography. *Analytical Chemistry*, **1976**, 48, 1106.
62. Snyder, L.; Dolan, J. and Gant, J. Gradient Elution in High-Performance Liquid Chromatography:: I. Theoretical Basis for Reversed-Phase Systems. *Journal of Chromatography A*, **1979**, 165, 3.
63. Fullmer, C. and Wasserman, R. Analytical Peptide Mapping by High Performance Liquid Chromatography. Application to Intestinal Calcium-Binding Proteins. *Journal of Biological Chemistry*, **1979**, 254, 7208.
64. Tarr, G. and Crabb, J. Reverse-Phase High-Performance Liquid Chromatography of Hydrophobic Proteins and Fragments Thereof. *Analytical Biochemistry*, **1983**, 131, 99.
65. Tiselius, A. A New Apparatus for Electrophoretic Analysis of Colloidal Mixtures. *Transactions of the Faraday Society*, **1937**, 33, 524.
66. Shapiro, A.L.; Vinuela, E. and Maizel Jr, J.V. Molecular Weight Estimation of Polypeptide Chains by Electrophoresis in SDS-Polyacrylamide Gels. *Biochemical and Biophysical Research Communications*, **1967**, 28, 815.
67. Weber, K. and Osborn, M. The Reliability of Molecular Weight Determinations by Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *The Journal of Biological Chemistry*, **1969**, 244, 4406.
68. Chrambach, A.; Reisfeld, R.; Wyckoff, M. and Zaccari, J. A Procedure for Rapid and Sensitive Staining of Protein Fractionated by Polyacrylamide Gel Electrophoresis. *Analytical Biochemistry*, **1967**, 20, 150.
69. Burnette, W. "Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate-Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A. *Analytical Biochemistry*, **1981**, 112, 195.

70. Moeremans, M.; Daneels, G. and De Mey, J. Sensitive Colloidal Metal (Gold or Silver) Staining of Protein Blots on Nitrocellulose Membranes. *Analytical Biochemistry*, **1985**, *145*, 315.
71. Gronborg, M.; Kristiansen, T.; Stensballe, A.; Andersen, J.; Ohara, O.; Mann, M.; Jensen, O. and Pandey, A. A Mass Spectrometry-Based Proteomic Approach for Identification of Serine/Threonine-Phosphorylated Proteins by Enrichment with Phospho-Specific Antibodies: Identification of a Novel Protein, Frigg, as a Protein Kinase a Substrate. *Molecular & Cellular Proteomics*, **2002**, *1*, 517.
72. O'Farrell, P. High Resolution Two-Dimensional Electrophoresis of Proteins. *Journal of Biological Chemistry*, **1975**, *250*, 4007.
73. Everaerts, F. and Verheggen, T. Isotachopheresis: Electrophoretic Analysis in Capillaries. *Journal of Chromatography A*, **1970**, *53*, 315.
74. Niessen, W.; Tjaden, U. and Van der Greef, J. Capillary Electrophoresis-Mass Spectrometry. *Journal of Chromatography*, **1993**, *636*, 3.
75. Cai, J. and Henion, J. Capillary Electrophoresis-Mass Spectrometry. *Journal of Chromatography A*, **1995**, *703*, 667.
76. Gordon, M.; Huang, X.; Pentoney, S. and Zare, R. Capillary Electrophoresis. *Science*, **1988**, *242*, 224.
77. Stroink, T.; Paarlberg, E.; Waterval, J.; Bult, A. and Underberg, W. On-Line Sample Preconcentration in Capillary Electrophoresis, Focused on the Determination of Proteins and Peptides: A Review. *Electrophoresis*, **2001**, *22*, 2374.
78. Edman, P. Method for Determination of the Amino Acid Sequence in Peptides. *Acta Chemica Scandinavica*, **1950**, *4*, 283.
79. Niall, H. Automated Edman Degradation: The Protein Sequenator. *Methods in Enzymology*, **1973**, *27*, 942.
80. Shen, S.; Matsubae, M.; Takao, T.; Tanaka, N. and Komatsu, S. A Proteomic Analysis of Leaf Sheaths from Rice. *Journal of Biochemistry*, **2002**, *132*, 613.
81. Bruschi, S.; West, K.; Crabb, J.; Gupta, R. and Stevens, J. Mitochondrial Hsp60 (P1 Protein) and a Hsp70-Like Protein (Mortalin) Are Major Targets for Modification During S-(1, 1, 2, 2-Tetrafluoroethyl)-L-Cysteine-Induced Nephrotoxicity. *Journal of Biological Chemistry*, **1993**, *268*, 23157.
82. Stolowitz, M. Chemical Protein Sequencing and Amino Acid Analysis. *Current Opinion in Biotechnology*, **1993**, *4*, 9.
83. Wüthrich, K. Protein Structure Determination in Solution by NMR Spectroscopy. *Journal of Biological Chemistry*, **1990**, *265*, 22059.
84. Franklin, R.E. and Gosling, R.G. Molecular Configuration in Sodium Thymonucleate. *Nature*, **1953**, *171*, 740.
85. Wang, B.-C., X-Ray Crystallography, Site: *Centre for Metalloenzyme Studies*. Date Accessed: 17/02/2010. URL: <http://www.uga.edu/cms/XCry.html>
86. Nagayama, K.; Kumar, A.; Wüthrich, K. and Ernst, R. Experimental Techniques of Two-Dimensional Correlated Spectroscopy. *Journal of Magnetic Resonance*, **1980**, *40*, 321.
87. Wüthrich, K.; Wider, G.; Wagner, G. and Braun, W. Sequential Resonance Assignments as a Basis for Determination of Spatial Protein Structures by High

- Resolution Proton Nuclear Magnetic Resonance. *Journal of Molecular Biology*, **1982**, 155, 311.
88. Degtyarev, M.Y.; Spiegel, A.M. and Jones, T.L.Z. The G Protein .Alpha.S Subunit Incorporates [3h]Palmitic Acid and Mutation of Cysteine-3 Prevents This Modification. *Biochemistry*, **1993**, 32, 8057.
 89. Roth, A.F.; Feng, Y.; Chen, L. and Davis, N.G. The Yeast Dhhc Cysteine-Rich Domain Protein Akr1p Is a Palmitoyl Transferase. *Journal of Cell Biology*, **2002**, 159, 23.
 90. Cooks, R.G.; Glish, G.L.; Mc Luckey, S.A. and Kaiser, R.E. Ion Trap Mass Spectrometry. *Chemical and Engineering News*, **1991**, 69.
 91. Wells, J.M. and McLuckey, S.A. Collision-Induced Dissociation (CID) of Peptides and Proteins. *Methods Enzymology*, **2005**, 402, 148.
 92. Yost, R.A. and Enke, C.G. Selected Ion Fragmentation with a Tandem Quadrupole Mass Spectrometer. *Journal of the American Chemical Society*, **1978**, 100, 2274.
 93. March, R.E. An Introduction to Quadrupole Ion Trap Mass Spectrometry. *Journal of Mass Spectrometry*, **1997**, 32, 351.
 94. Laskin, J. and Futrell, J.H. Collisional Activation of Peptide Ions in FT-ICR Mass Spectrometry. *Mass Spectrometry Reviews*, **2003**, 22, 158.
 95. Tang, X.; Ens, W.; Mayer, F.; Standing, K.G.; Westmore, J.B. and Boyd, R.K. Measurement of Unimolecular Decay in Peptides of Masses Greater Than 1200 Units by a Reflecting Time-of-Flight Mass Spectrometer. *Rapid Communications in Mass Spectrometry*, **1989**, 3, 443.
 96. Cole, R.B. Some Tenets Pertaining to Electrospray Ionization Mass Spectrometry. *Journal of Mass Spectrometry*, **2000**, 35, 763.
 97. Dole, M.; Mack, L.L.; Hines, R.L.; Mobley, R.C.; Ferguson, L.D. and Alice, M.B. Molecular Beams of Macroions. *The Journal of Chemical Physics*, **1968**, 49, 2240.
 98. Yamashita, M. and Fenn, J.B. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *The Journal of Physical Chemistry*, **1984**, 88, 4451.
 99. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F. and Whitehouse, C.M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*, **1989**, 246, 64.
 100. Fenn, J.B. Electrospray Wings for Molecular Elephants in *Nobel Lectures 2002*. 2002.
 101. Wang, G. and Cole, R.B. Charged Residue Versus Ion Evaporation for Formation of Alkali Metal Halide Cluster Ions in ESI. *Analytica Chimica Acta*, **2000**, 406, 53.
 102. Thomson, B. and Iribarne, J. Field Induced Ion Evaporation from Liquid Surfaces at Atmospheric Pressure. *Journal of Chemical Physics*, **1979**, 71, 4451.
 103. Scripps, What Is Mass Spectrometry? Site: *Scripps Centre for Mass Spectrometry*. Date Accessed: 15/02/2010. (Senior director: Gary Siuzdak) URL: http://masspec.scripps.edu/mshistory/whatisms_details.php
 104. Wilm, M. and Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry*, **1996**, 68, 1.

105. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yochida, Y. and Toshida, T. Protein and Polymer Analyses up to m/z 100 000 by Lazer Ionization Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*, **1988**, 2, 151.
106. Karas, M. and Hillenkamp, F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons. *Analytical Chemistry*, **1988**, 60, 2299.
107. Wiley, W.C. and McLaren, I.H. Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments*, **1955**, 26, 1150.
108. Spengler, B. and Cotter, R.J. Ultraviolet Laser Desorption/Ionization Mass Spectrometry of Proteins above 100,000 Daltons by Pulsed Ion Extraction Time-of-Flight Analysis. *Analytical Chemistry*, **1990**, 62, 793.
109. Juhasz, P.; Roskey, M.T.; Smirnov, I.P.; Haff, L.A.; Vestal, M.L. and Martin, S.A. Applications of Delayed Extraction Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry to Oligonucleotide Analysis. *Analytical Chemistry*, **1996**, 68, 941.
110. Kaufmann, R.; Kirscha, D. and Spengler, B. Sequencing of Peptides in a Time-of-Flight Mass Spectrometer: Evaluation of Postsource Decay Following Matrix-Assisted Laser Desorption Ionisation (MALDI). *International Journal of Mass Spectrometry and Ion Processes*, **1994**, 131, 355.
111. Thomson, J.J. Rays of Positive Electricity. *Proceedings of the Royal Society A*, **1913**, 89, 1.
112. Bateman, R. and Burns, P., Magnetic Sector Mass Spectrometer. **1988**, United States, Patent Number 4,727,249.
113. Wollnik, H. Time-of-Flight Mass Analyzers. *Mass Spectrometry Reviews*, **1993**, 12, 89.
114. Wollnik, H. Energy-Isochronous Time-of-Flight Mass Analyzers. *International Journal of Mass Spectrometry and Ion Processes*, **1994**, 131, 387.
115. Comisarow, M.B. and Marshall, A.G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chemical Physics Letters*, **1974**, 25, 282.
116. Marshall, A.G.; Hendrickson, C.L. and Jackson, G.S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrometry Reviews*, **1998**, 17, 1.
117. Jonscher, K.R. and Yates III, J.R. The Quadrupole Ion Trap Mass Spectrometer—a Small Solution to a Big Challenge. *Analytical Biochemistry*, **1997**, 244, 1.
118. Douglas, D.; Frank, A. and Mao, D. Linear Ion Traps in Mass Spectrometry. *Mass Spectrometry Reviews*, **2005**, 24, 1.
119. Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry*, **2000**, 72, 1156.
120. Fei, X.; Wei, G. and Murray, K.K. Aerosol MALDI with a Reflectron Time-of-Flight Mass Spectrometer. *Analytical Chemistry*, **1996**, 68, 1143.
121. Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V. and Zagulin, V.A. The Mass-Reflectron, a New Nonmagnetic Time-of-Flight Mass Spectrometer with High Resolution. *Soviet Physics JETP*, **1973**, 37, 45.

122. Shevchenko, A.; Loboda, A.; Shevchenko, A.; Ens, W. and Standing, K.G. MALDI Quadrupole Time-of-Flight Mass Spectrometry: A Powerful Tool for Proteomic Research. *Analytical Chemistry*, **2000**, 72, 2132.
123. Louris, J.; Wright, L.; Cooks, R. and Schoen, A. New Scan Modes Accessed with a Hybrid Mass Spectrometer. *Analytical Chemistry*, **1985**, 57, 2918.
124. Schlosser, A.; Pipkorn, R.; Bossemeyer, D. and Lehmann, W. Analysis of Protein Phosphorylation by a Combination of Elastase Digestion and Neutral Loss Tandem Mass Spectrometry. *Analytical Chemistry*, **2001**, 73, 170.
125. Shi, S.; Drader, J.; Freitas, M.; Hendrickson, C. and Marshall, A.G. Comparison and Interconversion of the Two Most Common Frequency-to-Mass Calibration Functions for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *International Journal of Mass Spectrometry*, **2000**, 195-196, 591.
126. Zubarev, R.; Kelleher, N. and McLafferty, F. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society*, **1998**, 120, 3265.
127. McLafferty, F.; Horn, D.; Breuker, K.; Ge, Y.; Lewis, M.; Cerda, B.; Zubarev, R. and Carpenter, B. Electron Capture Dissociation of Gaseous Multiply Charged Ions by Fourier-Transform Ion Cyclotron Resonance. *Journal of the American Society for Mass Spectrometry*, **2001**, 12, 245.
128. Stone, D.; Hemling, M.; Carr, S.; Horn, D.; Lindh, I. and McLafferty, F. Phosphopeptide/Phosphoprotein Mapping by Electron Capture Dissociation Mass Spectrometry. *Analytical Chemistry*, **2001**, 73, 19.
129. Hakansson, K.; Cooper, H.; Emmett, M.; Costello, C.; Marshall, A. and Nilsson, C. Electron Capture Dissociation and Infrared Multiphoton Dissociation MS/MS of an N-Glycosylated Tryptic Peptide to Yield Complementary Sequence Information. *Analytical Chemistry*, **2001**, 73, 4530.
130. Mirgorodskaya, E.; Roepstorff, P. and Zubarev, R. Localization of O-Glycosylation Sites in Peptides by Electron Capture Dissociation in a Fourier Transform Mass Spectrometer. *Analytical Chemistry*, **1999**, 71, 4431.
131. Creese, A. and Cooper, H. Liquid Chromatography Electron Capture Dissociation Tandem Mass Spectrometry (LC-ECD-MS/MS) Versus Liquid Chromatography Collision-Induced Dissociation Tandem Mass Spectrometry (LC-CID-MS/MS) for the Identification of Proteins. *Journal of the American Society for Mass Spectrometry*, **2007**, 18, 891.
132. Jebanathirajah, J.A.; Pittman, J.L.; Thomson, B.A.; Budnik, B.A.; Kaur, P.; Rape, M.; Kirschner, M.; Costello, C.E. and O'Connor, P.B. Characterization of a New qQq-FTICR Mass Spectrometer for Post-Translational Modification Analysis and Top-Down Tandem Mass Spectrometry of Whole Proteins. *Journal of the American Society for Mass Spectrometry*, **2005**, 16, 1985.
133. Paul, W. and Steinwedel, H., Apparatus for Separating Charged Particles of Different Specific Charges. **1956**, Germany, 944900.
134. Paul, W. Electromagnetic Traps for Charged and Neutral Particles. *Reviews of Modern Physics*, **1990**, 62, 531.
135. Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K. and Horning, S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Analytical Chemistry*, **2006**, 78, 2113.

136. Dawson, P.; Hedman, J. and Whetten, N. A Simple Mass Spectrometer. *Review of Scientific Instruments*, **1969**, 40, 1444.
137. Lawson, G.; Bonner, R. and Todd, J. The Quadrupole Ion Store (Quistor) as a Novel Source for a Mass Spectrometer. *Journal of Physics E: Scientific Instruments*, **1973**, 6, 357.
138. Stafford, G.; Kelley, P.; Syka, J.; Reynolds, W. and Todd, J. Recent Improvements in and Analytical Applications of Advanced Ion Trap Technology. *International Journal of Mass Spectrometry and Ion Processes*, **1984**, 60, 85.
139. Louris, J.; Cooks, R.G.; Syka, J.; Kelley, P.; Stafford Jr, G. and Todd, J. Instrumentation, Applications, and Energy Deposition in Quadrupole Ion-Trap Tandem Mass Spectrometry. *Analytical Chemistry*, **1987**, 59, 1677.
140. Kaiser, R.; Cooks, R.G.; Stafford, G.; Syka, J. and Hemberger, P. Operation of a Quadrupole Ion Trap Mass Spectrometer to Achieve High Mass/Charge Ratios. *International Journal of Mass Spectrometry and Ion Processes*, **1991**, 106, 79.
141. Louris, J.; Brodbelt-Lustig, J.; Cooks, R.G.; Glish, G.; van Berkel, G. and McLuckey, S. Ion Isolation and Sequential Stages of Mass Spectrometry in a Quadrupole Ion Trap Mass Spectrometer. *International Journal of Mass Spectrometry and Ion Processes*, **1990**, 96, 117.
142. Louris, J.; Amy, J.; Ridley, T. and Cooks, R.G. Injection of Ions into a Quadrupole Ion Trap Mass Spectrometer. *International Journal of Mass Spectrometry and Ion Processes*, **1989**, 88, 97.
143. Soni, M. and Cooks, R.G. Selective Injection and Isolation of Ions in Quadrupole Ion Trap Mass Spectrometry Using Notched Waveforms Created Using the Inverse Fourier Transform. *Analytical Chemistry*, **1994**, 66, 2488.
144. Church, D. Storage Ring Ion Trap Derived from the Linear Quadrupole Radio Frequency Mass Filter. *Journal of Applied Physics*, **1969**, 40, 3127.
145. Schwartz, J.; Senko, M. and Syka, J. A Two-Dimensional Quadrupole Ion Trap Mass Spectrometer. *Journal of the American Society for Mass Spectrometry*, **2002**, 13, 659.
146. Senko, M.; Hendrickson, C.; Emmett, M.; Shi, S. and Marshall, A. External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, **1997**, 8, 970.
147. Syka, J.; Marto, J.; Bai, D.; Horning, S.; Senko, M.; Schwartz, J.; Ueberheide, B.; Garcia, B.; Busby, S. and Muratore, T. Novel Linear Quadrupole Ion Trap/FT Mass Spectrometer: Performance Characterization and Use in the Comparative Analysis of Histone H3 Post-Translational Modifications. *Journal of Proteome Research*, **2004**, 3, 621.
148. Perry, R.; Cooks, R.G. and Noll, R. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrometry Reviews*, **2008**, 27, 661.
149. Thermo, Pamphlet: *Thermo Scientific LTQ Orbitrap Velos*. 2009.
150. Hardman, M. and Makarov, A. Interfacing the Orbitrap Mass Analyzer to an Electrospray Ion Source. *Analytical Chemistry*, **2003**, 75, 1699.
151. Anderson, J.S.; Svensson, B. and Roepstorff, P. Electrospray Ionization and Matrix Assisted Laser Desorption/Ionisation Mass Spectrometry: Powerful

