WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

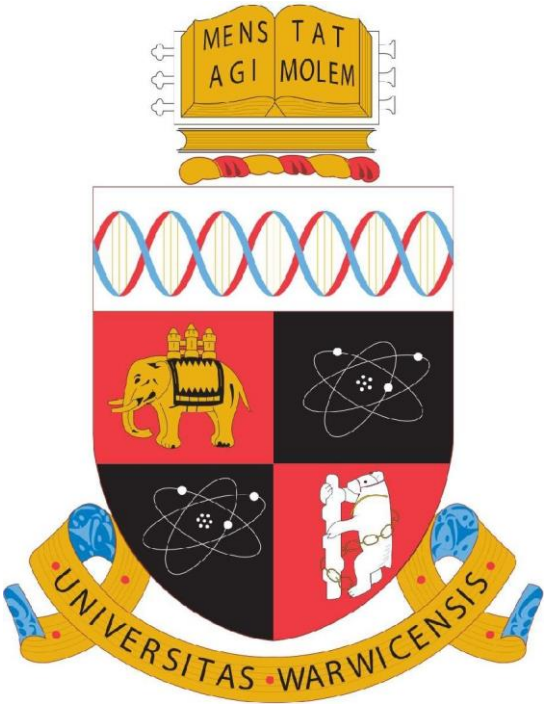
Permanent WRAP URL:

http://wrap.warwick.ac.uk/169840

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications



Advanced Applications of Mass Spectrometry for Isomer Differentiation and Analysis of Biomolecules

Anisha Haris

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry

University of Warwick

January 2022

Table of Contents

Acknowledgements

First and foremost, I would like to thank my PhD supervisor, Professor Peter B. O'Connor, for giving me the incredible opportunity to undertake this PhD as part of the FT-ICR MS group at the University of Warwick. My sincere thanks to Prof. O'Connor for always sharing his wealth of knowledge on everything related to mass spectrometry. Your advice, guidance, and continued support throughout my PhD has been invaluable. You have profoundly shaped the way in which I approach all aspects of research in this field, as you have taught us to critically analyse data so that we can consistently deliver results of the highest quality. I am most grateful for your kindness, encouragement, and the numerous opportunities that you have provided so that I could gain confidence through presenting my work to a wider audience. It is an honour to be your student and I thank you for reminding me that the pursuit of science is truly a collective human endeavour.

I would like to acknowledge all group members, past and present, in both the O'Connor group and the Barrow group. Thank you to Dr. Mark Barrow for your support, lending your expertise on FT-ICR MS, and for scheduling instrument time. Special thanks to Dr. Yuko P. Y. Lam, who has taught me so much. It has been a privilege to work with you and learn from you. Thank you for all your help and guidance in tackling the different research projects. My sincere thanks to Dr. Christopher Wootton, who generously gave his time to teach me and others about FT-ICR MS and for providing me with valuable advice and feedback on my work. Thank you to Dr. Tomos Morgan for supervising me on the instrument when I first started my PhD. You were always approachable and willing to teach or supervise when needed, for which I am grateful. I have also thoroughly enjoyed working with Dr. Alina Theisen on everything related to lasers from laser beam alignment to optimising UVPD experiments on the instrument. Thank you for your patience and for sharing your valuable knowledge and skills on this technique. Thank you to Dr. Diana Catalina Palacio Lozano, Dr. Cookson Chiu, Dr. Meng Li, Dr. Mary Thomas, and Hugh Jones, for your help and support. I also thank Bryan Marzullo and Johanna Paris, who I started the PhD journey with, thank you for the camaraderie and motivation.

I extend my thanks and appreciation to all the co-authors and collaborators for their valuable contributions to the projects presented in this thesis.

I would also like to take this opportunity to thank my parents, my siblings, my entire family, and friends, for their support and advice. I am grateful to my dear friends and housemates, Zainab Rehman and Natasha Sanjrani, who always encouraged me and were there for me throughout my PhD. Without the continued support and patience of my family and friends, this journey would not have been possible.

Finally, I would like to dedicate this thesis to the memory of my kind-hearted maternal grandparents in Bangladesh, Bedena Begum and Okil Uddin, as well as my caring paternal grandmother, Husnara Begum. To be in their presence, was to feel warmth and know kindness. Their life, and more recently, the loss of my grandparents motivated the completion of this thesis.

Declaration

I hereby declare that except where specifically references/stated are made to other sources, the thesis entitled “Advanced Applications of Mass Spectrometry for Isomer Differentiation and Analysis of Biomolecules” is the original work of the named Author. It has been composed by myself and co-authors where stated and has not been submitted in whole or in part for any other degree, diploma, or qualification.

Anisha Haris

January 2022

Abstract

The work presented herein focuses on the implementation of advanced fragmentation techniques with Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to distinguish between isomeric species, including small metabolites and peptides. Applications of matrix-assisted laser desorption ionisation- time of flight mass spectrometry (MALDI-TOF MS) and FT-ICR MS for the detection of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins are also investigated in this thesis.

The differentiation and relative quantification of isomeric species is of importance as the subtle changes in their physical structures may significantly impact their biological function. Current studies have demonstrated the potential of applying tandem mass spectrometry (MS/MS) techniques for direct isomer characterisation via generation of diagnostic fragment ions. Thus, the application of MS/MS methods has been explored in this thesis to characterize and relatively quantify the isomeric products of deamidation (chapter 2), modified tau and pi N-methylated actin peptides (chapter 3), and dihydroxylated vitamin D_3 isomers (chapter 4).

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus, responsible for causing coronavirus disease 2019 (COVID-19). In this work, MALDI-TOF MS was primarily used for the optimisation experiments to detect SARS-CoV-2 biomarker proteins including the nucleocapsid (N-protein) and the spike glycoprotein (S-protein). Focus was placed on various viral protein enrichment and extraction methods, which were applied to the standard SARS-CoV-2 proteins and then to COVID-19 negative and positive patient swab samples.

The final chapter of this thesis provides a conclusion on all the results presented herein and provides an outlook for future research. This can be used to further develop the current experimental work on the use of MS/MS techniques for the differentiation and relative quantification of various isomeric compounds as well the improvement of viral protein enrichment methods for MALDI-TOF MS analysis of SARS-CoV-2 biomarker proteins in complex patient samples.

Author Information and Curriculum Vitae

Anisha Haris

Email: a.haris@warwick.ac.uk

Education

Doctor of Philosophy, Chemistry

University of Warwick October 2017 – January 2022

Thesis title: Advanced Applications of Mass Spectrometry for Isomer Differentiation and Analysis of Biomolecules

Supervisor: Professor Peter B. O'Connor

Master of Chemistry with Industrial Experience (MChem)

Including Degree of Bachelor of Science in Chemistry (BSc)

University of Bradford September 2013 – July 2017

Thesis title: Application of Differential Ion Mobility Spectrometry – Mass Spectrometry to the Analysis of Isomeric Lipids

Supervisor: Dr. Richard Telford

MChem award: graduated with distinction

BSc Chemistry award: graduated with first class honours

Publications

Accepted papers

- Alexandre A. Shvartsburg, **Anisha Haris**, Roch Andrzejewski, Andrew Entwistle, and Roger Giles. Differential Ion Mobility Separations in the Low-Pressure Regime. Analytical Chemistry 2018 90 (1), 936-943, DOI: 10.1021/acs.analchem.
- Bryan P. Marzullo, Tomos E. Morgan, Alina Theisen, **Anisha Haris**, Christopher A. Wootton, Simon J. Perry, Mansoor Saeed, Mark P. Barrow, and Peter B. O'Connor. Combining Ultraviolet Photodissociation and Two-Dimensional Mass Spectrometry: A Contemporary Approach for Characterizing Singly Charged Agrochemicals. Analytical Chemistry 2021 93 (27), 9462-9470, DOI: 10.1021/acs.analchem.1c01185
- "Differentiation of Dihydroxylated Vitamin D3 Isomers using Tandem Mass Spectrometry" by **Anisha Haris**, Yuko P. Y. Lam, Christopher A. Wootton, Alina Theisen, Bryan P. Marzullo, Pascal Schorr, Dietrich A. Volmer, and Peter B. O'Connor. Accepted by Journal of American Society of Mass Spectrometry (April 2022).

Manuscripts in preparation

- "Differentiation and Relative Quantification of the Isomeric Products of Deamidation using ECD and UVPD Tandem Mass Spectrometry" by **Anisha Haris**, Yuko P. Y. Lam, Alina Theisen, Christopher A. Wootton, Tomos E. Morgan, Mark P. Barrow, and Peter B. O'Connor
- "Distinguishing between methylated histidine isomers generated as a post-translational modification of actin" by **Anisha Haris**, Yuko P. Y. Lam, Christopher A. Wootton, Hamdi Hussain, Mohan Balasubramanian, and Peter B. O'Connor.
- "Enhancing top-down and 2DMS experiments by implementation of (activated ion-) 193 nm UVPD on a FT-ICR mass spectrometer" by Alina Theisen, Christopher A. Wootton, **Anisha Haris**, Tomos E. Morgan, Yuko P. Y. Lam, Mark P. Barrow, and Peter B. O'Connor

- "Multimodal tandem mass spectrometry techniques for the analysis of phosphopeptides" by Johanna Paris, Alina Theisen, **Anisha Haris**, Bryan P. Marzullo, Tomos E. Morgan, Andrew Kerr, Alina Theisen, Sean Ellacott, Anisha Haris, Christopher A. Wootton, Mark P. Barrow, John O'Hara, and Peter B. O'Connor
- "UVPD of biologically compatible copolymers" by Tomos E. Morgan, Andrew Kerr, Alina Theisen, Sean Ellacott, **Anisha Haris**, Christopher A. Wootton, Mark P. Barrow, Anthony W. T. Bristow, Sébastien Perrier, and Peter B. O'Connor

Awards

- July 2017 – Awarded the JEOL UK ltd. prize for best overall performance in Stage 4 of an MChem course
- September 2018 - BMSS John Beynon Travel Award for British Mass Spectrometry Society Conference in Cambridge, UK
- October 2019 - Awarded for the best oral presentation at the East Midlands Proteomics Workshop in Sheffield, UK
- March 2020 - BMSS John Beynon Travel Award for B for DGMS Conference in Münster, Germany

Conference Presentations

Oral Presentations

- August 2019 - 4th Short Course EU FT-ICR MS Network. University of Warwick, UK
- October 2019 - East Midlands Proteomics Workshop (EMPW). Sheffield, UK
- September 2021 - British Mass Spectrometry Society (BMSS). Sheffield, UK

Poster Presentations

- July 2018 - Uppsala Conference on Electron Capture and Transfer Dissociation Mass Spectrometry (Uppcon) in Leeds, UK
- August 2018 - EU FT-ICR MS Network Workshop in Joensuu, Finland
- September 2018 - BMSS in Cambridge, UK
- October 2018 - East Midlands Proteomics Workshop (EMPW). Lincoln, UK
- March 2019 - Celebration of Native Mass Spectrometry in Oxford, UK
- September 2019 - BMSS in Manchester, UK
- March 2020 - German Society for Mass Spectrometry (DGMS) in Münster, Germany
- June 2020 – American Society for Mass Spectrometry (ASMS) Online Reboot Program

List of Figures, Tables, and Schemes

Chapter 1: Introduction

Chapter 2: Differentiation and Relative Quantification of the Isomeric Products of Deamidation using ECD and UVPD MS/MS

Chapter 2: Supplementary Information

Chapter 4: Exploring tandem mass spectrometry methods for the analysis of dihydroxylated vitamin D_3 isomers

Chapter 4: Supplementary information

Chapter 5: Development of matrix-assisted laser desorption ionisation mass spectrometry for the detection of SARS-COV-2 proteins

Chapter 5: Supplementary information

Chapter 6: Conclusions and Future Work

Abbreviations

Aβ	Amyloid beta
ABC	Ammonium bicarbonate
AC	Alternating current
ACN	Acetonitrile
AD	Alzheimer's disease
APCI	Atmospheric pressure chemical ionisation
APPI	Atmospheric pressure photo-ionisation
Aβ1-40	Amyloid beta 1-40
BSA	Bovine serum albumin
CAD	Collisionally activated dissociation
CCS	Collision-cross section
CEM	Chain ejection model
CHCA	α-Cyano-4-hydroxycinnamic acid
CHEF	Correlated harmonic excitation field
CI	Chemical ionisation
CID	Collisionally induced dissociation aka CAD
CRM	Charge residue model
CRS	Charge reduced species
Da	Daltons
DC	Direct current
DHB	2,5-Dihydroxybenzoic acid
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid

DTT	Dithiothreitol
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EI	Electron ionisation
EID	Electron induced dissociation
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
ExD	Electron based dissociation
FA	Formic acid
FDR	False discovery rate
FFT	Fast-Fourier transform
FID	Free induction decay
FT	Fourier transform
FT-ICR	Fourier transform-ion cyclotron resonance
FTMS	Fourier transform mass spectrometry
FWHM	Full width half maximum
GC	Gas chromatography
GC-MS	Gas chromatography mass spectrometry
HDX	Hydrogen-deuterium exchange
hECD	hot-electron capture dissociation
hIAPP	Human islet amyloid polypeptide
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High performance liquid chromatography
HSA	Human serum albumin
i.d.	Internal diameter
IAA	Iodoacetamide
ICR	Ion cyclotron resonance

IEM	Ion ejection model
IM-MS	Ion mobility-mass spectrometry
IR-ECD	Infrared electron capture dissociation
IRMPD	Infrared multiple photon dissociation
kV	Kilovolts
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
M	Mega-word
MALDI	Matrix assisted laser desorption ionisation
MeOH	Methanol
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multistage tandem mass spectrometry
MW	Molecular weighted average
m/z	Mass-to-charge ratio
nESI	Nano electrospray ionisation
nLC	nano-LC
NMR	Nuclear magnetic resonance
PD	Parkinson's disease
PEG	Polyethylene glycol
ppm	Parts-per-million
PTM	Post translation modification
QIT	Quadrupole ion trap
RF	Radio frequency
RNA	Ribose nucleic acid
RP	Reversed phase
R.P.	Resolving power

S/N	Signal-to-noise
SA	3,5-Dimethoxy-4-hydroxycinnamic acid
SAX	Strong anion exchange chromatography
SCL	Side-chain loss
SCX	Strong cation exchange chromatography
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodcyl sulfate polyacrylamide
SORI	Sustained off-resonance irradiation
SPE	Solid phase extraction
SSNMR	Solid-state nuclear magnetic resonance
SWIFT	Stored waveform inverse Fourier transform
T	Tesla
TD	Acquisition time/time duration of transient
TEM	Transmission electron microscopy
TOF	Time-of-flight
UVPD	Ultraviolet photodissociation
UV/vis	Ultra-violet/visible spectrometry
UW-	Utah-Washington-
v/v	volume per volume
μM	Micro-molar

1. Introduction

1.1. Mass Spectrometry

Mass spectrometry (MS) is a well-established analytical technique used for the qualitative and quantitative analysis of a wide range of molecules. MS focuses on the separation of the masses of molecules based on measuring the mass-to-charge (*m/z*) ratio of the ions generated because to detect the species of interest, they must first be charged. The earliest reported mass spectrometer was successfully constructed by British Physicist Sir J.J Thomson in 1912. He used the instrument, originally called a parabola spectrograph, to detect non-radioactive isotopes.[1] Francis Aston, who was a student of Thomson's, continued to develop on the original instrument and he built a mass spectrometer with improvements made to the speed, mass accuracy, and resolving power.[2] The studies taking place at the beginning of the 20th century were essential to our understanding of the fundamentals of MS and provide a basis for the modern mass spectrometers that we use today.

