Characteristic Negative Ion Fragmentations of Deprotonated Peptides Containing Post-Translational Modifications

by

Hayley Jo Andreazza, B.Sc (Hons)

A thesis submitted for the Degree of
Doctor of Philosophy

in the
School of Chemistry and Physics
The University of Adelaide

THE UNIVERSITY
of ADELAIDE

July 2014

For my darling son Rainer Vincent Andreazza Stranz

21/09/2009 –

I love you.

“Once you make a decision, the universe conspires to make it happen.” ― Ralph Waldo Emerson

&

In loving memory of Rolf Walch

26/05/1941 – 29/03/2013

TABLE OF CONTENTS

Abstract

The identification and characterisation of post-translational modifications (PTMs) of proteins and peptides is vital to understanding their functional properties and complex biological problems. The work presented in Chapters 2-4 of this thesis describes the development and application of a joint negative ion electrospray ionisation tandem mass spectrometry (ESI-MS/MS) and theoretical study into the identification and characterisation of several PTMs.

Chapter 2 deals with the characteristic negative ion fragmentations of deprotonated peptides containing mono- and di-phosphorylated systems. The characteristic fragmentations of monophosphorylated peptides containing pSer and pThr are the loss of H_3PO_4 from the $(M-H)^-$ anions and the formation of $H_2PO_4^-$ (previously identified by Lehmann, *et al.*). These characteristic cleavages were found to be more energetically favourable than the negative ion backbone cleavages of peptides described previously. The characteristic loss of HPO_3 from pTyr-containing peptides was found to be comparable with those already reported for negative ion backbone cleavages. The $(M-H)^-$ anions from selected diphosphopeptides show characteristic peaks corresponding to *m/z* 177 ($H_3P_2O_7^-$), 159 ($HP_2O_6^-$) and sometimes $[(M-H)^- - H_4P_2O_7]^-$.

The characteristic fragmentations of a pTyr group in the negative ion electrospray tandem mass spectrum of the $(M-H)^-$ parent anion of a peptide or protein involve the formation of PO_3^- (*m/z* 79) and the corresponding $[(M-H)^- - HPO_3]^-$ species. In some tetrapeptides where pTyr is the third residue, these characteristic anion fragmentations are accompanied by peaks corresponding to $H_2PO_4^-$ and $[(M-H)^- - H_3PO_4]^-$ (these are fragmentations normally indicating the presence of pSer or pThr). These fragment ions are formed by rearrangement processes which involve initial nucleophilic attack of a C-terminal $-CO_2^-$ [or $-C(=NH)O^-$] group at the phosphorus of the Tyr side chain [an $S_N2(P)$ reaction].

Chapter 3 describes how the negative ion ESI-MS of the peptides produced by tryptic and chymotrypsin digests of bovine insulin, and from the tryptic digest of lysozyme identify at least 80% of the sequences of these proteins as well as the positions of disulfide moieties.

Chapter 4 reports on the experimental and theoretical investigation into the negative ion fragmentations of Asp and *iso*Asp. It was found that it is not possible to differentiate between Asp and *iso*Asp containing peptides (used in this study) using negative ion ESI-MS because the *iso*Asp residue cleaves to give the same fragment anions as those formed by **δ** and **γ** backbone cleavage of Asp. No diagnostic cleavage cations were observed in the electrospray mass spectra of the MH^+ of the Asp and *iso*Asp containing peptides (used in this study) to allow differentiation between these two amino acid residues.

Chapter 5 describes the mass spectrometric component of a study showing that some selected His-containing anuran peptides, namely caerin 1.8, caerin 1.2, Ala_{15} maculatin 1.1, fallaxidin 4.1, riparin 5.1 and signiferin 2.1, all form $MMet^{2+}$ and $(M+Met^{2+}-2H^+)^{2+}$ cluster ions (where Met is Cu, Mg and Zn) following ESI. Peaks due to Cu(II) complexes are always the most abundant relative to other metal complexes. Information concerning $metal^{2+}$ connectivity in a complex was obtained using **b** and **y** fragmentation data from CID ESI-MS/MS.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

_____

Hayley Jo Andreazza

18th of July 2014

Acknowledgments

First and foremost, I wish to sincerely thank my principle supervisor, Prof. John H. Bowie for his continuous support, assistance and patience, throughout my studies. I am extremely appreciative of the opportunities and challenges that this research project has presented me with, under his knowledgeable guidance. Professor Bowie, it is a privilege to have worked with you. Thank you.

I would like to thank the past and present members of the Bowie research group, who have helped immensely over the years, providing me with support and friendship. In particular, many thanks to Dr. Tara Pukala, Dr. Anton Calabrese, Dr. Daniel Bilusich, Dr. Mark Fitzgerald, Dr Tianfang (Gavin) Wang, Dr. Patrick Sherman, Dr. Micheal (Mike) Maclean and Dr. Peter Eichinger for many helpful discussions, proofreading, teaching me laboratory techniques and providing advice on mass spectrometry and computational chemistry. Thank you also to Dr Yanqin Liu and Nha Tran for your support and friendship.