- Analytical Tools in Recombinant Protein Chemistry. *Nature Biotechnology*, **1996**, *14*, 449.
152. Mann, M.; Hendrickson, R.C. and Pandey, A. Analysis of Proteins and Proteomes by Mass Spectrometry. *Annual Review of Biochemistry*, **2001**, *70*, 437.
 153. Mann, M. and Wilm, M. Electrospray Mass Spectrometry for Protein Characterization. *Trends in Biochemical Sciences*, **1995**, *20*, 219.
 154. Nguyen, D.N.; Becker, G.W. and Rigglin, R.M. Protein Mass Spectrometry: Applications to Analytical Biotechnology. *Journal of Chromatography A*, **1995**, *705*, 21.
 155. Biemann, K. Mass Spectrometry of Peptides and Proteins. *Annual Review of Biochemistry*, **1992**, *61*, 977.
 156. Shimadzu, Accuspot-Nano Spotter for MALDI-TOF-MS. Date Accessed: 17-02-2010. URL: <http://www.shimadzu.com/products>
 157. Beuhler, R.J.; Flanigan, E.; Greene, L.J. and Friedman, L. Proton Transfer Mass Spectrometry of Peptides. Rapid Heating Technique for Underivatized Peptides Containing Arginine. *Journal of the American Chemical Society*, **1974**, *96*, 3990.
 158. McLafferty, F.W. and Bockhoff, F.M. Separation/Identification System for Complex Mixtures Using Mass Separation and Mass Spectral Characterization. *Analytical Chemistry*, **1978**, *50*, 69.
 159. McLafferty, F.W. Tandem Mass Spectrometry (MS/MS): A Promising New Analytical Technique for Specific Component Determination in Complex Mixtures. *Accounts of Chemical Research*, **1980**, *13*, 33.
 160. Hunt, D.F.; Yates III, J.; Shabanowitz, J.; Winston, S. and Hauer, C.R. Protein Sequencing by Tandem Mass Spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, **1986**, *83*, 6233.
 161. Papayannopoulos, I.A. The Interpretation of Collision-Induced Dissociation Tandem Mass Spectra of Peptides. *Mass Spectrometry Reviews*, **1995**, *14*, 49.
 162. Biemann, K. Contributions of Mass Spectrometry to Peptide and Protein Structure. *Biomedical and Environmental Mass Spectrometry*, **1988**, *16*, 99.
 163. Roepstorff, P. and Fohlman, J. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomedical Mass Spectrometry*, **1984**, *11*, 601.
 164. Johnson, R.S. and Biemann, K. Computer Program (Seqpep) to Aid in the Interpretation of High-Energy Collision Tandem Mass Spectra of Peptides. *Biological Mass Spectrometry*, **1989**, *18*, 945.
 165. Larsen, M.R.; Trelle, M.B.; Thingholm, T.E. and Jensen, O.N. Analysis of Posttranslational Modifications of Proteins by Tandem Mass Spectrometry. *BioTechniques*, **2006**, *40*, 790.
 166. Ogorzalek Loo, R.R.; Cavalcoli, J.D.; VanBogelen, R.A.; Mitchell, C.; Loo, J.A.; Moldover, B. and Andrews, P.C. Virtual 2-D Gel Electrophoresis: Visualization and Analysis of the E. Coli Proteome by Mass Spectrometry. *Analytical Chemistry*, **2001**, *73*, 4063.
 167. Di Guilmi, A.M.; Mouz, N.; Pétilot, Y.; Forest, E.; Dideberg, O. and Vernet, T. Deacylation Kinetics Analysis of *Streptococcus pneumoniae* Penicillin-Binding

- Protein 2x Mutants Resistant to Beta-Lactam Antibiotics Using Electrospray Ionization- Mass Spectrometry. *Analytical Biochemistry*, **2000**, 284, 240.
168. Fenaille, F.; Tabet, J.-C. and Guy, P.A. Study of Peptides Containing Modified Lysine Residues by Tandem Mass Spectrometry: Precursor Ion Scanning of Hexanal-Modified Peptides. *Rapid Communications in Mass Spectrometry*, **2004**, 18, 67.
 169. Annan, R.S.; Huddleston, M.J.; Verma, R.; Deshaies, R.J. and Carr, S.A. A Multidimensional Electrospray MS-Based Approach to Phosphopeptide Mapping. *Analytical Chemistry*, **2000**, 73, 393.
 170. Smith, C.M.; Gafken, P.R.; Zhang, Z.; Gottschling, D.E.; Smith, J.B. and Smith, D.L. Mass Spectrometric Quantification of Acetylation at Specific Lysines within the Amino-Terminal Tail of Histone H4. *Analytical Biochemistry*, **2003**, 316, 23.
 171. Zhang, K.; Yau, P.M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B.S. and Bradbury, M.E. Differentiation between Peptides Containing Acetylated or Tri-Methylated Lysines by Mass Spectrometry: An Application for Determining Lysine 9 Acetylation and Methylation of Histone H3. *Proteomics*, **2004**, 4, 1.
 172. Suckau, D.; Mak, M. and Przybylski, M. Protein Surface Topology-Probing by Selective Chemical Modification and Mass Spectrometric Peptide Mapping. *Proceedings of the National Academy of Sciences of the United States of America*, **1992**, 89, 5630.
 173. Glocker, M.O.; Borchers, C.; Fiedler, W.; Suckau, D. and Przybylski, M. Molecular Characterization of Surface Topology in Protein Tertiary Structures by Amino-Acylation and Mass Spectrometric Peptide Mapping. *Bioconjugate Chemistry*, **1994**, 5, 583.
 174. Jagannadham, M.V. and Nagaraj, R. Detection of Peptides Covalently Modified with Multiple Fatty Acids by MALDI-TOF Mass Spectrometry. *Journal of Peptide Research*, **2005**, 66, 94.
 175. Ogawa, Y.; Quagliarotti, G.; Jordan, J.; Taylor, C.W.; Starbuck, W.C. and Busch, H. Structural Analysis of the Glycine-Rich, Arginine-Rich Histone III. *Journal of Biological Chemistry*, **1969**, 244, 4387.
 176. Van Rhijn, I.; Young, D.C.; De Jong, A.; Vazquez, J.; Cheng, T.-Y.; Talekar, R.; Barral, D.C.; Leon, L.; Brenner, M.B.; Katz, J.T.; Riese, R.; Ruprecht, R.M.; O'Connor, P.B.; Costello, C.E.; Porcelli, S.A.; Briken, V. and Moody, D.B. CD1c Bypasses Lysosomes to Present a Lipopeptide Antigen with 12 Amino Acids. *Journal of Experimental Medicine*, **2009**, 206, 1409.
 177. O'Connor, P.B.; Budnik, B.A.; Ivleva, V.B.; Kaur, P.; Moyer, S.C.; Pittman, J.L. and Costello, C.E. A High Pressure Matrix-Assisted Laser Desorption Ion Source for Fourier Transform Mass Spectrometry Designed to Accommodate Large Targets with Diverse Surfaces. *Journal of the American Society for Mass Spectrometry*, **2004**, 15, 128.
 178. Pittman, J.L.; Thomson, B.A.; Budnik, B.A.; Cournoyer, J.J.; Fallows, E.; Jebanathirajah, J.A.; Moyer, S.C.; Costello, C.E. and O'Connor, P.B. A Novel Hybrid Instrument Using a Commercial Electrospray Ionization Source with a High-Performance FTMS for Proteomics Applications in *Proceedings of the 52nd ASMS Conference on Mass Spectrometry and Allied Topics*. 2004, Nashville, Tennessee.