A mass spectrometer generally consists of three major components including an ionisation source, a mass analyser, and a detector. An ionisation source is the region where neutral molecules are converted to charged ions prior to entrance into the MS. The mass analyser takes the ionised species and separates them using electric and/or magnetic fields, based on their *m/z* ratio, and then outputs them to the detector which records a mass spectrum displaying the ions that correspond to each *m/z* value. The main ionisation methods, types of mass analysers and fragmentation methods studied will be discussed in this section.

1.1.1 Ionisation Methods

Ionisation is a process by which a neutral molecule is converted to an electrically charged molecule/ion. As mentioned above, the species of interest need to be charged and in the gaseous state for MS detection and samples are normally ionised in the source region at the front end of the instrument. Ionisation methods can be classed as hard or soft ionisation. Hard ionisation methods produce ions with high internal energies, which fragment before leaving the ion source producing complex spectra; therefore, these methods are best for the structural elucidation of small molecules. Soft ionisation on the other hand minimises further fragmentation and is most suitable for characterising mixtures or larger, more complex molecules.

1.1.1.1 Electron Ionisation

Electron ionisation (EI) was first described by A.J. Dempster in 1918.[3] It is considered to be one of the earliest, hard ionisation techniques developed for MS. In EI, a beam of high energy negatively charged electrons (accelerated by a potential of approximately 70 eV) are generated by a heated filament and attracted towards a positively charged anode in an ionisation chamber. Vaporised/gaseous molecules are introduced from the inlet which is placed orthogonally to the heated filament and collide with the high energy electron beam to displace a bound electron from the analyte species, forming a radical cation. The overall EI process can be represented by the chemical equation below:

$$M + e^- \rightarrow M^{+\bullet} + 2e^- \qquad \text{[Eqn. 1]}$$

EI produces unstable radical ions which can rapidly dissociate or rearrange with low internal barriers to generate significant fragmentation; therefore, it is classed as a hard ionisation method. The fragment ions produced are useful for the structural elucidation of volatile small molecules. EI-MS is commonly coupled with gas chromatography (GC) for the analysis of molecules prevalent and important in the industries of food, clinical analysis, and the analysis of environmental samples. However, the use of EI mass spectra for distinguishing between isomeric compounds can be difficult due to the generation of radicals, which cause multiple rearrangements in the molecule and result in the subsequent loss of isomeric information.

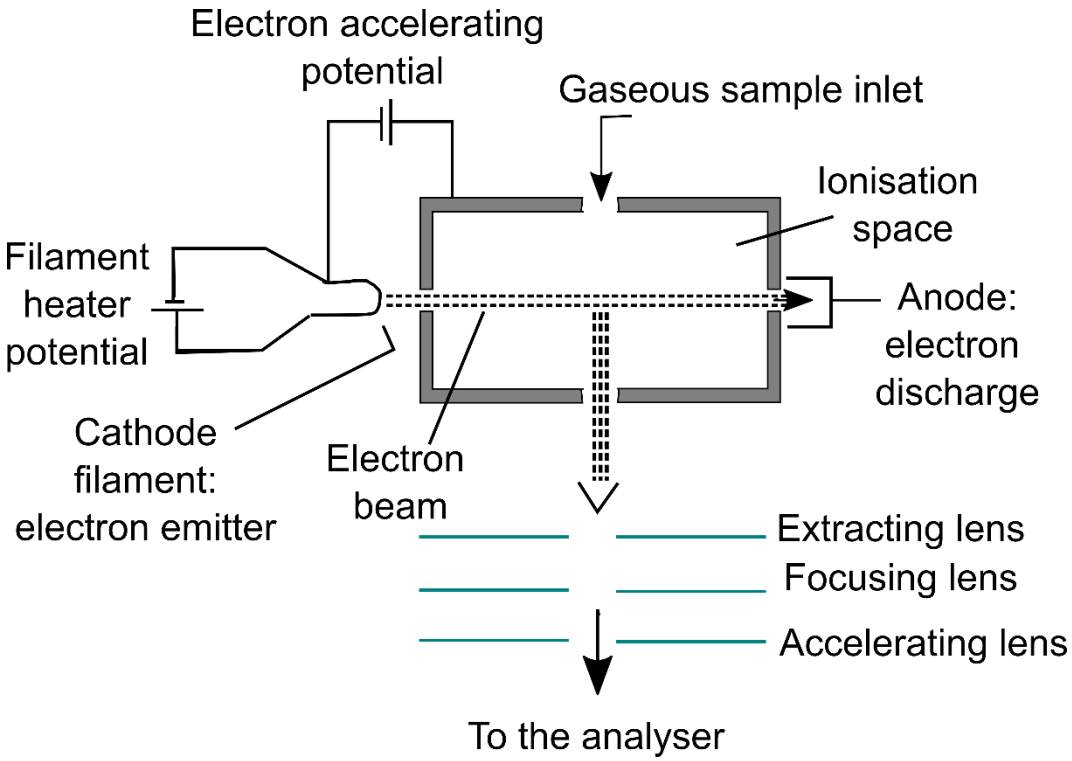


Figure 1. 1 Schematic representation of an electron ionisation (EI) source (adapted from De Hoffmann *et. al.* 2007)[4]

1.1.1.2 Chemical Ionisation

Chemical ionisation (CI) is a soft ionisation method first introduced by the Russian scientist Victor Talrose in the early 1950's.[5] Following on from the work of Talrose, significant contributions to the development of CI were made by American scientists, Burnaby Munson and Frank Field in 1966.[6,7] In CI, the sample is introduced into a chemical ionisation chamber, which is filled with excess reagent gas such as methane. Reagent gas molecules are ionised by EI and then the protons are transferred to the analyte molecule. The CI process is shown in chemical equations 2-3 below:

Step 1: EI of the reagent gas – methane (CH_4)

$$CH_4 + e^- \rightarrow {CH_4}^{+\bullet} + 2e^- \quad \text{[Eqn. 2]}$$

$${CH_4}^{+\bullet} + CH_4 \rightarrow {CH_5}^+ + {CH_3}^\bullet \quad \text{[Eqn. 3]}$$

Step 2: Proton transfer to the analyte molecule

$$M + {CH_5}^+ \rightarrow MH^+ + CH_4 \quad \text{[Eqn. 4]}$$

Unlike EI, CI is a softer ionisation method and does not involve the formation of radicals on the analyte molecule, M, so fragmentation of the sample is generally much lower than that of EI. This results in little to no fragmentation observed in CI spectra and the molecular ion (MH^+) is more readily identified due to the reduced complexity. CI can also be operated in negative mode (to generate anions) by using different reagent gases. As this method generally produces singly charged molecular ion species and as the precursor molecules must already be in the gas phase, the ionisation method is limited to smaller molecules and is not readily amenable to analysis of biomolecules such as peptides and proteins.

1.1.1.3 Electrospray Ionisation

Electrospray ionisation (ESI) was first developed by John Fenn at Yale University in 1984, who went on to receive the Nobel Prize in Chemistry for the invention in 2002.[8] ESI is a widely recognised and implemented soft-ionisation method in MS, particularly in the field of proteomics as this ionisation method is commonly applied to the analysis of peptides, proteins, and large biological macromolecules.[9]

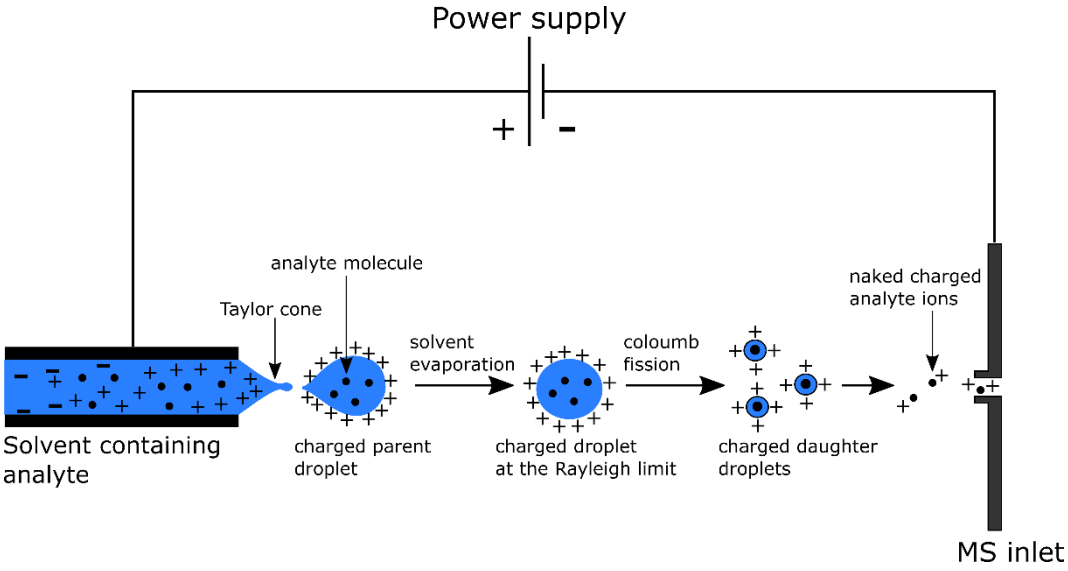


Figure 1. 2 Schematic representation of the ESI process (adapted from Banerjee *et. al.* 2012).[10]

One of the first steps of ESI involves application of a high potential difference (approximately 2-6 kilovolts (kV)) between the flow of the analyte solution and the inlet of the mass spectrometer.[10-12] The charged solution passes through an ESI needle and a nebulising gas such as nitrogen is applied, which increases sample flow rate and aids the dispersion of the solution into a fine spray. The charges within the droplet migrate to the surface and the accumulation of charge distorts the droplet into a conical shape known as the Taylor cone.[13-15] As the charged droplets move under the influence of the strong electric field, solvent evaporation occurs and the density of charges on the surface of the droplet increases until it reaches a critical value known as the Rayleigh limit.[16,17] At this point due to the instability, Coulombic repulsion/explosion occurs to overcome the surface tension, producing multiple smaller droplets, which travel along the gradient of the electric field and ultimately form individual desolvated ions, which are analysed by the mass spectrometer.

The specific mechanism of forming ions from the charged droplets is still under consideration. However, the following proposed models are generally accepted and recognised; the ion ejection model (IEM), the charged residue model (CRM), and the chain ejection model (CEM).[18,19]

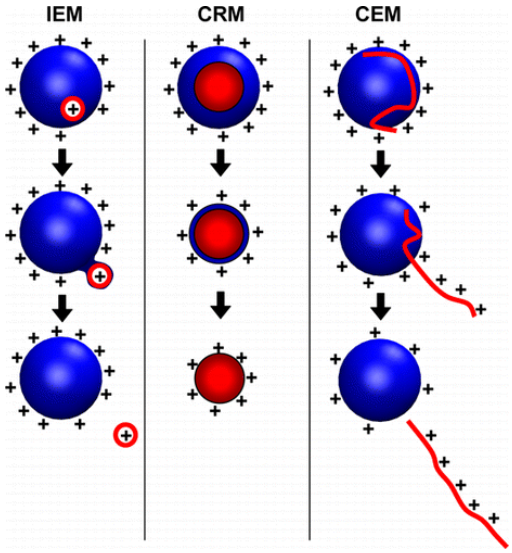


Figure 1. 3 Summary of the ESI models for ion formation. From left to right, the Ion Ejection Model, the Charge Residue Model, and the Chain Ejection Model. (Reproduced from Konermann *et. al.* 2013).[19]

Konermann *et. al.*[19] suggested that low molecular weight molecules follow the IEM model, usually getting charged via proton/salt ions already present in the solution. CRM is mainly applicable to large globular species such as folded proteins and protein complexes which require a lot of energy to overcome the energy barrier for IEM to occur. In CRM model, the protein ions are released by droplet evaporation to dryness. The CEM model was proposed for unfolded proteins (disordered/extended chain structure), such as denatured biomolecules and polymers. In the CEM model, the protein ions are ejected from the protein surface.

ESI is one of the softest ionisation methods, enabling ionisation of large molecules that are characterised by non-covalent interactions. Another major advantage of ESI is that the ions generated are multiply charged (depending on their molecular mass), and the analyte remains intact (with no fragmentation) forming molecular ions $[M+nH]^{n+}$ in positive ionisation mode or $[M-nH]^{n-}$ in negative ionisation mode. This improves detector sensitivity as high molecular weight molecules can be analysed at lower mass limits and aids accurate qualitative and relative quantitative measurements.

In addition to increased sensitivity and preservation of large analytes via generation of multiply charged molecular ions, combining ESI with liquid chromatography (LC) enables separation of complex mixtures prior to MS analysis. Some disadvantages include susceptibility to contamination when high concentrations of salt

and other buffers are used, requiring the use of chromatography or an offline desalting step prior to ESI-MS analysis.

1.1.1.4 Nano Electrospray Ionisation

Nano electrospray ionisation (nESI) is a form of ESI as the fundamental ionisation processes are the same but the experimental setup for nESI is different. A smaller diameter orifice emitter (~ 0.5-5 µm) is utilised in nESI compared to normal ESI (~10-100 µm). As a result, this causes smaller droplets to form during the electrospray process, better desolvation of the charged droplets, and more efficient conversion of the analyte solution to ions leading to overall better sensitivity compared to ESI.[20,21]

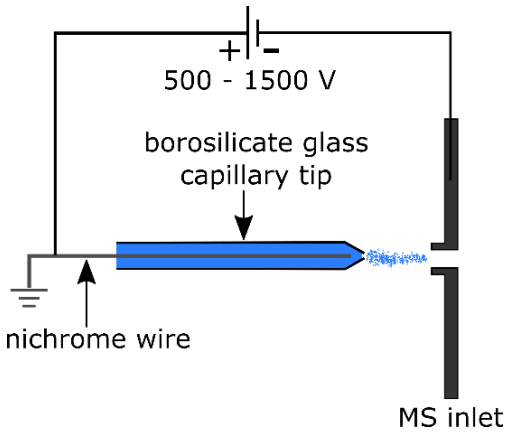


Figure 1. 4 Schematic diagram of a nanoelectrospray ionisation setup.