I would like to express my gratitude to my co-supervisor, Dr. Peter Hoffmann and members of his research group at the Adelaide Proteomics Centre (APC), in particular Dr. Mark Condina and Dr. Christopher Bagley, for teaching me many useful laboratory techniques and mass spectrometry methods associated with working in a proteomics laboratory.

Much appreciation goes to all the academic, teaching, technical and administrative staff whose presence make the School of Chemistry and Physics at The University of Adelaide and enjoyable and productive environment to work. In particular, my gratitude goes to Phil Clements for his help with mass spectrometry, and Dr. Danielle Williams for proofreading.

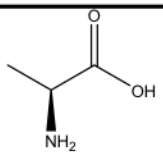
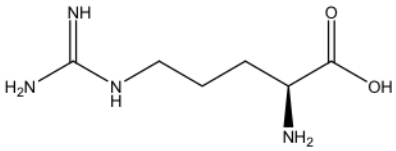
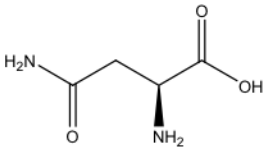
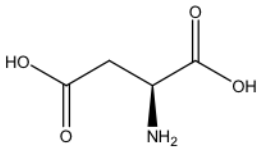
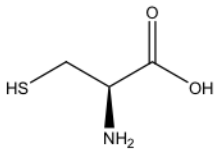
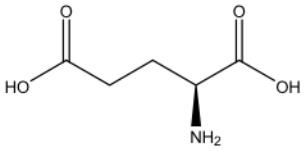
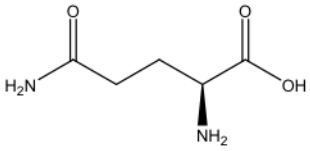
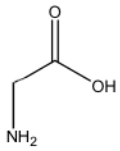
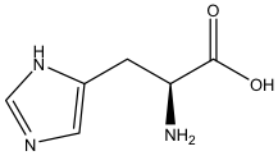
I sincerely thank all my friends both in and out of the Chemistry Department. To Sara, Ciarne, Kelly, Sandy, Jonathan, Danielle, Melanie and Mark, thank you for your unwavering friendship, love and support through the testing times and the fun times. Thank you also to Courtney, Scott, Tom, Ondrej, Alex, Paddy and Claire for the sharing

the PhD journey with me. Thank you to Jennifer for her help with formatting. A big thank you to Oscar and Carmelina for their support over the years.

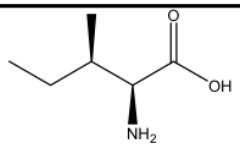
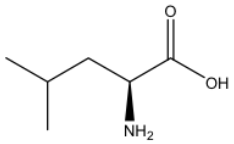
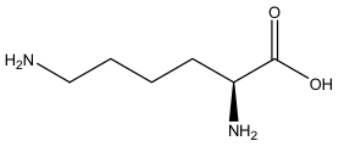
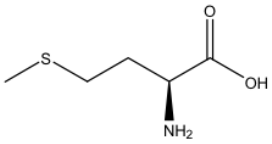
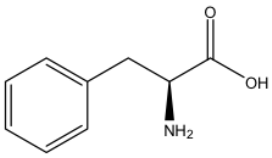
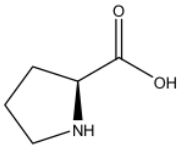
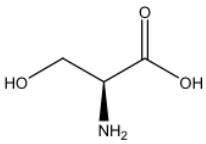
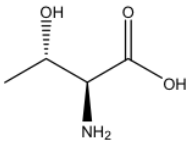
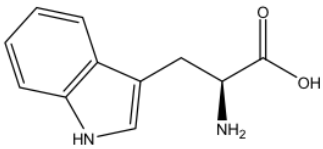
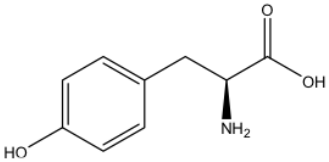
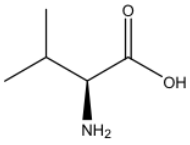
Finally, my love and gratitude goes to my family for the amazing support they have provided throughout my life. To Mum, Dad, Rolf, Dieter and Annie, your love, encouragement, advice and assistance have allowed me to reach my goals and live my dreams. I couldn't have done this without you. To Rainer, my darling son, my little ray of sunshine. Thank you for bringing me so much joy and love. I am so very proud of you! I love you big sky!

Twenty Common Amino Acids

The structure and nominal masses of the 20 common amino acids.

Amino Acid	Structure	Nominal Mass
Alanine Ala A		71
Arginine Arg R		156
Asparagine Asn N		114
Aspartic Acid Asp D		115
Cysteine Cys C		103
Glutamic Acid Glu E		129
Glutamine Gln Q		128
Glycine Gly G		57
Histidine His H		137

Twenty Common Amino Acids

Amino Acid	Structure	Nominal Mass
Isoleucine Ile I		113
Leucine Leu L		113
Lysine Lys K		128
Methionine Met M		131
Phenylalanine Phe F		147
Proline Pro P		97
Serine Ser S		87
Threonine Thr T		101
Tryptophan Trp W		186
Tyrosine Tyr Y		163
Valine Val V		99