179. O'Connor, P.B., Boston University School of Medicine, **2000**, Boston, MA. (Personal communication.)
180. Yousefi-Salakdeh, E.; Johansson, J. and Stramberg, R. A Method for S- and O-Palmitoylation of Peptides: Synthesis of Pulmonary Surfactant Protein-C Models. *Biochemical Journal*, **1999**, 343, 557.
181. Burlingame, A.L., Protein Prospector. Date Accessed: 2003-2010. URL: <http://prospector.ucsf.edu>
182. Hosada, H.; Kojima, M.; Mizushima, T.; Shimizu, S. and Kangawa, K. Structural Divergence of Human Ghrelin: Identification of Multiple Ghrelin-Derived Molecules Produced by Post-Translational Processing. *Journal of Biological Chemistry*, **2003**, 278, 67.
183. Zhou, H.; Watts, J.D. and Aebersold, R. A Systematic Approach to the Analysis of Protein Phosphorylation. *Nature Biotechnology*, **2001**, 19, 375.
184. Zaia, J. Mass Spectrometry of Oligosaccharides. *Mass Spectrometry Reviews*, **2004**, 23, 161.
185. Bizzozero, O.A. and Lees, M.B. Fatty Acid Composition of Myelin Proteolipid Protein During Vertebrate Evolution. *Neurochemical Research*, **1999**, 24, 269.
186. Bizzozero, O.A.; Bixler, H.A.; Davis, J.D.; Espinosa, A. and Messier, A.M. Chemical Deacylation Reduces the Adhesive Properties of Proteolipid Protein and Leads to Decompaction of the Myelin Sheath. *Journal of Neurochemistry*, **2001**, 76, 1129.
187. Mumby, S.M. Reversible Palmitoylation of Signaling Proteins. *Current Opinion in Cell Biology*, **1997**, 9, 148.
188. Duncan, J.A. and Gilman, A.G. Autoacylation of G Protein α Subunits. *The Journal of Biological Chemistry*, **1996**, 271, 23594.
189. Bano, M.C.; Jackson, C.S. and Magee, A.I. Pseudo-Enzymatic S-Acylation of a Myristoylated Yes Protein Tyrosine Kinase Peptide in Vitro May Reflect Non-Enzymatic S-Acylation in Vivo. *Biochemical Journal*, **1998**, 330, 723–731.
190. Smotrys, J.E. and Linder, M.E. Palmitoylation of Intracellular Signaling Proteins: Regulation and Function. *Annual Review of Biochemistry*, **2004**, 73, 559.
191. Belanger, C.; Ansanay, H.; Qanbar, R. and Bouvier, M. Primary Sequence Requirements for S-Acylation of Beta(2)-Adrenergic Receptor Peptides. *FEBS Letters*, **2001**, 499, 59.
192. El-Husseini, A.E.-D. and Bredt, D.S. Protein Palmitoylation: A Regulator of Neuronal Development and Function. *Nature Reviews Neuroscience*, 2002, 3, 791.
193. Basar, T.; Havlicek, V.; Bezouskova, S.; Halada, P.; Hackett, M. and Sebo, P. The Conserved Lysine 860 in the Additional Fatty-Acylation Site of Bordetella Pertussis Adenylate Cyclase Is Crucial for Toxin Function Independently of Its Acylation Status. *Journal of Biological Chemistry*, **1999**, 274, 10777.
194. Gonzalo, S. and Linder, M.E. Snap-25 Palmitoylation and Plasma Membrane Targeting Require a Functional Secretory Pathway. *Molecular Biology of the Cell*, **1998**, 9, 585.
195. Gonzalo, S.; Greentree, W.K. and Linder, M.E. Snap-25 Is Targeted to the Plasma Membrane through a Novel Membrane-Binding Domain. *Journal of Biological Chemistry*, **1999**, 274, 21313.

196. Rouquette-Jazdanian, A.K.; Pelassy, C.; Breittmayer, J.-P.; Cousin, J.-L. and Aussel, C. Metabolic Labelling of Membrane Microdomains/Rafts in Jurkat Cells Indicates the Presence of Glycerophospholipids Implicated in Signal Transduction by the Cd3 T-Cell Receptor. *Biochemical Journal*, **2002**, 363, 645.
197. Corvi, M.M.; Soltys, C.-L.M. and Berthiaume, L.G. Regulation of Mitochondrial Carbamoyl-Phosphate Synthetase 1 Activity by Active Site Fatty Acylation. *Journal of Biological Chemistry*, **2001**, 276, 45704.
198. Veit, M.; Sachs, K.; Heckelmann, M.; Maretzki, D.; Hofmann, K.P. and Schmidt, M.F.G. Palmitoylation of Rhodopsin with S-Protein Acyltransferase: Enzyme Catalyzed Reaction Versus Autocatalytic Acylation. *Biochimica et Biophysica Acta: Lipids and Lipid Metabolism*, **1998**, 1394, 90.
199. Liang, X.; Lu, Y.; Neubert, T.A. and Resh, M.D. Mass Spectrometric Analysis of GAP-43/Neuromodulin Reveals the Presence of a Variety of Fatty Acylated Species. *Journal of Biological Chemistry*, **2002**, 277, 33032.
200. Fivaz, M. and Meyer, T. Specific Localization and Timing in Neuronal Signal Transduction Mediated by Protein-Lipid Interactions. *Neuron*, **2003**, 40, 319.
201. Resh, M.D. Fatty Acylation of Proteins: New Insights into Membrane Targeting of Myristoylated and Palmitoylated Proteins. *Biochimica et Biophysica Acta*, **1999**, 1451, 1.
202. Dietrich, L.E.; Gurezka, R.; Veit, M. and Ungermann, C. The SNARE Ykt6 Mediates Protein Palmitoylation During an Early Stage of Homotypic Vacuole Fusion. *The EMBO Journal*, **2004**, 23, 45.
203. De, B.K.; Woolfitt, A.R.; Barr, J.R.; Daneshvar, M.I.; Sampson, J.S.; Ades, E.W. and Carlone, G.M. Analysis of Recombinant Acylated Pneumococcal Surface Adhesin A of Streptococcus Pneumoniae by Mass Spectrometry. *Archives of Biochemistry and Biophysics*, **2003**, 419, 147.
204. Stevenson, F.T.; Bursten, S.L.; Locksley, R.M. and Lovett, D.H. Myristyl Acylation of the Tumor Necrosis Factor Alpha Precursor on Specific Lysine Residues. *The Journal of Experimental Medicine*, **1992**, 176, 1053.
205. Stevenson, F.T.; Bursten, S.L.; Fanton, C.; Locksley, R.M. and Lovett, D.H. The 31-kDa Precursor of Interleukin 1 Alpha Is Myristoylated on Specific Lysines within the 16-kDa N-Terminal Propiece. *Proceedings of the National Academy of Sciences of the United States of America*, **1993**, 90, 7245.
206. Wilcox, C.; Hu, J.S. and Olson, E.N. Acylation of Proteins with Myristic Acid Occurs Cotranslationally. *Science*, **1987**, 238, 1275.
207. Maurer-Stroh, S.; Gouda, M.; Novatchkova, M.; Schleiffer, A.; Schneider, G.; Sirota, F.L.; Wildpaner, M.; Hayashi, N. and Eisenhaber, F. Myrbase: Analysis of Genome-Wide Glycine Myristoylation Enlarges the Functional Spectrum of Eukaryotic Myristoylated Proteins. *Genome Biology*, **2004**, 5, R21.
208. Sowa, G.; Pypaert, M. and Sessa, W.C. Distinction between Signaling Mechanisms in Lipid Rafts vs. Caveolae. *Proceedings of the National Academy of Sciences of the United States of America*, **2001**, 98, 14072.
209. Selvakumar, P.; Pasha, M.K.; Ashakumary, L.; Dimmock, J.R. and R.K., S. Myristoyl-CoA:Protein N-Myristoyltransferase: A Novel Molecular Approach for Cancer Therapy (Review). *International Journal of Molecular Medicine*, **2002**, 10, 493.
210. Boutin, J.A. Myristoylation. *Cellular Signalling*, **1997**, 9, 15.

211. Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H. and Kangawa, K. Ghrelin Is a Growth-Hormone-Releasing Acylated Peptide from Stomach. *Nature*, **1999**, 402, 656.
212. Kaiya, H.; Kojima, M.; Hosoda, H.; Koda, A.; Yamamoto, K.; Kitajima, Y.; Matsumoto, M.; Minamitake, Y.; Kikuyama, S. and Kangawa, K. Bullfrog Ghrelin Is Modified by N-Octanoic Acid at Its Third Threonine Residue. *Journal of Biological Chemistry*, **2001**, 276, 40441.
213. Yoshihara, F.; Kojima, M.; Hosoda, H.; Nakazato, M. and Kangawa, K. Ghrelin: A Novel Peptide for Growth Hormone Release and Feeding Regulation. *Current Opinion in Clinical Nutrition and Metabolic Care*, **2002**, 5, 391.
214. Iglesias, M.; Piñeiro, R.; Blanco, M.; Gallego, R.; Diéguez, C.; Gualillo, O.; González-Juanatey, J. and Lago, F. Growth Hormone Releasing Peptide (Ghrelin) Is Synthesized and Secreted by Cardiomyocytes. *Cardiovascular Research*, **2004**, 62, 481.
215. Galas, L.; Chartrel, N.; Kojima, M.; Kangawa, K. and Vaudry, H. Immunohistochemical Localization and Biochemical Characterization of Ghrelin in the Brain and Stomach of the Frog *Rana Esculenta*. *The Journal of Comparative Neurology*, **2002**, 450, 34.
216. Lu, S.; Guan, J.-L.; Wang, Q.-P.; Uehara, K.; Yamada, S.; Goto, N.; Date, Y.; Nakazato, M.; Kojima, M. and Shioda, S. Immunocytochemical Observation of Ghrelin-Containing Neurons in the Rat Arcuate Nucleus. *Neuroscience Letters*, **2002**, 321, 157.
217. Baskina, D.G.; Figlewicz Lattemanna, D.; Seeley, R.J.; Woodse, S.C.; Porte Jr., D. and Schwartz, M.W. Insulin and Leptin: Dual Adiposity Signals to the Brain for the Regulation of Food Intake and Body Weight. *Brain Research*, **1999**, 848, 114.
218. Cummings, D. Ghrelin and the Short-and Long-Term Regulation of Appetite and Body Weight. *Physiology & Behavior*, **2006**, 89, 71.
219. Lindroos, A.; Lissner, L.; Carlsson, B.; Carlsson, L.; Torgerson, J.; Karlsson, C.; Stenlof, K. and Sjostrom, L. Familial Predisposition for Obesity May Modify the Predictive Value of Serum Leptin Concentrations for Long-Term Weight Change in Obese Women. *American Journal of Clinical Nutrition*, **1998**, 67, 1119.
220. Gentile, S. Contributing Factors to Weight Gain During Long-Term Treatment with Second-Generation Antipsychotics. A Systematic Appraisal and Clinical Implications. *Obesity Reviews*, **2009**, 10, 527.
221. James, W. The Epidemiology of Obesity: The Size of the Problem. *Journal of Internal Medicine*, **2008**, 263, 336.
222. Cummings, D.E.; Weigle, D.S.; Frayo, R.S.; Breen, P.A.; Ma, M.K.; Dellinger, E.P. and Purnell, J.Q. Plasma Ghrelin Levels after Diet-Induced Weight Loss or Gastric Bypass Surgery. *New England Journal of Medicine*, **2002**, 346, 1623.
223. Tschop, M.; Weyer, C.; Tataranni, P.A.; Devanarayan, V.; Ravussin, E. and Heiman, M.L. Circulating Ghrelin Levels Are Decreased in Human Obesity. *Diabetes*, **2001**, 50, 707.
224. Cummings, D.E.; Clement, K.; Purnell, J.Q.; Vaisse, C.; Foster, K.E.; Frayo, R.S.; Schwartz, M.W.; Basdevant, A. and Weigle, D.S. Elevated Plasma Ghrelin Levels in Prader Willi Syndrome. *Nature Medicine*, **2002**, 8, 643.

225. Haqq, A.M.; Farooqi, I.S.; O'Rahilly, S.; Stadler, D.D.; Rosenfeld, R.G.; Pratt, K.L.; LaFranchi, S.H. and Purnell, J.Q. Serum Ghrelin Levels Are Inversely Correlated with Body Mass Index, Age, and Insulin Concentrations in Normal Children and Are Markedly Increased in Prader-Willi Syndrome. *Journal of Clinical Endocrinology and Metabolism*, **2003**, 88, 174.
226. Blackman, S. The Enormity of Obesity. *The Scientist*, **2004**.
227. Esler, W.; Rudolph, J.; Claus, T.; Tang, W.; Barucci, N.; Brown, S.; Bullock, W.; Daly, M.; DeCarr, L. and Li, Y. Small-Molecule Ghrelin Receptor Antagonists Improve Glucose Tolerance, Suppress Appetite, and Promote Weight Loss. *Endocrinology*, **2007**, 148, 5175.
228. Pasternak, A.; Goble, S.; deJesus, R.; Hreniuk, D.; Chung, C.; Tota, M.; Mazur, P.; Feighner, S.; Howard, A. and Mills, S. Discovery and Optimization of Novel 4-[(Aminocarbonyl) Amino]-N-[4-(2-Aminoethyl) Phenyl] Benzenesulfonamide Ghrelin Receptor Antagonists. *Bioorganic & Medicinal Chemistry Letters*, **2009**, 19, 6237.
229. Liang, X.; Nazarian, A.; Erdjument-Bromage, H.; Bornmann, W.; Tempst, P. and Resh, M.D. Heterogeneous Fatty Acylation of Src Family Kinases with Polyunsaturated Fatty Acids Regulates Raft Localization and Signal Transduction. *Journal of Biological Chemistry*, **2001**, 276, 30987.
230. Nakayama, Y.; Odagaki, Y.; Fujita, S.; Matsuoka, S.; Hamanaka, N.; Nakai, H. and Toda, M. Clarification of Mechanism of Human Sputum Elastase Inhibition by a New Inhibitor, Ono-5046, Using Electrospray Ionization Mass Spectrometry. *Bioorganic and Medicinal Chemistry Letters*, **2002**, 12, 2349.
231. Guan, Z. Identification and Localization of the Fatty Acid Modification in Ghrelin by Electron Capture Dissociation. *Journal of the American Society for Mass Spectrometry*, **2002**, 13, 1443.
232. Bizzozero, O.A.; Malkoski, S.P.; Mobarak, C.; Bixler, H.A. and Evans, J.E. Mass-Spectrometric Analysis of Myelin Proteolipids Reveals New Features of This Family of Palmitoylated Membrane Proteins. *Journal of Neurochemistry*, **2002**, 81, 636.
233. Serebryakova, M.V.; Kordyukova, L.V.; Baratova, L.A. and Markushin, S.G. Mass Spectrometric Sequencing and Acylation Character Analysis of the C-Terminal Anchoring Segment from Influenza a Hemagglutinin. *European Journal of Mass Spectrometry*, **2006**, 12, 51.
234. Hoffman, M.D. and Kast, J. Mass Spectrometric Characterization of Lipid-Modified Peptides for the Analysis of Acylated Proteins. *Journal of Mass Spectrometry*, **2006**, 41, 229.
235. Sparkman, O.D., *Mass Spectrometry Desk Reference*. 2000: Global View Publishing.
236. Chaurand, P.; Luetzenkirchen, F. and Spengler, B. Peptide and Protein Identification by Matrix-Assisted Laser Desorption Ionization (MALDI) and MALDI-Post-Source Decay Time-of-Flight Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, **1999**, 10, 91.
237. Li, L.; Garden, R. and Sweedler, J. Single-Cell MALDI: A New Tool for Direct Peptide Profiling. *Trends in Biotechnology*, **2000**, 18, 151.
238. Schlosser, A. and Lehmann, W.D. Special Feature: Commentary - Five-Membered Ring Formation in Unimolecular Reactions of Peptides: A Key