In nESI, the solution of sample is drawn out from the glass capillary emitter tip to the MS inlet via a voltage difference applied between the two via the metal wire. Along with better sensitivity, the advantages of nESI also include lower sample consumption, lower flow rate (nL/min to few tens of nL/min), application of lower voltages (0.5-1.5 kV), and nESI is also more tolerant towards different buffer compositions.

1.1.1.5 Matrix-Assisted Laser Desorption Ionisation

Matrix assisted laser desorption ionisation (MALDI) is a soft ionisation method that uses laser irradiation on analyte molecules embedded in a matrix to generate gas-phase ions without causing fragmentation to the species of interest. Michael Karas and Franz Hillenkamp first introduced and developed MALDI in 1985.[22-24] In the initial studies, they found that alanine could be ionised more easily when mixed with tryptophan and irradiated with a pulsed laser at 266 nm.[23] Their research led to the works of Tanaka *et. al.* in 1988, who were able to demonstrate the potential of MALDI analysis for large

biomolecules and polymer species with high molecular weights, exceeding 100,000 Daltons (Da).[25] MALDI has since become a powerful laser ionisation method with applications to a broad range of molecules including peptides, proteins, oligonucleotides, carbohydrates, synthetic polymers, lipids, and other organic or labile macromolecules.

The two main steps to achieve MALDI are desorption and ionisation. The first step involves desorption, which can be described as the transfer of individual molecules from the condensed phase to the gas phase from the outermost layers of the sample.[26] In this step, the analyte of interest is mixed properly with the matrix solution. Matrices are generally small organic molecules, which need to have strong absorption at the laser wavelength. They are therefore composed of aromatic rings and tend to have a functional group for proton donation to the analyte molecule (examples of common matrices used for MALDI are shown in Figure 1.5).

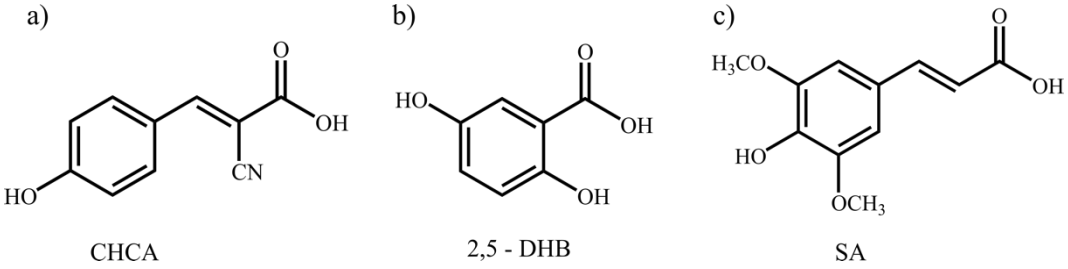


Figure 1. 5 Examples of three common matrices used in MALDI-MS for peptide and protein analysis a) α-cyano-4-hydroxycinnamic acid (CHCA), b) 2,5-dihydroxybenzoic acid (DHB), and c) 3,5-dimethoxy-4-hydroxycinnamic acid also known as sinapinic acid (SA).

After the analyte is mixed with the matrix, the mixed solution is spotted on to a stainless steel MALDI target plate and must then be allowed to dry prior to MALDI analysis. The second step occurs under vacuum conditions within the source of the mass spectrometer. This involves laser irradiation of the solid sample-matrix MALDI spot, resulting in rapid heating and ablation of the matrix crystals. The aromatic groups on the excited matrix molecules absorb a large amount of energy induced from the laser and deprotonate.[26] As the matrix molecules expand into the gas phase, the analyte molecules do the same, where the released protons transfer to the neutral analyte molecules and generate molecular ion species.

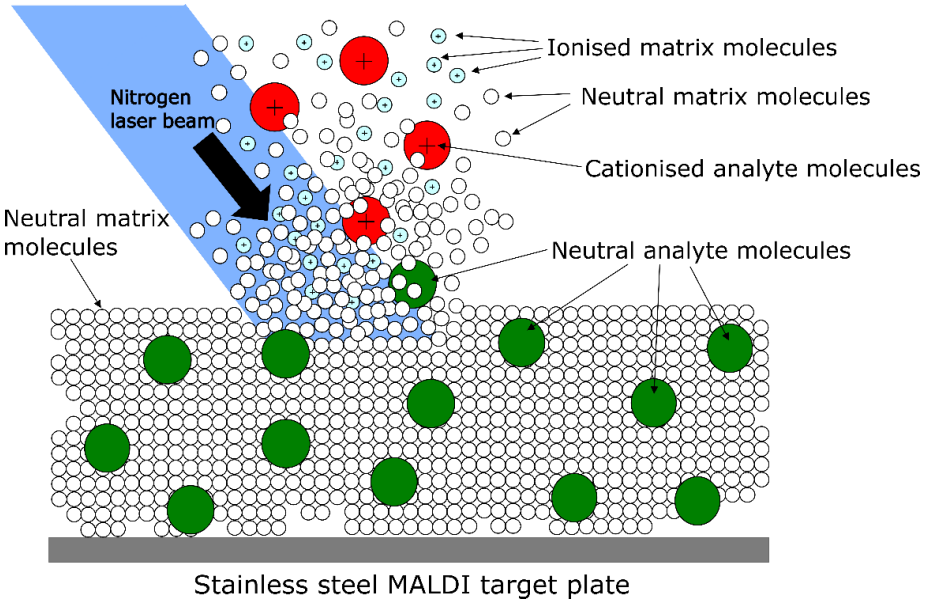


Figure 1. 6 Schematic representation of the MALDI ionisation process.

One of the main advantages of MALDI is that the process of soft ionisation enables observation of ionised molecules with minimal fragmentation of analytes as the generated ions have low internal energy. This is beneficial for the analysis of fragment proteins and other large biomolecules. However, in some cases, increasing the laser energy can still cause fragmentation of molecular ions.

Another benefit of MALDI is that the performance of this ionisation method is less affected by buffer components, detergents, and contaminants compared to other soft-ionisation methods such as ESI and nESI. On the other hand, some disadvantages include, low shot-to-shot reproducibility, which stems from the inhomogeneity of the sample and matrix distribution on the spot. This can significantly affect any quantitation results in MALDI. EI, ESI, nESI, and MALDI are the most common ionisation methods, and were the ionisation methods used in this thesis.

1.1.2. Mass Analysers

After a sample is ionised, the ions are sent through to the mass analyser, which is the second major component in the mass spectrometer after the ionisation source. The main purpose of the mass analyser is to separate the ionised species based on their *m/z* ratios using electric and/or magnetic fields. In this section, the following three mass analysers, the quadrupole, time of flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR) mass analysers, will be discussed in further detail.

The five main analytical criteria that are used to critically assess the performance of a mass analyser include the mass range limit, scan speed, ion transmission, mass accuracy, and the resolving power. The mass range determines the maximum lower and upper *m/z* limit over which the mass analyser can measure ions and is usually adjustable using the instrument parameter settings. The scan speed is the rate required to measure over a particular mass range in a mass analyser and simply put it is the rate at which we acquire mass spectra. The ion transmission can be defined as the ratio of the number of ions reaching the detector and the number of ions entering the mass analyser. It is common to observe ion losses as the ions go through the different sections in the mass analyser to the detector.

The mass accuracy is how close the measured *m/z* value is to the theoretical (exact) *m/z* value and this is normally calculated by determining the difference between theoretical (exact) and the measured *m/z* (Eqn. 5).

Parts per million (ppm) is the common unit for the expression of mass accuracy. Mass accuracy is also affected by the stability and the resolving power of a mass analyser (if other species interfere with the measurement).

$$Error\ (ppm) = \frac{measured\ m/z - theoretical\ m/z}{theoretical\ m/z} \times 10^6 \qquad \text{[Eqn. 5]}$$

Resolving power is the ability of a mass analyser to distinguish two signals from two ion packets with a small difference in their *m/z* ratios. In other words, it can be defined as how well resolved or separated the peaks are within the mass spectra. The resolving power can be calculated by m/Δ m 50 % (Figure 1.7), where m is the *m/z* value of the peak and Δ m 50% is the peak width measured at half maximum (FWHM).

A greater resolving power indicates the mass analyser has a higher ability to distinguish ions with small *m/z* differences, therefore, a high resolving power mass analyser is useful in complicated sample measurement.

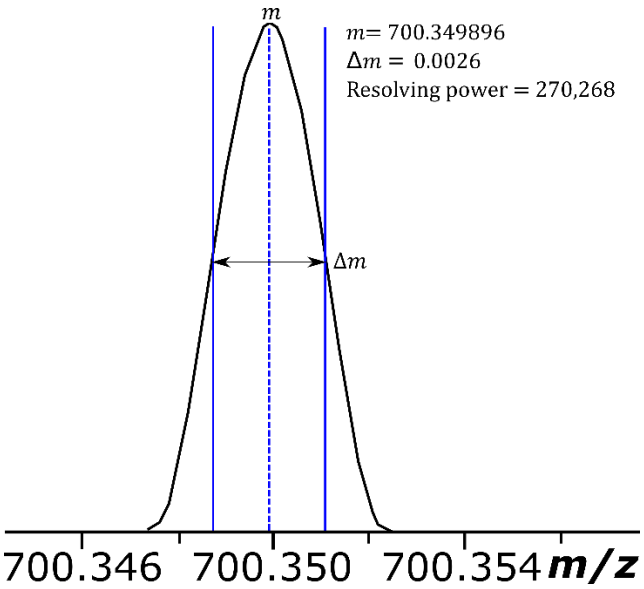


Figure 1. 7 Resolving power calculated by Full Width Half Maximum (FWHM).

1.1.2.1. Quadrupole

The quadrupole mass analyser was first described by German scientists Paul and Steinwegen in 1953.[27] This type of mass analyser separates ions according to their *m/z* ratios, based on the stability of the ion trajectories in an oscillating electric fields that is applied to the rods.

A quadrupole has four cylindrical or hyperbolic rods, set parallel to each other with each opposite pair being electronically connected (Figure 1.8). A fixed direct current (DC) and an alternating radio frequency (RF) potential is applied to the rods. The RF potential on each pair of rods is set to be completely out of phase by 180 ° and oscillate rapidly between the pairs of rods. The electric field generated by the RF voltage (V), and the DC voltage (U) on the rods is used to influence the oscillation of ions.

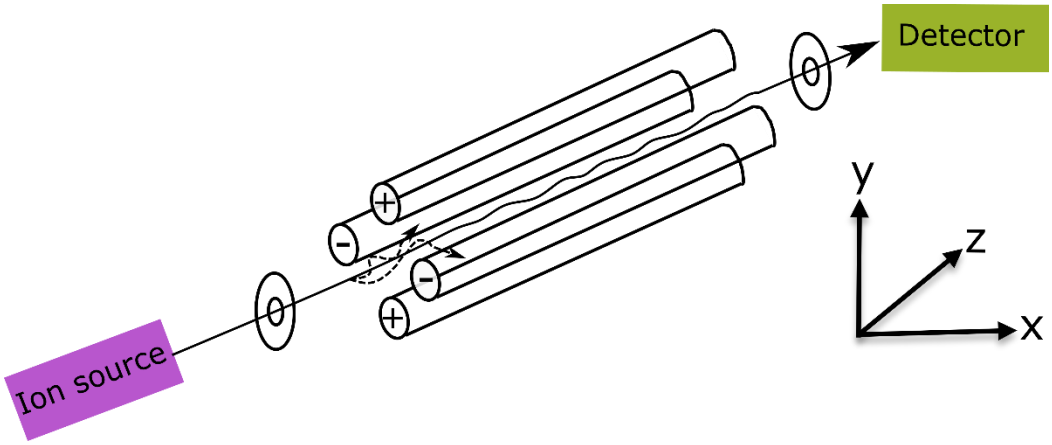


Figure 1. 8 Schematic of a quadrupole which contains four parallel rods, connecting diagonally in pairs of positive and negative terminals (adapted from El-Aneed *et. al.* 2009).[28]

Ions that are injected into the quadrupole travel along the z axis, while ions oscillating in the x and y axes and are separated based on the stability of their trajectories. The trajectory stability of ions can be represented by the following equations:[29]

$$a = \frac{8zeU}{m\omega^2 {r_0}^2} \qquad \text{[Eqn. 6]}$$

$$q = \frac{4zeV}{m\omega^2 {r_0}^2} \qquad \text{[Eqn. 7]}$$

Where ze represents charge of the ion, m represents mass of the ion, r_0 represents the field radius, and ω represents the angular frequency; and a and q are directly proportional to U and V respectively.

Only ions with stability at certain U and V values can be transmitted and detected, while all other unstable ions will be discharged against one of the rods for example and will not be detected.

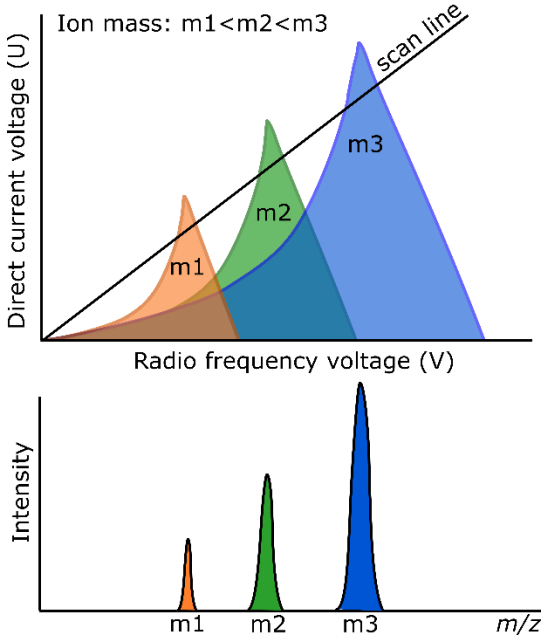


Figure 1. 9 Representation of the stability diagram of ions in a quadrupole. Ions are represented by m1, m2, and m3 with increasing the mass respectively.

For ions that are of different masses, the solution to the Mathieu equation will result in a different stability area, and the regions of stability and instability for the ions in a quadrupole can be represented via the Mathieu stability diagram (see above Figure). The scan line represents the U/V line where a constant U/V value is kept during the quadrupole operation. Since the scan line only crosses parts of the stable areas; only ions with specific *m/z* values are transmitted. By adjusting the U/V values to increase the slope of the scan line, where only the vertices of the stable regions are crossed, the peak width is reduced and the resolution in increased. In addition, by switching off the DC potential (U=0), the quadrupole will function as an RF ion guide, allowing all ions above a certain *m/z* value to pass through.

The resolving power of a quadrupole is approximately 2000 at 1000 *m/z*, with a maximum upper mass limit of approximately 4000 *m/z* and a mass accuracy of 100 ppm.[4] Quadrupoles are well known for their robustness and reliability; therefore, they have been employed in many different research fields.