- Structural Element Controlling Low-Energy Collision-Induced Dissociation of Peptides. *Journal of Mass Spectrometry*, **2000**, 35, 1382.
239. Vaisar, T. and Urban, J. Gas-Phase Fragmentation of Protonated Mono-N-Methylated Peptides. Analogy with Solution-Phase Acid-Catalyzed Hydrolysis. *Journal of Mass Spectrometry*, **1998**, 33, 505.
 240. Jedrzejewski, P.T. and Lehmann, W.D. Detection of Modified Peptides in Enzymatic Digests by Capillary Liquid Chromatography/Electrospray Mass Spectrometry and a Programmable Skimmer CID Acquisition Routine. *Analytical Chemistry*, **1997**, 69, 294.
 241. Francis, M.J. Peptide Vaccines for Viral Diseases. *Science Progress*, **1990**, 74, 115.
 242. Gahery-Segard, H.; Pialoux, G.; Charmeteau, B.; Sermet, S.; Poncelet, H.; Raux, M.; Tartar, A.; Levy, J.-P.; Gras-Masse, H. and Guillet, J.-G. Multiepitopic B- and T-Cell Responses Induced in Humans by a Human Immunodeficiency Virus Type 1 Lipopeptide Vaccine. *Journal of Virology*, **2000**, 74, 1694.
 243. Gahery-Segard, H.; Pialoux, G.; Figueiredo, S.; Igea, C.; Surenaud, M.; Gaston, J.; Gras-Masse, H.; Levy, J.-P. and Guillet, J.-G. Long-Term Specific Immune Responses Induced in Humans by a Human Immunodeficiency Virus Type 1 Lipopeptide Vaccine: Characterization of Cd8⁺-T-Cell Epitopes Recognized. *Journal of Virology*, **2003**, 77, 11220.
 244. Brigl, M. and Brenner, M.B. Cd1: Antigen Presentation and T Cell Function. *Annual Review of Immunology*, **2004**, 22, 817.
 245. Wikipedia, T Cells. Date Accessed: 22/12/2009. URL: http://en.wikipedia.org/wiki/T_cells
 246. Sinkovics, J.G. and Horvath, J.C. Vaccination against Human Cancers (Review). *International Journal of Oncology*, **2000**, 16, 81.
 247. Maeurer, M.J.; Storkus, W.J.; Kirkwood, J.M. and Lotze, M.T. New Treatment Options for Patients with Melanoma: Review of Melanoma-Derived T-Cell Epitope-Based Peptide Vaccines. *Melanoma Research*, **1996**, 6, 11.
 248. Francis, J.N. and Larche, M. Peptide-Based Vaccination: Where Do We Stand? *Current Opinion in Allergy and Clinical Immunology*, **2005**, 5, 537.
 249. Letsch, A.; Keilholz, U.; Fluck, M.; Nagorsen, D.; Asemissen, A.M.; Schmitt, A.; Thiel, E. and Scheibenbogen, C. Peptide Vaccination after Repeated Resection of Metastases Can Induce a Prolonged Relapse-Free Interval in Melanoma Patients. *International Journal of Cancer*, **2005**, 114, 936.
 250. Cox, J.C. and Coulter, A.R. Adjuvants—a Classification and Review of Their Modes of Action. *Vaccine*, **1997**, 15, 248.
 251. Hopp, T.P. Immunogenicity of a Synthetic HBsAg Peptide: Enhancement by Conjugation to a Fatty Acid Carrier. *Molecular Immunology*, **1984**, 21, 13.
 252. Brynestad, K.; Babbitt, B.; Huang, L. and Rouse, B.T. Influence of Peptide Acylation, Liposome Incorporation, Synthetic Immunomodulators on the Immunogenicity of a 1-23 Peptide of Glycoprotein D of Herpes Simplex Virus: Implications for Subunit Vaccines. *Journal of Virology*, **1990**, 64, 680–85.
 253. Andrieu, M.; Loing, E.; Desoutter, J.-F.; Connan, F.; Choppin, J.; Gras-Masse, H.; Hanau, D.; Dautry-Varsat, A.; Guillet, J.-G. and Hosmalin, A. Endocytosis of an HIV-Derived Lipopeptide into Human Dendritic Cells Followed by Class

I-Restricted Cd8⁺ T Lymphocyte Activation. *European Journal of Immunology*, **2000**, 30, 3256.

254. Sprengers, D. and Janssen, H.L.A. Immunomodulatory Therapy for Chronic Hepatitis B Virus Infection. *Fundamental and Clinical Pharmacology*, **2005**, 19, 17.
255. Andrieu, M.; Desoutter, J.-F.; Loing, E.; Gaston, J.; Hanau, D.; Guillet, J.-G. and Hosmalin, A. Two Human Immunodeficiency Virus Vaccinal Lipopeptides Follow Different Cross-Presentation Pathways in Human Dendritic Cells. *Journal of Virology*, **2003**, 77, 1564.
256. Aruna Seth, Y.Y., Helen Jacoby, Joan C. Callery, Stephen M. Kaminsky, Wayne C. Koff, Douglas F. Nixon, Norman L. Letvin Evaluation of a Lipopeptide Immunogen as a Therapeutic in HIV Type 1-Seropositive Individuals. *AIDS Research and Human Retroviruses*, **2000**, 16, 337.
257. Gahery, H., Choppin, J., Bourgault, I., Fischer, E., Maillere, B., Guillet, J.G. HIV Preventive Vaccine Research at the ANRS: The Lipopeptide Vaccine Approach. *Therapie*, **2005**, 60, 243.
258. Partidos, C.D.; Beignon, A.S.; Semetey, V.; Briand, J.P. and S., M. The Bare Skin and the Nose as Non-Invasive Routes for Administering Peptide Vaccines. *Vaccine*, **2001**, 19, 2708.
259. avert.org, The History of AIDS: 1981-1986. Date Accessed: 2004
260. Marx, J.L. New Disease Baffles Medical Community. *Science*, **1982**, 217, 618.
261. Beral, V., Peterman, T.A., Berkelman, R.L., Jaffe, H.W. Kaposi's Sarcoma among Persons with AIDS: A Sexually Transmitted Infection? *Lancet*, **1990**, 335, 123.
262. Lindegren, M.L., Steinberg, S., Byers Jr, R.H. Epidemiology of HIV/AIDS in Children. *Pediatric Clinics of North America*, **2000**, 47, 1.
263. Sher, R. HIV Infection in South Africa, 1982-1988--a Review. *South African Medical Journal*, **1989**, 76, 314.
264. Barre-Sinoussi, F.; J-C., C.; Rey, F.; Nugeyre, M.T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Brun-Vezinet, F.; Rouzioux, C.; Rozenbaum, W. and Montagnier, L. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*, **1983**, 220, 868.
265. Janeway, C.A.; Travers, P.; Walport, M. and Shlomchik, M., *Chapter 11. Failures of the Host Defense Mechanism*, in *Immunobiology*, Janeway, C.A., Editor. 2001, Garland Publishing: New York.
266. Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.; Sharp, P.M.; Shaw, G.M. and Hahn, B.H. Human Infection by Genetically Diverse SIVSM-Related HIV-2 in West Africa. *Nature*, **1992**, 358, 495.
267. Taylor, B.; Sobieszczyk, M.; McCutchan, F. and Hammer, S. The Challenge of HIV-1 Subtype Diversity. *New England Journal of Medicine*, **2008**, 358, 1590.
268. Mann, J.M., *AIDS: A Worldwide Pandemic*, in *Current Topics in AIDS*, Gottlieb, M.S., Jeffries, D.J., Mildvan, D., Pinching, A.J. and Quinn, T.C., Editors. 1989, John Wiley and Sons.
269. UNAIDS, *AIDS Epidemic Update:2009*. 2009, UNAIDS: Geneva.
270. Emlet, C. and Shippy, R. HIV/AIDS Treatments. *Journal of Gerontological Social Work*, **2008**, 50, 131.

271. Sattentau, Q.; Dalgleish, A.; Weiss, R. and Beverley, P. Epitopes of the CD4 Antigen and HIV Infection. *Science*, **1986**, 234, 1120.
272. Stevenson, M.; Stanwick, T.; Dempsey, M. and Lamonica, C. HIV-1 Replication Is Controlled at the Level of T Cell Activation and Proviral Integration. *The EMBO Journal*, **1990**, 9, 1551.
273. Johnston, M. and Fauci, A. An HIV Vaccine--Evolving Concepts. *The New England Journal of Medicine*, **2007**, 356, 2073.
274. Koup, R.; Safrit, J.; Cao, Y.; Andrews, C.; McLeod, G.; Borkowsky, W.; Farthing, C. and Ho, D. Temporal Association of Cellular Immune Responses with the Initial Control of Viremia in Primary Human Immunodeficiency Virus Type 1 Syndrome. *Journal of Virology*, **1994**, 68, 4650.
275. Richman, D.; Wrin, T.; Little, S. and Petropoulos, C. Rapid Evolution of the Neutralizing Antibody Response to HIV Type 1 Infection. *Proceedings of the National Academy of Sciences*, **2003**, 100, 4144.
276. Pantaleo, G.; Graziosi, C.; Demarest, J.F.; Butini, L.; Montroni, M.; Fox, C.H.; Orenstein, J.M.; Kotler, D.P. and Fauci, A.S. HIV Infection Is Active and Progressive in Lymphoid Tissue During the Clinically Latent Stage of Disease. *Nature*, **1993**, 362, 355.
277. Pantaleo, G., Graziosi, C., Fauci, A.S. The Role of Lymphoid Organs in the Pathogenesis of HIV Infection. *Seminars in Immunology*, **1993**, 5, 57.
278. Chun, T.; Engel, D.; Berrey, M.; Shea, T.; Corey, L. and Fauci, A. Early Establishment of a Pool of Latently Infected, Resting CD4⁺ T Cells During Primary HIV-1 Infection. *Proceedings of the National Academy of Sciences*, **1998**, 95, 8869.
279. Tolstrup, M.; Ostergaard, L.; Laursen, A.L.; Pedersen, S.F. and Duch, M. HIV / SIV Escape from Immune Surveillance: Focus on Nef. *Current HIV Research*, **2004**, 2, 141.
280. Vandamme, M., Van Laethem K., De Clercq, E. Managing Resistance to Anti-HIV Drugs: An Important Consideration for Effective Disease Management. *Drugs*, **1999**, 57, 337.
281. medscape, Choosing the Best Antiretroviral Regimen by HIV Resistance Testing?, Site: *Drug Therapy Perspective*. Date Accessed: 2005. URL: www.medscape.com/viewarticle/406397
282. Perelson, A.; Neumann, A.; Markowitz, M.; Leonard, J. and Ho, D. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science*, **1996**, 271, 1582.
283. Rowland-Jones, S., Sutton, J., Ariyoshi, K., Dong, T., Gotch, F., McAdam, S., Whitby, D., Sabally, S., Gallimore, A., Corrah, T. HIV-Specific Cytotoxic T-Cells in HIV-Exposed but Uninfected Gambian Women. *Nature Medicine*, **1995**, 1, 59.
284. Rowland-Jones, S.L.; Dong, T.; Fowke, K.R.; Kimani, J.; Krausa, P.; Newell, H.; Blanchard, T.; Ariyoshi, K.; Oyugi, J.; Ngugi, E.; Bwayo, J.; MacDonald, K.S.; McMichael, A.J. and Plummer, F.A. Cytotoxic T Cell Responses to Multiple Conserved HIV Epitopes in HIV-Resistant Prostitutes in Nairobi. *Journal of Clinical Investigation*, **1998**, 102, 1758.
285. Rathbun, R.; Lockhart, S. and Stephens, J. Current HIV Treatment Guidelines--an Overview. *Current Pharmaceutical Design*, **2006**, 12, 1045.

286. Flexner, C. HIV-Protease Inhibitors. *New England Journal of Medicine*, **1998**, 338, 1281.
287. Ndembu, N.; Abrahams, A.; Pilch, H.; Ichimura, H.; Mbanya, D.; Kaptue, L.; Salata, R. and Arts, E. Molecular Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) and HIV-2 in Yaounde, Cameroon: Evidence of Major Drug Resistance Mutations in Newly Diagnosed Patients Infected with Subtypes Other Than Subtype B. *Journal of Clinical Microbiology*, **2008**, 46, 177.
288. Schiller, D. and Youssef-Bessler, M. Etravirine: A Second-Generation Nonnucleoside Reverse Transcriptase Inhibitor (NNRTI) Active against NNRTI-Resistant Strains of HIV. *Clinical Therapeutics*, **2009**, 31, 692.
289. Appay, V., Nixon, D. F., Donahoe, S. M., Gillespie, G. M., Dong, T., King, A., Ogg, G. S., Spiegel, H. M., Conlon, C., Spina, C. A., Havlir, D. V., Richman, D. D., Waters, A., Easterbrook, P., McMichael, A. J., Rowland-Jones, S. L. HIV-Specific CD8^+ T Cells Produce Antiviral Cytokines but Are Impaired in Cytolytic Function. *Journal of Experimental Medicine*, **2000**, 192, 63–75.
290. Champagne, P., Ogg, G. S., King, A. S., Knabenhans, C., Ellefsen, K., Nobile, M., Appay, V., Rizzardi, G. P., Fleury, S., Lipp, M., Forster, R., Rowland-Jones, S., Sekaly, R. P., McMichael, A. J., Pantaleo, G. Skewed Maturation of Memory HIV-Specific CD8^+ T Lymphocytes. *Nature*, **2001**, 410, 106–111.
291. Kaul, R., Dong, T., Plummer, F. A., Kimani, J., Rostron, T., Kiama, P., Njagi, E., Irungu, E., Farah, B., Oyugi, J., Chakraborty, R., MacDonald, K. S., Bwayo, J. J., McMichael, A., Rowland-Jones, S. L. CD8^+ Lymphocytes Respond to Different HIV Epitopes in Seronegative and Infected Subjects. *Journal of Clinical Investigation*, **2001**, 107, 1303–1310.
292. Appay, V., Dunbar, P. R., Callan, M., Klenerman, P., Gillespie, G. M. A., Papagno, L., Ogg, G. S., King, A., Lechner, F., Spina, C. A., Little, S., Havlir, D. V., Richman, D. D., Gruener, N., Pape, G., Waters, A., Easterbrook, P., Salio, M., Cerundolo, V., McMichael, A. J., Rowland-Jones, S. L. Memory CD8^+ T Cells Vary in Differentiation Phenotype in Different Persistent Virus Infections. *Nature Medicine*, **2002**, 8, 379–385.
293. Mortara, L., Letourneur, F., Gras-Masse, H., Venet, A., Guillet, J. G., Bourgault-Villada, I. Selection of Virus Variants and Emergence of Virus Escape Mutants after Immunization with Epitope Vaccine. *Journal of Virology*, **1998**, 72, 1403–1410.
294. Peng, B. and Robert-Guroff, M. Deletion of N-Terminal Myristoylation Site of HIV Nef Abrogates Both Mhc-1 and CD4 Down-Regulation. *Immunology Letters*, **2001**, 78, 195.
295. Niederman, T.M.; Thielan, B.J. and Ratner, L. Human Immunodeficiency Virus Type 1 Negative Factor Is a Transcriptional Silencer. *Proceedings of the National Academy of Sciences of the United States of America*, **1989**, 86, 1128.
296. Miller, M.; Warmerdam, M.; Gaston, I.; Greene, W. and Feinberg, M. The Human Immunodeficiency Virus-1 Nef Gene Product: A Positive Factor for Viral Infection and Replication in Primary Lymphocytes and Macrophages. *Journal of Experimental Medicine*, **1994**, 179, 101.
297. Spina, C.; Kwok, T.; Chowder, M.; Guatelli, J. and Richman, D. The Importance of Nef in the Induction of Human Immunodeficiency Virus Type 1 Replication from Primary Quiescent CD4^+ Lymphocytes. *Journal of Experimental Medicine*, **1994**, 179, 115.

298. Geyer, M.; Fackler, O.T. and Peterlin, B.M. Structure-Function Relationships in HIV-1 Nef. *EMBO report*, **2001**, 2, 580.
299. Fackler, O.T., Baur, A.S. Live and Let Die - Nef Functions Beyond HIV Replication. *Immunity*, **2002**, 16, 493.
300. Kestler, H.W.; Ringler, D.J.; Mori, K.; Panicali, D.L.; Sehgal, P.K.; Daniel, M.D. and R.C., D. Importance of the Nef Gene for Maintenance of High Virus Loads and for Development of AIDS. *Cell*, **1991**, 65, 651.
301. Daniel, M.D.; Kirchhoff, F.; Czajak, S.C.; Sehgal, P.K. and Desrosiers, R.C. Protective Effects of a Live Attenuated SIV Vaccine with a Deletion in the Nef Gene. *Science*, **1992**, 258, 1938.
302. Whitney, J.B.R., Ruth M Live Attenuated HIV Vaccines: Pitfalls and Prospects. *Current Opinion in Infectious Diseases.*, **2004**, 17, 17.
303. Otake, K., Fujii, Y., Nakaya, Y., Nishino, T., Zhong, Q., Fujinaga, K., Kameola, M., Ohki, K., Ikuta, K. The Carboxyl-Terminal Region of HIV-1 Nef Protein Is a Cell Surface Domain That Can Interact with CD41 T Cells. *Journal of Immunology*, **1994**, 153, 5826–5837.
304. Fujinaga, K., Zhong, Q., Nakaya, T., Kameoka, M., Meguro, T., Yamada, K., and Ikuta, K. Extracellular Nef Protein Regulates Productive HIV-1 Infection from Latency. *Journal of Immunology*, **1995**, 155, 5289–5298.
305. Moody, D.B.; Young, D.C.; Cheng, T.-Y.; Rosat, J.-P.; Roura-mir, C.; O'Connor, P.B.; Zajonc, D.M.; Walz, A.; Miller, M.J.; Lavery, S.B.; Wilson, I.A.; Costello, C.E. and Brenner, M.B. T Cell Activation by Lipopeptide Antigens. *Science*, **2004**, 303, 527.
306. Burdick, D.J., Stults, J.T. Analysis of Peptide Synthesis Products by Electrospray Ionization Mass Spectrometry. *Methods in Enzymology*, **1997**, 289, 499.
307. Anderson, L.B.; Maderia, M.; Oullette, A.J.A.; Putnam-Evans, C.; Higgins, L.; Krick, T.; MacCoss, M.J.; Lim, H.; Yates III, J.R. and Barry, B.A. Posttranslational Modifications in the Cp43 Subunit of Photosystem II. *Proceedings of the National Academy of Sciences of the United States of America*, **2002**, 99, 14676.
308. Berlett, B.S. and Stadtman, E.R. Protein Oxidation in Aging, Disease, and Oxidative Stress. *The Journal of Biological Chemistry*, **1997**, 272, 20313.
309. Simat, T.J. and Steinhart, H. Oxidation of Free Tryptophan and Tryptophan Residues in Peptides and Proteins. *Journal of Agricultural and Food Chemistry*, **1998**, 46, 490.
310. Moody, B.; Rhijn, I.; Young, D. and Costello, C., Methods and Compositions for Immunomodulation. **2008**, United States, Patent number: 20080226587.
311. Gibney, B.R.; Rabanal, F.; Skalicky, J.J.; Wand, A.J. and Dutton, P.L. Design of a Unique Protein Scaffold for Maquettes. *Journal of the American Chemical Society*, **1997**, 119, 2323.
312. Arnold, F. and Haymore, B. Engineered Metal-Binding Proteins: Purification to Protein Folding. *Science*, **1991**, 252, 1796.
313. Håkansson, S.; Viljanen, J. and Broo, K.S. Programmed Delivery of Novel Functional Groups to the Alpha Class Glutathione Transferases. *Biochemistry*, **2003**, 42, 10260.