1.1.2.2. Time-of-Flight

The idea behind the time-of-flight (TOF) mass analyser was first proposed and discussed by W. E. Stephens in 1946 at a meeting of the American Physical Society and the advances made were published in the Journal of Review of Scientific Instruments in 1953[30] His initial ideas were soon transformed and implemented into the first commercial linear TOF mass spectrometer (TOF-MS) in 1955 by Wiley and McLaren.[31] The TOF mass analyser separates ions according to their *m/z* ratios, based on the kinetic energy and velocity of the ions.

In a TOF mass analyser, ions are accelerated through a flight tube by an electric field, induced by the potential difference between the electrodes and the extraction grid. In theory, all ions gain the same amount of kinetic energy, however different ions have different masses, therefore travel with different velocity, and reach the detector at different times.

Equation for the kinetic energy gained by the ions:

$$KE = \frac{1}{2}mv^2 = qV \qquad \text{[Eqn. 8]}$$

The equation for the velocity of the ion (after rearrangement of eqn.8):

$$v = \sqrt{\frac{2qV}{m}} \qquad \text{[Eqn. 9]}$$

where KE is the kinetic energy obtained by the ion, m is the mass of the ion, v is the ion velocity, q is the charge of the ion, and V is the electric potential generated from the potential difference between the electrodes and the extraction grid.

After initial acceleration, the ion travels in a straight line at constant velocity to the detector. The time t needed to cover the distance L in the field free region in the time-of-flight tube before reaching the detector is given by:

$$t = \frac{L}{v} \qquad \text{[Eqn. 10]}$$

Combining the last two equations (Eqn. 9 and Eqn. 10) shows that the flight time of an ion can be calculated from a measurement of

$$t = \sqrt{\frac{mL^2}{2qv}} = \sqrt{\frac{m}{z}}\sqrt{\frac{L^2}{2eV}} \qquad \text{[Eqn. 11]}$$

where z is the number of charges on the ion and e is the charge constant.

Theoretically, all ions get the same amount of kinetic energy, however in reality, the kinetic energy gained by each ion is slightly different to each other, leading to spreading of ions drift time, even ions with the same *m/z*, causing peak broadening and low peak resolution. To overcome this, multiple solutions have been developed. The reflectron is a common technique employed to correct the difference in kinetic energy between ions. The idea of reflectron in TOF was first introduced by Mamyrin *et. al.* in 1973.[32] The reflectron acts like a mirror which reflects the ions travelling from the ionisation source to the detector.

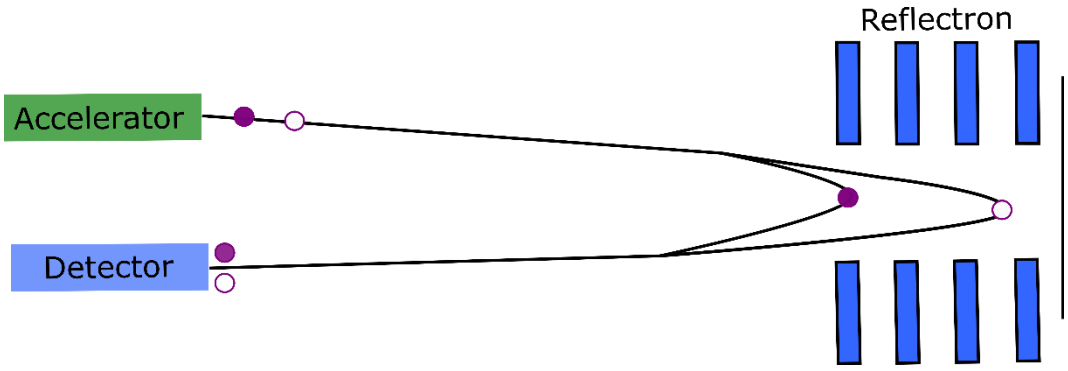


Figure 1. 10 Schematic representation of a reflectron TOF (adapted from de Hoffmann *et. al.* 2013).[4]

Ions with higher kinetic energy travel faster and penetrate deeper into the reflectron compared to ions with less kinetic energy, meaning ions with higher kinetic energy will spend a longer time in the reflectron than that of ions with lower kinetic energy. Therefore, ions with the same *m/z*, even with different initial kinetic energy after leaving the source can arrive the detector at the same time. With the application of reflectron in TOF-MS, the mass accuracy can be enhanced significantly from 200 ppm to 10 ppm (depending on the calibration) and resolving power at 1000 *m/z* can be improved greatly from 5,000 to 20,000 or higher, resolving power of 100,000 can be achieved with some commercial setups now.

As the mass resolution is proportional to flight time and the flight path, a potential solution to improve the resolution is to increase the length of the flight tube. However, too long a flight tube decreases the performance of TOF analysers because of the loss of ions by scattering after collisions with gas molecules or by angular dispersion of the ion beam. It is also possible to increase the flight time by lowering the acceleration voltage but lowering this voltage reduces the sensitivity. Therefore, the only way to have

both high resolution and high sensitivity is to use a long flight tube with a length of 1 to 2m for a higher resolution and an acceleration voltage of at least 20 kV to keep the sensitivity high.

The mass accuracy of a linear TOF mass analyser is around 200 ppm with the resolving power around 5000 at 1000 *m/z*, whereas the mass accuracy for the reflectron TOF is 10 ppm with the resolving power around 20,000 (at 1000 *m/z*).

1.1.2.3. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

The theoretical concept of ion cyclotron resonance (ICR) can be dated back to Ernest O. Lawrence in 1930, who applied the theory to a cyclotron resonator and also won the Nobel prize in 1923 for the invention of the cyclotron.[33-35] Sommer *et. al.* used Lawrence's proposed concept of ICR and implemented this into a mass spectrometer, called the omegatron in the early 1950's.[36,37] Use of inductive detection and the Fourier transform was then developed by Comisarow and Marshall to ICR-MS and the first Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer was built in 1974.[38,39] FT-ICR MS is well known for the ability to provide consistently high resolving power and high mass accuracies and therefore has been applied to a wide range of molecules in the fields of proteomics,[40-45] petroleomics,[46,47] metabolomics,[48,49] glycomics,[50] lipidomics[48] and more. In this section, the main principles of FT-ICR MS will be discussed.

Ion Cyclotron Motion

Cyclotron motion refers to the motion that ions experience in the presence of an applied magnetic field. In modern FT-ICR MS instruments, the analyte ions are produced from an ion source, accumulated in the ion optics, and then transferred to the ion cyclotron resonance (ICR) analyser cell. The ICR cell is placed in the centre of a uniform, unidirectional and homogenous magnetic field generated by a superconducting magnet. When the ions are transferred to the ICR cell, the ions experience a force called the "Lorentz force", which causes an ion to travel in a circular orbit that is perpendicular to the magnetic field. The Lorentz force can be determined by the equation below:

$$F_{Lorentz} = qv \times B \qquad \text{[Eqn. 12]}$$

where q is the charge of the ion, v is the perpendicular speed of the ion and B is the magnetic field strength.

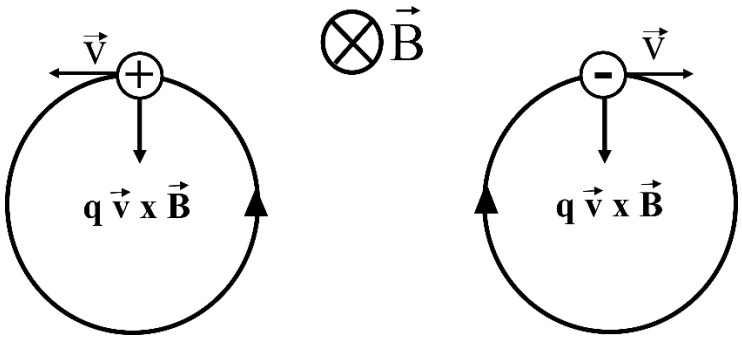


Figure 1. 11 Lorentz force acting on positively and negatively charged ions moving perpendicular to the direction of the generated homogenous magnetic field (B) (adapted from Marshall *et. al.* 1998).[51]

The ions orbiting inside the magnetic field have a unique frequency, called the cyclotron frequency, which is based upon their *m/z* ratio and the magnetic field strength. The *m/z* ratio of an ion is determined by measuring the cyclotron frequency.

$$\omega_c = \frac{qB}{2\pi m} \qquad \text{[Eqn. 13]}$$

where ω_c is the cyclotron frequency of the ion in Hertz (Hz), q is the charge of the ion, B is the magnetic field strength in Tesla (T), and m is the mass of the ion in Daltons (Da).

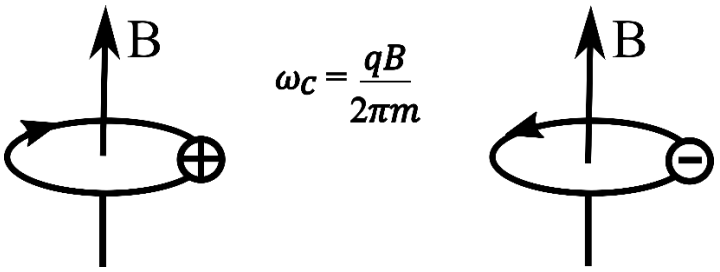


Figure 1. 12 Cyclotron motion of a positively charged ion (left) and a negatively charged ion (right) in the presence of magnetic field (B) (adapted from Marshall *et. al.* 2002).[52]

The cyclotron frequency of an ion is inversely proportional to the *m/z* of the ion, meaning ions with a higher *m/z* precess around the centre axis of the magnetic field, B at a lower cyclotron frequency and lower *m/z* ions precess around the axis of the magnetic field at higher frequencies. Ions with different *m/z* can therefore be separated due to the differences in their cyclotron frequencies. The cyclotron frequency of an ion is also independent of its velocity and kinetic energy according to Eqn. 13, which is a key factor for the ability of FT-ICR MS in achieving high resolving power.

Trapping motion

Ions moving parallel to the magnetic field in the ICR cell move unconstrained along this direction, which can be a problem as the ions can escape along the z-axis even if they are contained radially by the magnetic field. The ions can be prevented from leaving the cell by applying trapping plates perpendicular to the magnetic field to create a potential well. A small, positive voltage is applied to the trapping plates to store the positive ions, and a small negative voltage is applied to trap the negative ions. When the energies of the ions are lower than the trapping voltage, the ions are trapped in the ICR cell and oscillate back and forth between the plates. The angular frequency of the ion's trapping motion, ω_z can be described by the following equation:

$$\omega_z = \sqrt{\frac{2qV_{trap}\alpha}{ma^2}} \qquad \text{[Eqn. 14]}$$

where ω_z is the trapping oscillating frequency induced by the trapping motion, q is the charge of the ion, V_{trap} is the trapping potentials applied to the trapping electrodes, m is the mass of the ion, α is the trapping scale factor which depends on the geometry of the ICR cell, and a is the distance between the two trapping plates of the ICR cell.

Magnetron motion

Application of electric potentials to the trapping plates is an effective way to prevent analyte ions from leaving the ICR cell along the z-axis but this generates an electric field between the plates, which also, therefore, creates a small radial electric field. A combination of the electric field and magnetic field induces a third unwanted motion, known as magnetron motion [Eqn. 15].

$$\omega_m = \frac{V_{trap}\alpha}{2\pi B a^2} \qquad \text{[Eqn. 15]}$$

where ω_m is the magnetron frequency of the ion, V_{trap} is the trapping potential applied to the trapping plates, α is trapping scale factor which depends on the geometry of the ICR cell, B is the strength of magnetic field, and a is the distance between the two trapping plates of the ICR cell.

The magnetron motion acts along the same axis as the cyclotron motion but at a much lower frequency (<10 Hz) and since the ions are affected by the electric fields generated from the trapping potentials, the actual frequency measured is known as the

"reduced" cyclotron frequency, $\omega_{measured}$. This can be represented by Eqn. 16 shown below:

$$\omega_{measured} = \omega_c - \omega_m \quad \text{[Eqn. 16]}$$

where ω_c is the pure cyclotron frequency and ω_m is the magnetron frequency.

The summarised ion motions (cyclotron motion, trapping motion, and magnetron motion) in a penning trap is depicted below.

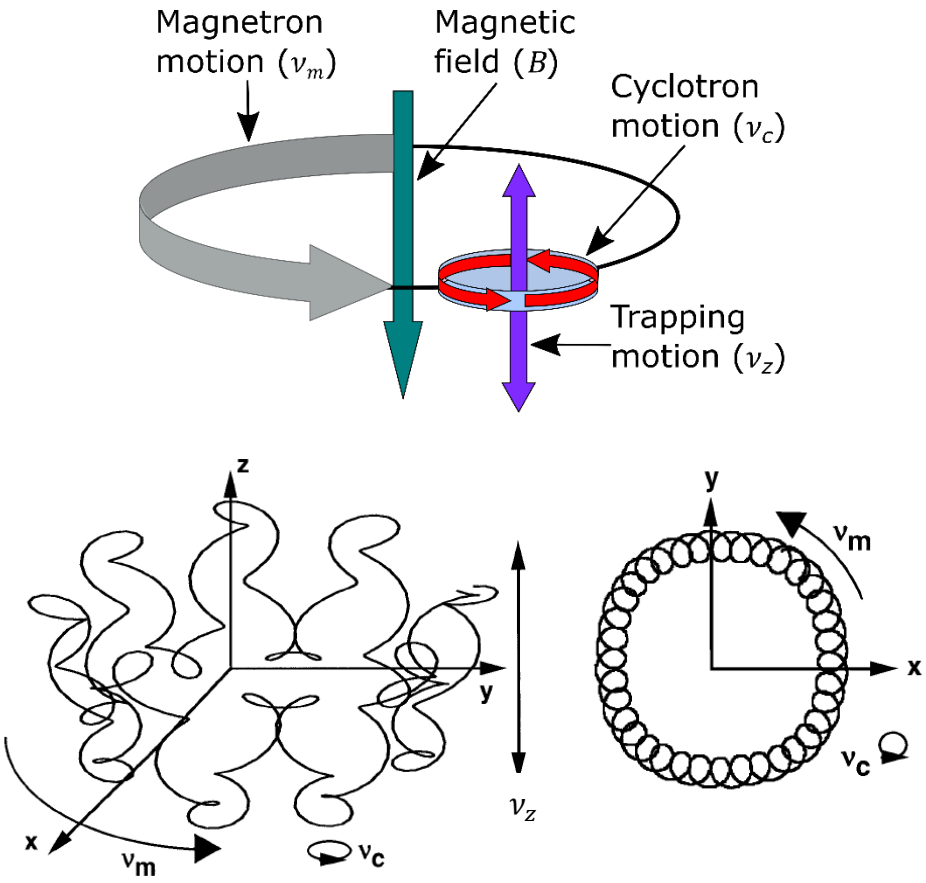


Figure 1. 13 Top figure: Summarised modes of ion motion in the FT-ICR analyser cell including cyclotron motion (ν_c), trapping motion (ν_z), and magnetron motion (ν_m) Bottom figures: Representation of the ion path with contributions from the cyclotron motion (ν_c), trapping motion (ν_z), and magnetron motion (ν_m) (reproduced and adapted from Marshall *et. al.*[51] 1998 and Amster *et. al.*[53] 1998).