314. Viljanen, J.; Tegler, L. and Broo, K.S. Combinatorial Chemical Reengineering of the Alpha Class Glutathione Transferases. *Bioconjugate Chemistry*, **2004**, *15*, 718.
315. Raleigh, D.P.; Betz, S.F. and Degrado, W.F. A De-Novo Designed Protein Mimics the Native-State of Natural Proteins. *Journal of the American Chemical Society*, **1995**, *117*, 7558.
316. Shirahama, K.; Tsujii, K. and Takagi, T. Free-Boundary Electrophoresis of Sodium Dodecyl Sulfate-Protein Polypeptide Complexes with Special Reference to SDS-Polyacrylamide Gel Electrophoresis. *Journal of Biochemistry*, **1974**, *75*, 309.
317. Gudiksen, K.L.; Urbach, A.R.; Gitlin, I.; Yang, J.; Vazquez, J.A.; Costello, C.E. and Whitesides, G.M. Influence of the Zn(II) Cofactor on the Refolding of Bovine Carbonic Anhydrase after Denaturation with Sodium Dodecyl Sulfate. *Analytical Chemistry*, **2004**, *76*, 7151.
318. Falick, A.; Hines, W.; Medzihradszky, K.; Baldwin, M. and Gibson, B. Low-Mass Ions Produced from Peptides by High-Energy Collision-Induced Dissociation in Tandem Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, **1993**, *4*, 882.
319. Kim, J.; Kim, K.; Kwon, H. and Yoo, J. Probing Lysine Acetylation with a Modification-Specific Marker Ion Using High-Performance Liquid Chromatography/Electrospray-Mass Spectrometry with Collision-Induced Dissociation. *Analytical Chemistry*, **2002**, *74*, 5443.
320. Dookeran, N.; Yalcin, T. and Harrison, A. Fragmentation Reactions of Protonated-Amino Acids. *Journal of Mass Spectrometry*, **1996**, *31*, 500.
321. Whitesides, G.M. and Krishnamurthy, V.M. Designing Ligands to Bind Proteins. *Quarterly Reviews of Biophysics*, **2005**, *38*, 385.
322. Day, Y.S.N.; Baird, C.L.; Rich, R.L. and Myszk, D.G. Direct Comparison of Binding Equilibrium, Thermodynamic, and Rate Constants Determined by Surface- and Solution-Based Biophysical Methods. *Protein Science*, **2002**, *11*, 1017.
323. Cleland, J.L.; Hedgcock, C. and Wang, D.I. Polyethylene Glycol Enhanced Refolding of Bovine Carbonic Anhydrase B. Reaction Stoichiometry and Refolding Model. *Journal of Biological Chemistry*, **1992**, *267*, 13327.
324. Henderson, L.E.; Henriksson, D. and Nyman, P.O. Primary Structure of Human Carbonic Anhydrase C. *Journal of Biological Chemistry*, **1976**, *251*, 5457.
325. Clarke, S. Propensity for Spontaneous Succinimide Formation from Aspartyl and Asparaginyl Residues in Cellular Proteins. *International Journal of Peptide and Protein Research*, **1987**, *30*, 808.
326. Capasso, S.; Mazzarella, L.; Sica, F. and Zagari, A. Deamidation Via Cyclic Imide in Asparaginyl Peptides. *Peptide Research*, **1989**, *2*, 195.
327. Patel, K. and Borchardt, R.T. Chemical Pathways of Peptide Degradation. II. Kinetics of Deamidation of an Asparaginyl Residue in a Model Hexapeptide. *Pharmaceutical Research*, **1990**, *7*, 703.
328. Geiger, T. and Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides. Succinimide-Linked Reactions That Contribute to Protein Degradation. *Journal of Biological Chemistry*, **1987**, *15*, 785.

329. Meinwald, Y.C.; Stimson, E.R. and Scheraga, H.A. Deamidation of the Asparaginyl-Glycyl Sequence. *International Journal of Peptide and Protein Research*, **1986**, 28, 79.
330. Chelius, D.; Rehder, D.S. and Bondarenko, P.V. Identification and Characterization of Deamidation Sites in the Conserved Regions of Human Immunoglobulin Gamma Antibodies. *Analytical Chemistry*, **2005**, 77, 6004.
331. Cournoyer, J.J.; Pittman, J.L.; Ivleva, V.B.; Fallows, E.; Waskell, L.; Costello, C.E. and O'Connor, P.B. Deamidation: Differentiation of Aspartyl from Isoaspartyl Products in Peptides by Electron Capture Dissociation. *Protein Science*, **2005**, 14, 452.
332. Bischoff, R.; Lepage, P.; Jaquinod, M.; Cauet, G.; Acker-Klein, M.; Clesse, D.; Laporte, M.; Bayol, A.; Van Dorsselaer, A. and Roitsch, C. Sequence-Specific Deamidation: Isolation and Biochemical Characterization of Succinimide Intermediates of Recombinant Hirudin. *Biochemistry*, **2002**, 32, 725.
333. Volkin, D.; Mach, H. and Middaugh, C. Degradative Covalent Reactions Important to Protein Stability. *Molecular Biotechnology*, **1997**, 8, 105.
334. Sun, A.Q.; Yuksel, K.U. and Gracy, R.W. Terminal Marking of Triosephosphate Isomerase: Consequences of Deamidation. *Archives of Biochemistry and Biophysics*, **1995**, 322, 361.
335. Linder, H. and Hellinger, W. Age-Dependent Deamidation of Asparagine Residues in Proteins. *Experimental Gerontology*, **2001**, 36, 1551.
336. Lehmann, W.D.; Schlosser, A.; Erben, G.; Pipkorn, R.; Bossemeyer, D. and Kinzel, V. Analysis of Isoaspartate in Peptides by Electrospray Tandem Mass Spectrometry. *Protein Science*, **2000**, 9, 2260.
337. Sargaeva, N.; Lin, C. and O'Connor, P. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid Peptides, Including a 1- 42, Using Electron- Ion Reactions. *Analytical Chemistry*, **2009**, 489.
338. Sadakane, Y.; Yamazaki, T.; Nakagomi, K.; Akizawa, T.; Fujii, N.; Tanimura, T.; Kaneda, M. and Hatanaka, Y. Quantification of the Isomerization of Asp Residue in Recombinant Human α -Crystallin by Reversed-Phase HPLC. *Journal of Pharmaceutical and Biomedical Analysis*, **2003**, 30, 1825.
339. Chan, W.; Chan, T. and O'Connor, P. Electron Transfer Dissociation with Supplemental Activation to Differentiate Aspartic and Isoaspartic Residues in Doubly Charged Peptide Cations. *Journal of the American Society for Mass Spectrometry*, **2010**, 21, 1012.
340. Wolfenden, R.; Andersson, L.; Cullis, P.M. and Southgate, C.C.B. Affinities of Amino Acid Side Chains for Solvent Water. *Biochemistry*, **2002**, 20, 849.
341. Kyte, J. and Doolittle, R.F. A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology*, **1982**, 157, 105.
342. Beavis, R.C., Prowl - Amino Acid Information, Site: *Hydrophobicity Scales*. Date Accessed: 2009. URL: <http://prowl.rockefeller.edu/aainfo/hydro.htm>