Excitation and detection in the FT-ICR analyser cell

After the analyte ions are transferred to the ICR cell, the cyclotron radii of the trapped ions are generally small compared to the dimensions of the cell and therefore too small to generate a detectable signal as they only have a small amount of kinetic energy. To detect the ions, a radio frequency (RF) potential can be applied to the two excitation plates, an opposing pair of plates which lie parallel to the magnetic field axis, to excite the ions. The applied RF pulse will transfer energy to the ions trapped in the ICR cell. Since the cyclotron frequency of an ion is independent of its kinetic energy, any added energy from the RF pulse will allow the ions to travel faster and excite out to a larger cyclotron radius, while the cyclotron frequency remain the same around the ICR cell. The radius of the post-excited ion cloud can be shown by the following equation:

$$r = \frac{E_0 T_{excite}}{2B} \qquad \text{[Eqn. 17]}$$

where r is the radius of the post excited ion packet, E_0 is the electric field applied to excite the ions, T_{excite} is the excitation duration, and B is the magnetic field strength.

The excitation and detection events are shown in Figure 1.14 below:

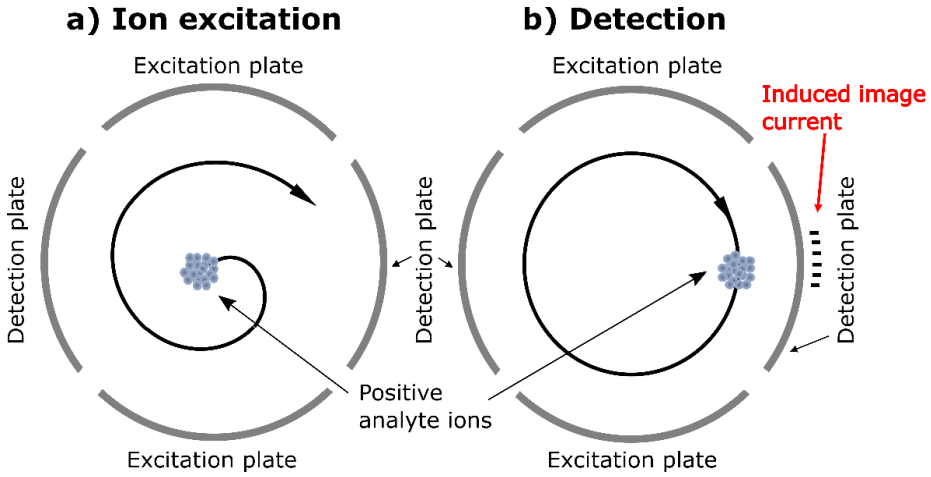


Figure 1. 14 a) excitation and b) subsequent detection of a species within an ICR cell. Courtesy of Bruker Daltonics, Bremen, Germany.

The two most common excitation methods applied to FT-ICR MS are frequency-sweep excitation (RF chirp) and stored waveform inverse Fourier transform (SWIFT).

The RF chirp was developed by Marshall *et. al.* in 1974.[39,54] In this method, many frequencies are applied during the excitation event and the ions are excited in a stepwise

manner with a broad frequency sweep across the frequency range of interest. However, the RF chirp also generates profiles which are not completely flat forming horns at the beginning and end of each pulse (Figure 1.15a).

To overcome the issue with the non-uniform excitation profile generated by the RF chirp, the SWIFT method of ion excitation in the ICR cell was proposed by Marshall in 1985.[55] SWIFT predicts the desired frequency profile (rectangular excitation profile) and then uses an inverse FT to calculate the time domain signal needed to get a perfect rectangular excitation profile.

SWIFT can also be used to isolate ions of interest by overexciting the unwanted species to a larger cyclotron orbit so that once they hit the electrodes, they are neutralised and ejected while the desired ions are retained in the ICR cell.

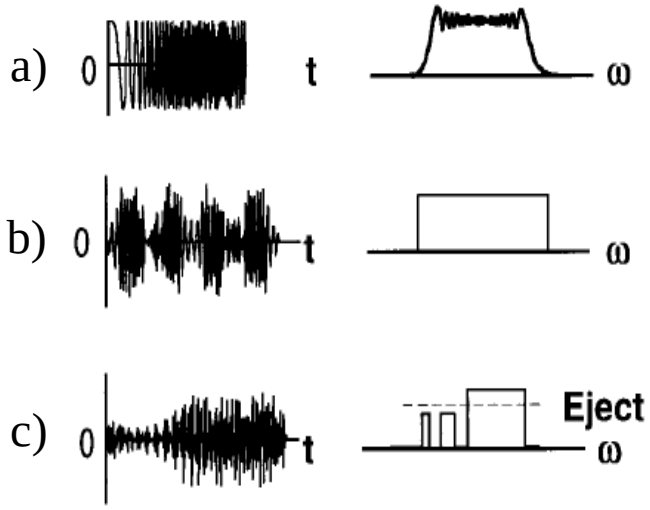


Figure 1. 15 excitation methods applied in FT-ICR cell using a) RF Chirp and b) SWIFT c) ejection of ions from the ICR cell by applying the SWIFT excitation pulse (Reproduced from Marshall *et. al.* 1998).[51]

Post ion excitation to a large enough orbital, where ions can be detected within the ICR cell, the RF pulse, which is applied to the excitation plates shown in Figure 1.14 is turned off. The coherent ion packets continue to precess at the excited radius around the ICR cell, where they attract electrons to the first detection plate and then the second one, through an external circuit. By monitoring the alternating current i.e., the image current, the cyclotron motion of ions produces an image signal which can be detected, amplified, and digitised. Thus, a time-domain spectrum is recorded, with the fast Fourier transform (FFT) applied to generate a spectrum of the signal intensity against the frequency. A mass

calibration function is then applied to convert the frequency spectrum into a mass spectrum, which is displayed on the user control software.

ICR cells

The ICR cell, also known as the analyser cell, is a significant component of any FT-ICR MS instrument. In the ICR cell, analyte ions are stored, excited, mass analysed, and detected. Over the years, many different types of analyser cells have been designed and implemented into FT-ICR mass spectrometers. The different types of analyser cells include the cubic cell, cylindrical cell and open-ended cylindrical cell, Infinity Cell, ParaCell and many others can be found in a recently published review article evaluating ICR cell designs of the past and present by Nikolaev and Lioznov.[56]

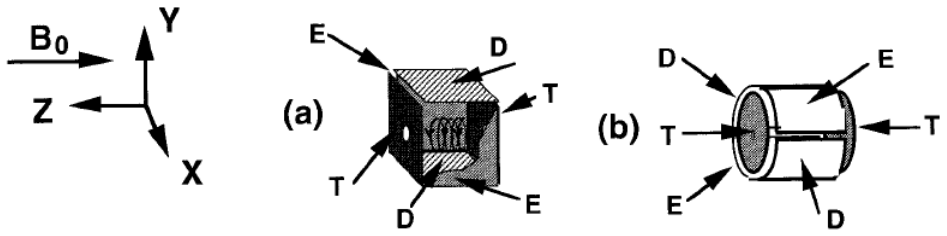


Figure 1. 16 Ion trap configurations of a a) cubic ICR cell and b) cylindrical ICR cell where E is the excitation plate; D is the detection plate and T is the trapping plate (reproduced from Marshall *et. al.* 1998).[51]

The cubic ICR cell was one of the earliest designs for FT-ICR cells[57] whereas the general cylindrical cell, which can be further classified into different configurations is more commonly used. The cubic cell is composed of six plates, with each pair of opposing plates used for excitation, detection, and ion trapping. The cylindrical cell also has six plates that have the same purpose as those of the cubic cell.

Cylindrical ICR cells can be further classified into an open or closed configuration with circular end cap electrodes, one of which has segmented end caps, known as the infinity cell (Figure 1.17). The open cylindrical cell uses cylindrical trapping plates for ion confinement in the cell and this design allows for effective trapping of the ions in the z axis with the minimum influence of the ion cyclotron frequencies. The curved detection plates also allow for longer interactions with the ions in the cell. The end-capped closed cylindrical cells however are more likely to be affected by electric field permeation resulting in the ion-packets orbiting away from the centre of the ICR cell and unwanted ejection of ions along the z-axis. The infinity cell introduced by Caravatti *et. al.*[58] was a solution to this problem caused by standard end-cap electrodes. The other solution was

the Beu *et. al.*[59] capacitively coupled open cylindrical cell, which was a modified form of Gabrielse *et. al.*[60] open cell design. The problem of the z-axis ejection of ions is thought to come from the finite dimension of the standard ion traps hence the concept of the infinity cell is based on modelling the electric excitation RF field of an infinitely long cell with a cell of finite dimensions.

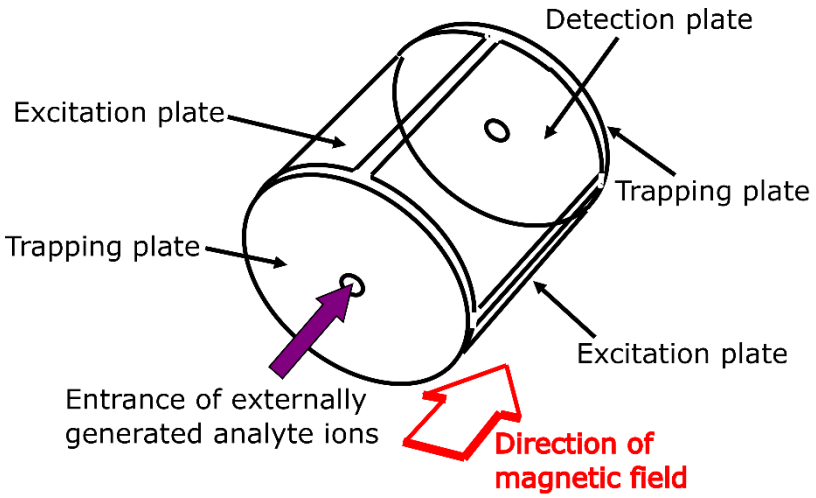


Figure 1. 17 Cylindrical ICR cell design (Infinity Cell), courtesy of Bruker Daltonik, Bremen, Germany.

High mass accuracy and high resolving power

High mass accurate measurements and high mass resolving power can be achieved with FT-ICR MS. Resolving power is a measure of how effectively the instrument can separate mass spectral peaks. The higher the number, the more effective the mass spectrometer is at separating closely spaced peaks. In FT-ICR MS, resolving power increases linearly with the magnetic field strength.

$$Resolving\ power = \frac{m}{\Delta m} \qquad \text{[Eqn. 18]}$$

where m is the *m/z* of the peak in question and Δm is the width of the peak at half its own height, also known as full width half maximum (FWHM).

Mass accuracy refers to the deviation of the measured *m/z* value of a certain analyte from the theoretically calculated *m/z* for that species.

$$Mass\ accuracy = \frac{\left(measured\frac{m}{z} - exact\frac{m}{z}\right)}{exact\frac{m}{z}} \times 10^{6} \text{ in parts per million (ppm)} \qquad \text{[Eqn. 19]}$$

Accurate measurement of the cyclotron frequency dictates the mass accuracy of FT-ICR MS.

$$Mass\ accuracy\ (ppm) \propto B^{2} \qquad \text{[Eqn. 20]}$$

As mass accuracy is directly proportional to the square of the magnetic field strength, it is therefore also beneficial to use higher magnetic fields as this enables much more accurate mass analysis of the target compounds.

1.1.3. SolariX 12 T FT-ICR Mass Spectrometer

The Bruker 12 T SolariX FT-ICR MS used for the work presented in this thesis has the configuration shown in Figure 1.18. Samples are introduced into the front source region of the instrument, where ions can be generated via ESI, nESI or MALDI. The quadrupole can simply be used for transmission of ions or it be used for isolating ions of a selected range of *m/z* (for MS/MS purposes) before transfer to the hexapole collision cell. In the collision cell, ions can be accumulated and/or fragmented via collisions with neutral gas molecules (e.g., for CAD). A hexapole transfer optic then enables ion transfer to the ICR cell. Once the ions enter the ICR cell, which sits in a homogenous magnetic field provided by the superconducting magnet, the ions can be trapped via application of voltages on the certain plates of the ICR cell, such as the front and back trap plates. An indirectly heated hollow dispenser cathode can be used to generate electrons for electron-based fragmentation, such as ECD or EID in the ICR cell. Photodissociation methods can also be implemented and applied to the trapped ions in the ICR cell as the laser pulses from an ArF 193 nm excimer laser and a CO_2 laser can be directed to the ions in the cell for IRMPD and UVPD experiments, respectively. All ions of interest are thus excited and detected in the ICR cell as mentioned in detail above.

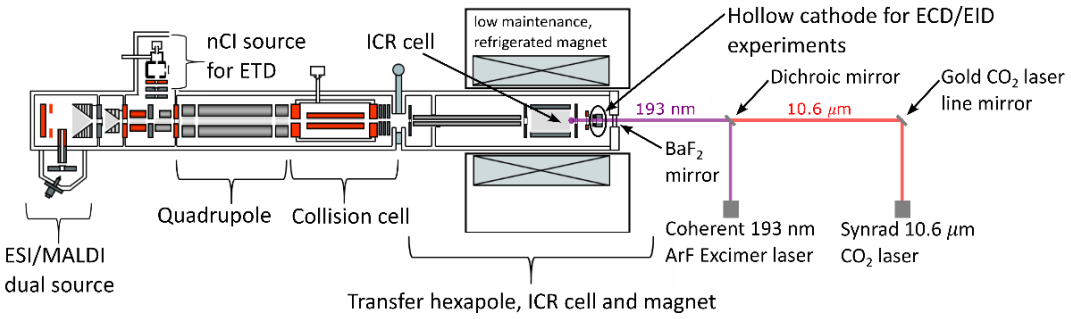


Figure 1. 18 Schematic representation of the 12 T Bruker SolariX FT-ICR MS (courtesy of and adapted from Bruker Daltonik, Bremen, Germany).

1.1.4. Tandem Mass Spectrometry Techniques

To generate structural information, unimolecular dissociation is a crucial reaction in mass spectrometry. Various tandem mass spectrometry (MS/MS or MS^n) techniques are developed to break ions apart causing fragmentation.

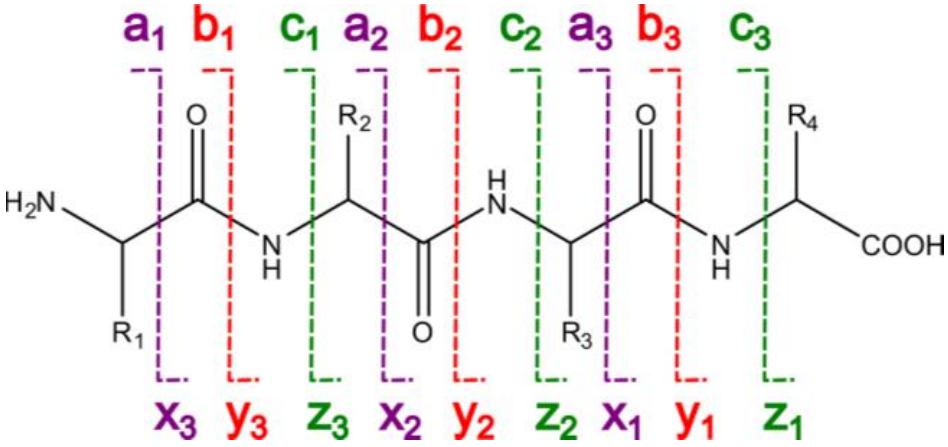


Figure 1. 19 Roepstorff nomenclature as modified by Biemann of possible fragments generated from peptides and proteins.[61]

1.1.4.1. Collisionally Activated Dissociation

Collisionally activated dissociation tandem mass spectrometry (CAD MS/MS), also known as collision induced dissociation (CID) was first discovered and implemented by Keith Jennings in 1968.[62,63] In this process, an electric potential is applied to analyte ions to accelerate the ions, which then undergo collisions with neutral gas molecules such as argon, neon, helium or nitrogen. Multiple collisions with the neutral gas molecules result in a kinetic energy transfer to the analyte ions. The kinetic energy is converted to internal energy, which is rapidly distributed across the analyte ion, causing a "slow heating" effect because the ions are activated and deactivated via the low energy collisions, resulting in fragmentation. For peptides and proteins, the amide bond usually dissociates first because it is the weakest bond. This takes place via a proton transfer rearrangement generating mostly b and y ions with losses of neutral molecules such as H_2O, NH_3 and CO_2 also observed in the CAD MS/MS spectrum.[64]

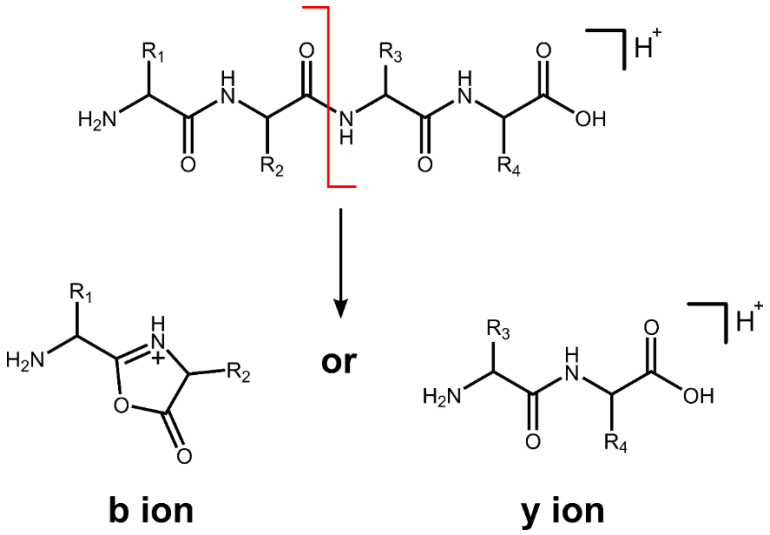


Figure 1. 20 Diagram showing the amide bond breakage to generate b and y ions in CAD MS/MS.

CAD MS/MS is a standard fragmentation method that is compatible with most commercial mass analysers such as quadrupole, ToF, FT-ICR and Orbitrap. Hence to this day, this method has been commonly and effectively used to sequence peptides, although it has shown a lower effectiveness for sequencing larger peptides and proteins. Fragments are generally observed near the N and C-termini region, with less coverage of the core parts of the proteins. As mentioned previously, since CAD breaks the weakest bond first (amide bond in peptides and proteins), labile modifications such as phosphorylation, which are weaker than the peptide backbone, are not preserved and will be dissociated during the process of CAD MS/MS. This makes it difficult to determine the location of the modification site and this information will be lost.

1.1.4.2. Infrared Multiphoton Dissociation

Infrared multiphoton dissociation tandem MS (IRMPD MS/MS) is another commonly used fragmentation method that has shown wide use over the years for the study and characterisation of biomolecules.[65] It is usually applied to ion trap instruments such as quadrupole ion trap mass spectrometers[66-71] and FT-ICR MS.[65,72-77] IRMPD, as the name suggests is a photon-based step-wise activation process. It is very similar to CAD, as they are both "slow heating" techniques, which can be achieved via heating of vibrational modes and internal vibrational redistribution of energy (IVR).[78] For biomolecules, IRMPD also generates b and y ions as observed with CAD, and often accompanied with losses of small neutral molecules, such as H_2O and CO_2.

In an IRMPD MS/MS experiment, for example in an FT-ICR MS, a continuous-wave carbon dioxide (CO_2) laser generates photons at a wavelength of 10.6 μm, which can be absorbed by many types of molecules, especially biomolecules such as peptides and proteins. After the CO_2 laser is used to irradiate the trapped ions of interest in the ICR cell, this results in preferential cleavage of the most labile bonds (amide bonds in biomolecules), generating fragments which can be readily detected. Absorption of dozens or even hundreds of photons are normally required to cause dissociation since the absorption of each photon generated by a CO_2 laser corresponds to ~ 0.117 eV of energy, so a high laser power or a long laser pulse length (irradiation time) is often required for dissociation. Although CAD and IRMPD generate similar fragmentation spectra, they are tuned differently. CAD MS/MS requires optimisation of the collision energy (acceleration voltage applied to the ions of interest into the collision gas) whereas IRMPD MS/MS requires fine tuning of the laser power and irradiation time.

1.1.4.3. Ultraviolet Photodissociation

Ultraviolet photodissociation (UVPD) is a laser photodissociation technique like IRMPD, where samples are irradiated with photons to excite molecules and cause dissociation. However, UVPD relies on the absorption of a UV photon instead of an IR photon, which leads to electronic excitation of a suitable chromophore. It is a higher energy activation method, shown to result in extensive fragmentation of peptides and proteins and has been carried out at several different wavelengths such as 266 nm, 213 nm, 193 nm, and 157 nm. The corresponding energy per photon is around 3-8 eV depending on the wavelength used. At 266nm, the absorption in peptides and proteins generally occurs at the aromatic side chains of tyrosine and tryptophan. At 193 nm and 213 nm, which correspond to higher energy photons, excitation of the peptide backbone can take place and at 157 nm, excitation of most bonds becomes possible, including the molecules in the air.

After absorption of a UV photon, and the resulting promotion of an electron to an excited electronic state, two main dissociation mechanisms for UVPD have been suggested, namely direct dissociation and internal conversion. Direct dissociation takes place if the electron is excited into or can relax into a dissociative orbital. Internal conversion is based on IVR (as mentioned in section 1.1.4.2 IRMPD), where the photon energy is converted into vibrational modes and fragmentation occurs in the ground state. Hence the fragments produced will be like those generated by CAD and IRMPD. For

UVPD, it has been suggested that both dissociation pathways may occur simultaneously.[79]

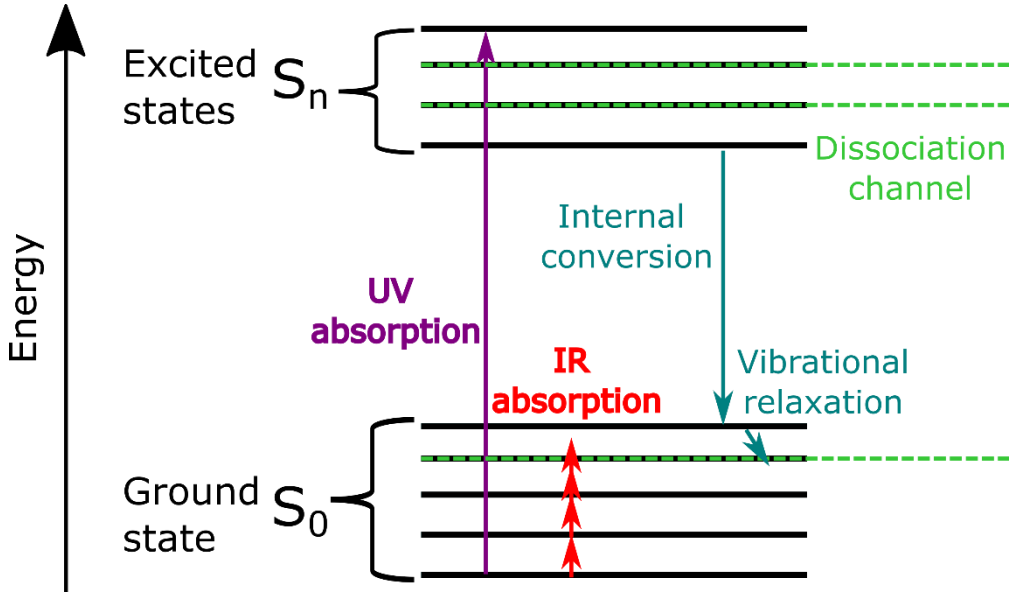


Figure 1. 21 Energy diagram of photodissociation mechanisms.

UVPD primarily generates a/x ions by cleavage of Cα - C bonds in addition to b/y and c/z ions,[80,81] increasing the complexity of UVPD MS/MS spectra but this may also provide more fragmentation information for the sample of interest and more sequence information for biomolecules.

Over the past few decades, UVPD has been successfully applied to the analysis of a wide range of molecules including peptides,[80-83] proteins[84-86], lipids,[87-89] oligonucleotides, and small molecules. Additional benefits of UVPD include the ability to access both high and low energy product ions due to the combination of the dissociation pathways, optimisation of the photon energy deposited based on the laser wavelength and fast scanning speed of UV lasers, which enables ease of implementation on a variety of mass analysers such as quadrupole ion traps, TOF, FT-ICR MS and Orbitrap.

1.1.4.4. Electron Capture Dissociation

Electron capture dissociation tandem MS (ECD MS/MS) was first developed in 1998 by Zubarev *et. al.* and it was used to fragment multiply charged gas-phase protein ions in an FT-ICR MS.[90] To perform ECD, a beam of low energy electrons (~ 0.1-3 eV) produced by a hollow dispenser cathode is applied to multiply charged positive ions, which are normally trapped in the MS (for example in the ICR cell in FT-ICR MS). The

electrons are captured by the positively charged ions creating a radical cation species known as the charged reduced species (CRS) (Eqn. 21). Once the electrons are captured, the CRS can dissociate into different fragmentation pathways and for ECD MS/MS of peptides and proteins, this results in breakage of the N-C_α bond (NH-CHR bond), producing c (enolamine) and z• (α-amide radical) fragment ions (Figure 1.22).[90,91] The electron capture process and generation of the CRS can be represented by the equation shown below:

$$[M + nH]^{n+} + e^- \ (0.1 - 3\ eV) \rightarrow \ [M + nH]^{n-1+\bullet} + fragments \qquad \text{[Eqn. 21]}$$

where M is the precursor ion, n is the charge of the ion, and e is the electron.

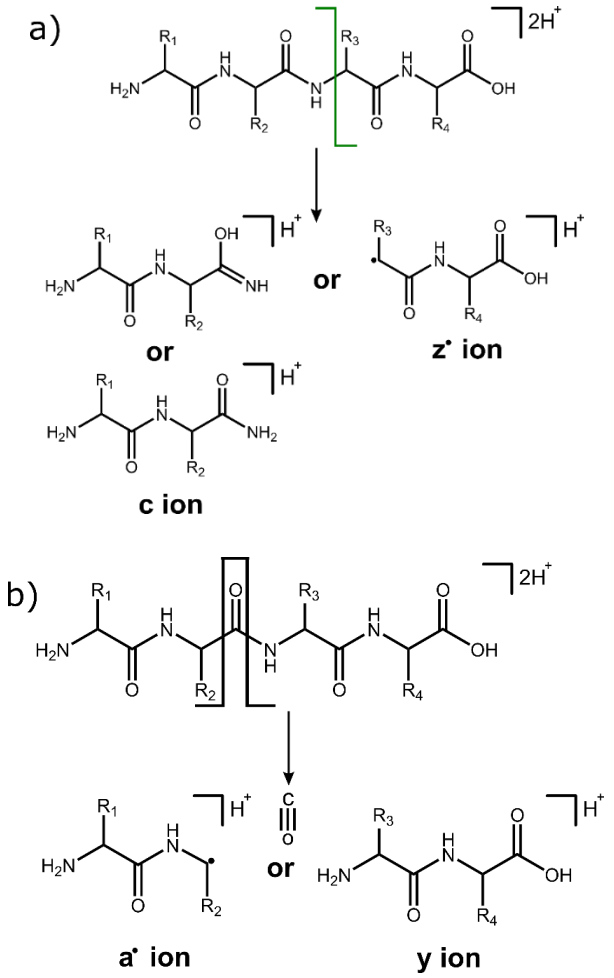


Figure 1. 22 Diagram to show the generation of the a) c and z ions and b) a and y ions in ECD MS/MS fragmentation.

ECD is only suitable for multiply charged precursor ions because at least one electron is captured by the precursor ion during the ECD fragmentation process resulting in a charge neutralisation effect (Eqn. 18). For singly charged precursor ions, after electron capture, this would generate neutral species which cannot be detected by MS.

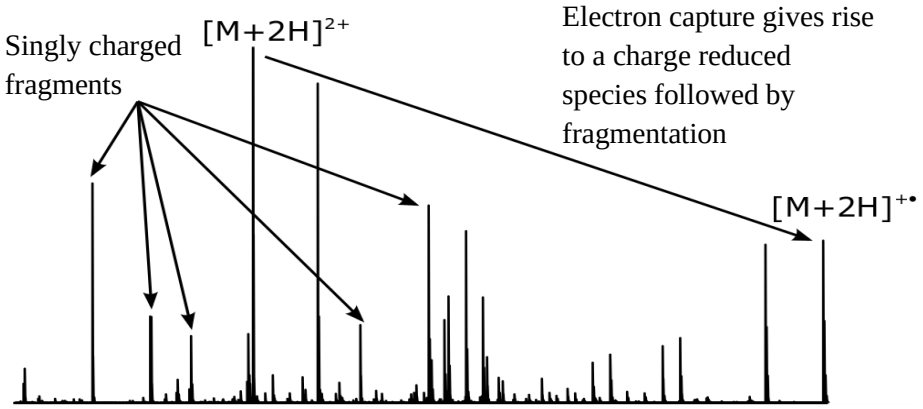


Figure 1. 23 Example of an ECD MS/MS spectrum of a doubly protonated peptide.

The dissociation mechanism for the generation of the c/z fragment ions from the CRS has been under serious discussion for many years. However, the most widely accepted mechanisms for ECD include the Cornell Mechanism and the Utah-Washington mechanism (UW mechanism).

The Cornell mechanism was proposed by McLafferty *et. al.* in 1998.[90] This mechanism suggests that electrons are captured at a positively charged site, such as the amino group on the N-terminus and the side chains of basic amino acids (such as histidine, arginine, and lysine); forming a hydrogen atom. This hydrogen atom is attracted to an amide oxygen found at the peptide or protein backbone. This results in the formation of a carbon centred aminoketyl radical precursor ion. The N-Cα bond that is adjacent to the carbonyl group is then cleaved and generates the c and z ions (Figure 1.24).

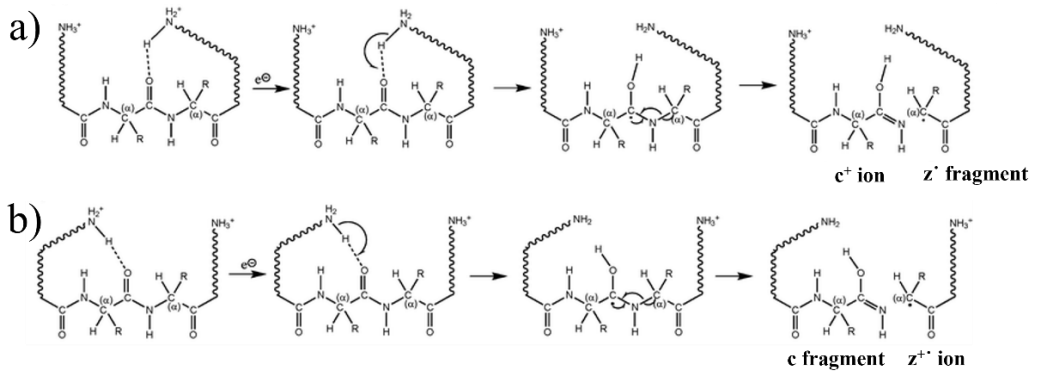


Figure 1. 24 The Cornell mechanisms of ECD MS/MS with the electron captured at the a) C-terminus to generate the c^+ and $z^{\cdot}$ ions and b) N-terminus to generate the c and $z^{+\cdot}$ ions.

The Utah-Washington (UW) mechanism is a combination of two previously proposed mechanisms to explain ECD: the Utah mechanism by Simons *et. al.* in 2003,[92] and this mechanism was further expanded on by Turecek *et. al.* in 2005, known as the Washington method.[93] The UW mechanism shows that electrons are directly captured at the amide group at the peptide or protein backbone, forming a carbon-centred aminoketyl radical precursor ion, cleaving the N-Cα bond adjacent to the carbonyl group. The same product ions as shown by the Cornell mechanism are formed, primarily c and z• fragments from the breakage of the N-Cα bond in peptides/proteins.

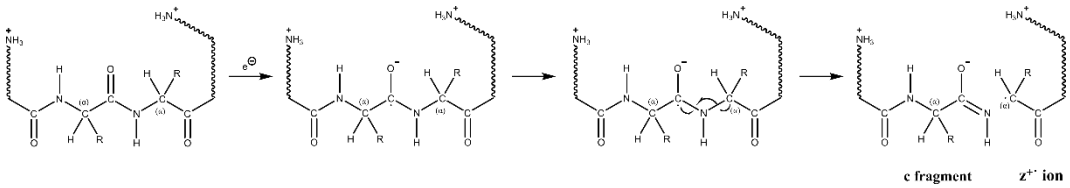


Figure 1. 25 The Utah-Washington mechanism of ECD MS/MS to generate the c and z ions.

Applications of ECD MS/MS

ECD is a radical based fragmentation method, where it is believed that dissociation of the bonds occurs before energy is redistributed among the molecules (non-ergodic hypothesis by Zubarev *et. al.*[90]) although others suggest due to it's low-energy radical rearrangements, the nonergodic hypothesis is not necessary.[93,94] Regardless, ECD offers several advantages over the 'slow heating' activation methods such as the useful contribution to de novo sequencing of biomolecules, preservation of label PTMs and non-covalent interactions, preferential cleavage of disulphide bonds, and differentiation of isomeric peptides and proteins.

De Novo sequencing

De novo sequencing is a method to determine the peptide or protein sequence by using fragment ion information available from an MS/MS spectrum. With ECD MS/MS, a high sequence coverage has consistently been achieved in the top-down spectra of proteins.[90,95,96] Extensive cleavage coverage can be obtained with ECD alone or with a combination of ECD and other MS/MS methods such as CAD or IRMPD.

Preservation of fragile PTMs and non-covalent interactions

As a result of the radical-based process, if the species of interest contains weakly-bound molecules/interactions, these may be preserved in the fragment ions detected, which is a huge benefit over threshold techniques such as CAD and IRMPD. The ability to retain delicate/fragile modifications/interactions has been an exceptional feature of ECD MS/MS and has enabled the study of fragile PTM's such as carboxylation, sulphation, O-glycosylation, N-glycosylation, disulphide bonds, and phosphorylation.

Disulphide bonds

A disulphide bond is a covalent bond between two sulphur atoms. They can form via oxidation between thiol (-SH) groups on two adjacent cysteine residues in peptides and proteins. Disulphide bonds provide stability to the protein because only large amounts of energy can break the covalent cross-link hence it is difficult to fragment protein ions when CAD and IRMPD MS/MS is used. ECD on the other hand, has been shown to prefer electron capture at the disulphide bond site, which results in a reduction of the S-S bond and allows fragmentation on the protein backbone (Figure 1.26).[92,97-99]

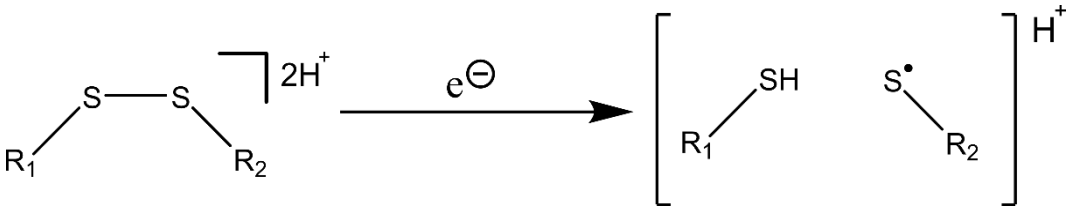


Figure 1. 26 Diagram to show the electron is captured by the disulphide bond and results in the S-S bond cleavage.

Differentiation of the isomeric products of deamidation

Isoaspartic acid/aspartic acid and γ-glutamic acid/ α-glutamic acid are the common products obtained during deamidation of peptides and proteins (detailed discussion in section 1.2.2.1).[100] The mass of isoaspartic acid and γ-glutamic acid are exactly the same as aspartic acid and glutamic acid respectively, the only difference is the shift of one bond to put a methylene group (CH_2-) on the backbone. ECD can result in the cleavage of Cα-Cβ bond and the formation of specific fragments for isoaspartic acid and γ-glutamic acid from aspartic acid and α-glutamic acid, respectively.[101-105]

For isoaspartic acid residues, cleavage of the C_α-C_β bond results in the breakage of the polypeptide chain and the formation of c+57 (C_2O_2H) and z-57 (C_2O_2H) fragment ions (Figure 1.27).

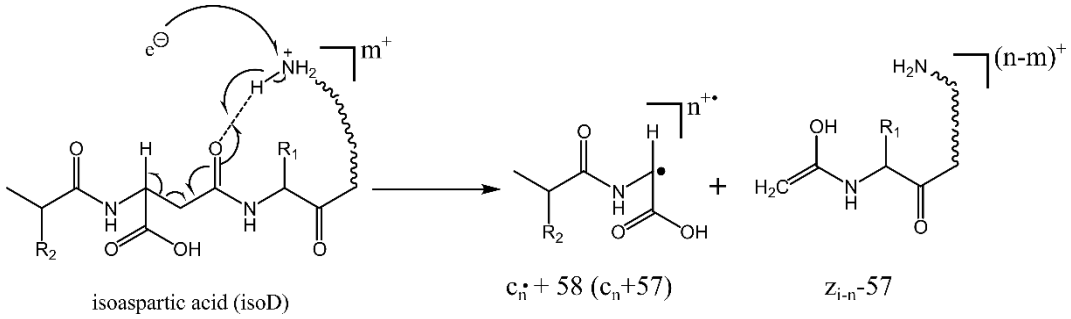


Figure 1. 27 Diagram to show the ECD fragmentation mechanism at the iso aspartic acid deamidation site to generate a (c+58 Da) ion and a (z- 57 Da) ion.

1.1.4.5. Electron Induced Dissociation

For the analysis of singly charged peptides and small molecules, ECD would not work as the precursor ion needs to be at least a 2+ charge to produce 1+ detectable fragments after the electron capture process otherwise for 1+ precursor ions this would result in a charge neutralisation effect, generating neutral species that cannot be detected by the MS. To combat the charge limitation problem with ECD for singly charged species, Budnik *et. al.*[106] and Fung *et. al.*[107] investigated the effect of tuning electron energy during ECD to find a region for electron-based dissociation of 1+ ions.

Electron induced dissociation (EID) was first proposed by Cody *et. al.* in 1979,[108] and the technique was named electron induced excitation of ions from organics (EIEIO). EID involves the irradiation of singly charged ions with higher energy electrons causing further ionisation/excitation followed by dissociation of the precursor ion.

The EID process is shown by Equations 22 and 23.

$$[M+H]^{+} + e^{-}_{fast} \rightarrow [M+H]^{2+\bullet} + e^{-}_{slow} \quad \text{[Eqn. 22]}$$

$$[M+H]^{+} + e^{-}_{slow} \rightarrow [M+H]^{+} + fragments \quad \text{[Eqn. 23]}$$

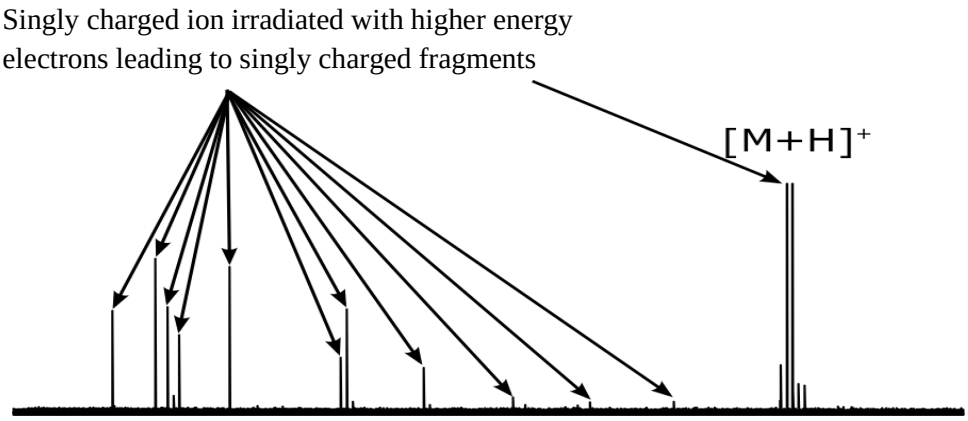


Figure 1. 28 Example of an EID MS/MS spectrum of a small molecule.

For singly charged small molecules this produced a range of fragments, akin to EI, but available to selected MS/MS of selected ions, which allowed soft ionisation followed by MS/MS of various analytes. For singly charged biomolecules EID produced c/z• and a/x type fragments. EID has also been shown to cause a series of cross-ring cleavages and enable detailed characterisation of singly charged species as well as small molecules such as lipids[109], pharmaceutical compounds,[110,111] and other small molecules.

1.1.4.6. Other Electron-Based Fragmentation Methods

Another fragmentation analogous to the fragments generated by ECD is electron transfer dissociation (ETD), which was invented in 2004 by Syka *et. al.*[112] ETD uses gas-phase ion/ion chemistry to transfer an electron from singly charged aromatic anions to multiply charged ions. Reactive radical anions are generated by ionising the reagent molecules, i.e., fluoranthene or anthracene through a negative chemical ionisation process. The precursor ions and radical anions are then transferred to the ion trap MS where the ion-ion interaction takes place. During this process, electrons from the radical anions are transferred onto the precursor analyte ions, which results in the formation of radical cations and leads to further fragmentation, producing mainly c and z• fragment ions.

Hot ECD MS/MS is an interaction between precursor ions and more energetic electrons (~ 3-13 eV). In hot ECD, the chance of secondary fragmentation increases

which results in extensive fragmentation, due to the release of excess energy. Kjeldsen *et. al.* demonstrated that with the use of hot ECD, isomeric leucine and isoleucine residues could be distinguished due to the presence of w and d fragments that were generated.[113] Eqn. 24 demonstrates the hot-ECD process.

$$[M + nH]^{n+} + e^-_{fast}(3 - 13\ eV) \rightarrow [M + H]^{n-1+\bullet} + fragments \qquad \text{[Eqn. 24]}$$

As ECD can only be used on positively charged precursor ions, for negative precursor ions, electron detachment dissociation (EDD) was developed by Budnik and Zubarev *et. al.* in 2000.[114] EDD is the same as ECD in principle but is applied to negatively charged precursor ions and uses energetic electrons (~ 10-25 eV), resulting in Cα-C backbone cleavage and generates a, x, c, and z ions. EDD has been applied for the characterisation of biomolecules such as proteins,[115,116] carbohydrates such as oligosaccharides[117,118] and glycosaminoglycans.[119]

Electron-photodetachment dissociation (EPD) is a photon-based activation method but it is included in this section as it involves the detachment of electrons from multiply charged anions after absorption of UV photons. This results in the production of charge-reduced radical anions. The EPD process can be represented by Eqn. 25:[120]

$$[M - nH]^{n-} + h\nu \rightarrow [M - nH]^{(n-1)+\bullet} + e^- \rightarrow fragments \qquad \text{[Eqn. 22]}$$

The first EPD experiments were shown by Guan *et. al.* in 1996 on nucleotide polyanions using a 193 nm ArF excimer laser on an FT-ICR MS.[121] Kjeldsen *et. al.* also demonstrated the potential of EPD for peptide sequencing and mapping PTM's.[122] This was applied to dianions of tryptic peptides using an F_2 excimer laser (wavelength of 157 nm) on a quadrupolar ion trap (QIT) mass spectrometer.

1.2. Applications of Mass Spectrometry

The applications of MS are extensive and span a wide range of areas from environmental analysis, drug characterisation, protein biomarker identification and quantification, metabolomics, discovery proteomics for viral analysis and more. In this work, the highlighted applications of MS include the differentiation and quantification of disease-related isomeric metabolites and post translationally modified biomolecules. An additional focus of this thesis will also be placed on the topical development of MS methods and sample preparation techniques for detection of viral proteins from severe

acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for the coronavirus disease (COVID-19).

1.2.2. Differentiation of Isomeric Post-Translationally Modified Peptides

Post-translational modifications (PTMs) are reversible or irreversible chemical modifications that a protein undergoes via the addition or conversion of a modifying group to one or more amino acid in the sequence. These changes can modulate how a protein functions and are known to play a major role in controlling protein stability, localisation, and protein-protein interactions. Hence PTMs are often discussed in the context of understanding the roles that they play in diseases such as cancer and neurological disorders such as Alzheimer's disease and Parkinson's disease. The most common and frequently discussed PTMs include phosphorylation, glycosylation, deamidation, ubiquitination, methylation, and sulfation. In this section, the significance of deamidation and methylation are discussed further as these PTMs are observed in the isomeric biomolecules studied in this thesis.

1.2.2.1. Isomeric Products of Deamidation

Deamidation is a non-enzymatic post-translational modification (PTM) of peptides and proteins. It is generally referred to as a common degradation or aging mechanism that takes place in proteins, often but not always, resulting in the reduction or complete loss of biological activity of the protein. Deamidation has been used as a marker for protein ageing, often referred to as molecular clock,[123] particular in disease related proteins such as amyloid-beta,[42,124] α-synuclein[125] and α-crystallin.[126-128] Studies have shown that deamidation can result in the enhanced aggregation or these long-lived proteins resulting in changes in their structure and reduced activity.

All peptides and proteins containing asparagine (Asn) and glutamine (Gln) are susceptible to deamidation, which is initiated at the amide ($-NH_2$) functional group on the side chain of both amino acids as shown by Figure 1.29. A nucleophilic attack from the side chain nitrogen to the carbonyl carbon atom results in the loss of ammonia ($-NH_3$). The reaction then proceeds via the formation of cyclic imide intermediates (succinimide for Asn and glutarimide for Gln). The imide intermediates of Asn and Gln hydrolyse at either one of the two carbonyls generating aspartic acid (D)/isoaspartic acid (isoD) and glutamic acid (α-Glu)/isoglutamic acid (γ-Glu), respectively.[100,123,129-131]

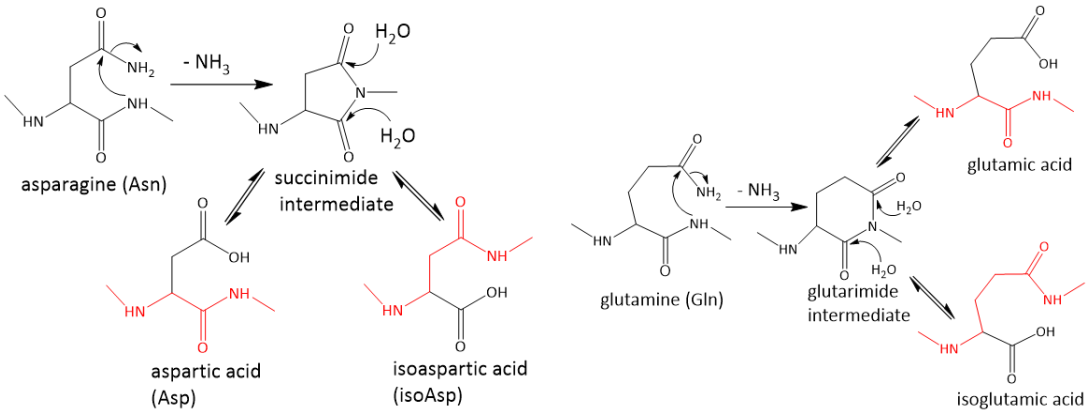


Figure 1. 29 Deamidation mechanisms of (a) asparagine and (b) glutamine.

Asn deamidation is much faster than Gln deamidation. Generally, the first order deamidation half-life of a linear peptide with Asn at neutral pH and at physiological temperature (37 °C) can be from half a day to approximately 1.4 years, whereas the deamidation half-life of Gln can range from approximately 1.6 years to 55 years.[100,123] The reason given for the drastic difference in the rate of deamidation of Asn and Gln is based on the structure of the amino acids. The distance between the backbone amido group (-NH-) and the side chain (-NH_2-) group for Asn is shorter compared to that of Gln.[100,132]

Deamidation can either be acid or base catalysed. At acidic pH, the rate determining step is the formation of the cyclization product (the succinimide intermediate) but at neutral or alkaline pH, the rate determining step is the removal of the leaving group (NH_3).[133] Under acidic conditions (low pH), deamidation by direct hydrolysis of the amide side chain becomes more favourable and Asp is formed directly as major degradation product by protonation of the amide leaving group. Under basic conditions, at a higher pH, the formation of the succinimide intermediate increases because of greater deprotonation of the peptide bond nitrogen.

Although pH is one of the most significant factors driving the process of deamidation, other factors such as temperature, buffer type, buffer ionic strength, and neighbouring residues to the deamidation site can also impact the rate of deamidation. For example, the rate of deamidation in proteins is enhanced with increasing temperature as weak interactions in proteins are broken, becoming increasingly unstable.[126,134-136] Hence, the temperature dependence of deamidation is affected by the preservation of covalent bonds in proteins as well as the size of sterically hindering groups adjacent to the deamidation site.[123] The rate of deamidation and succinimide intermediate formation also

increases for amino acids with less bulky and highly polar side chains, such as glycine and histidine. In a linear peptide with an -NG- in the peptide sequence, the deamidation half-life is 1 day (at pH 7.4, 37 °C, in 0.15 M Tris HCl) compared to a deamidation half-life of 53 days for -NC- peptides. Smaller and more flexible side chains of an adjacent residue to the deamidation site lowers the steric hindrance and more polar side chains help to stabilize the ionized transition state leading to the succinimide intermediate formation.

In linear and unstructured peptides, the ratio of the products of deamidation generally form in a ratio of 3:1, favouring isoD formation at pH 7.4.[100,123] At this pH, isoD formation is favoured over D formation due to the higher acidity of the isoD side chain.[137] As pK_a is lower for isoD than that of D, at neutral pH, that means less isoD is in the carboxylic acid form, which is generally favoured in the back-reaction due to the absence of the negative charge.[123] Hence, isoD is the major product at equilibrium.

Previous methods used for the study of deamidation and the products of deamidation can be separated into chemical and non-MS instrument-based methods. Some chemical-based detection methods include Edman degradation,[138-140] proteolytic digestion with enzymes such as endoproteinase Asp-N,[141] and enzymatic detection using protein L-isoaspartyl methyltransferase (PIMT).[142-144] On the other hand, instrument-based detection methods include nuclear magnetic resonance (NMR) spectroscopy,[145] fluorescence spectroscopy,[128,146] and liquid chromatography (LC), which is a commonly used method for separation of the non-deamidated and deamidated variants.[147-149]

Deamidation of peptides and proteins can easily be detected using most mass spectrometers, as the deamidation reaction produces a mass difference of + 0.984 Da, resulting from the mass conversion of Asn (-NH_2) to a mixture of D and isoD (-OH).[43,102,150-152]

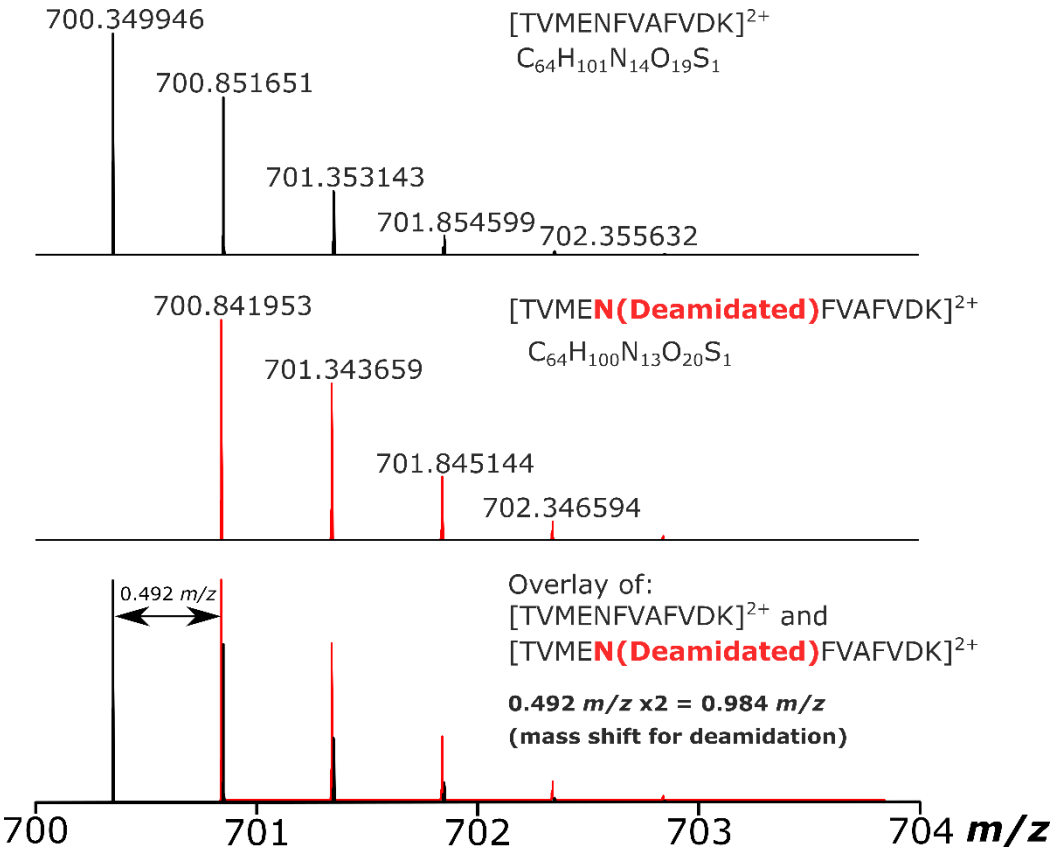


Figure 1. 30 Theoretical isotopic distributions of a tryptic peptide from bovine serum albumin (BSA) $[TVMENFVAFVDK+2H]^{2+}$. The top trace shows the MS with 0 % deamidation, the middle trace showing 100 % deamidation, and the bottom trace is an overlay of the non-deamidated and fully deamidated peptide MS spectrum.

Although the conversion of Asn to isoD and D can be detected with ease, the isomerization of D to isoD, where there are only minor structural differences between the isomers, proves to be more challenging as there is no mass difference observed in the mass spectrum.

Tandem mass spectrometry (MS/MS) methods used for the differentiation of isoD and D are provided in further detail in chapter 2.2. These methods include CAD, which has shown to generate diagnostic fragments and groups have utilised the differences in the abundance of immonium and b/y ions to discriminate between isoD and D peptides.[153,154]

As mentioned earlier in section 1.1.4.4, where the applications of ECD, particularly to differentiate the isomeric products of deamidation is discussed, Cournoyer *et. al.* found specific fragments for isoD (z_{i-n}- C_2O_2H and z_{i-n}+C_2O_2H), which were absent for D

peptides.[101,102] These characteristic fragments were used and continue to be used to confidently differentiate between isoD and D for peptides and proteins.

1.2.2.2. Methylated Histidine Isomers

Much like deamidation, methylation is another PTM that is commonly observed in proteins. Methylation involves the transfer of one methyl group to a nitrogen or oxygen (N-methylation or O-methylation respectively) on the amino acid side chains or the N-termini and C-termini in the protein sequence.[155] S-Adenosyl methionine (SAM/AdoMet) is a substrate that is a primary methyl group donor hence it is referred to as a methyltransferase. Protein methylation has been implicated in various biological processes including transcriptional regulation,[156] cellular signalling,[157] processing of RNA,[158] protein ageing/repair,[159,160] and regulation of protein-protein interactions.[161] Lysine and arginine residues predominantly undergo methylation and the role of lysine and arginine-specific methylation on histone proteins, has been extensively studied.[157,162] However, methylation can also occur on histidine, proline, and carboxyl residues.

Methylation of histidine was first observed and noted in actin[163] and myosin,[164,165] which are proteins found in almost every type of muscle tissue. Actin and myosin play an important role in cellular processes, especially in muscle contraction as skeletal muscle tissues are composed of repeating units of these proteins. Actin forms the thin filament in muscle cells, whereas myosin is known as the motor protein involved in generating the force in muscle contractions.

3-methylhistidine (3-MeH) is the modified amino acid and the product of actin and myosin methylation.[163,166,167] It is excreted via the urine and has often been used as a biomarker for measuring the rate of skeletal muscle protein breakdown.[168,169] Histidine methylation, however, can take place at two different positions, which are noted as 1-methylhistidine (1-MeH) and 3-methylhistidine (3-MeH). The isomeric metabolite to 3-MeH is 1-MeH, which has been used as a biomarker for meat consumption as it is derived from dietary sources.[170]

Different systems for numbering the atoms on the imidazole ring of histidine have been used by biochemists and organic chemists, where the nitrogen atom adjacent to the side chain was numbered as 1 by biochemists and the same nitrogen atom was numbered as 3 by organic chemists. Hence for simplification and to avoid any confusion, numbers will not be used and instead in accordance with the IUPAC guidelines, the position of the methyl group will be referred to as tele/τ-MeH and pros/π-MeH. The nitrogen atom on

the imidazole ring of histidine is denoted by tele ('far', abbreviated τ) as it is further away from the sidechain compared to the nitrogen atom which is denoted as pros ('near', abbreviated as π), which is closer to the side chain as shown in Figure below.

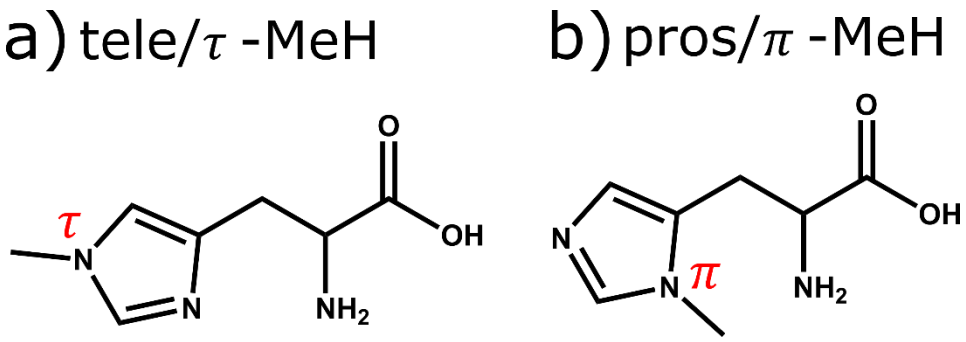


Figure 1. 31 Chemical structures for a) tele-methylhistidine (τ-MeH) and b) pros-methylhistidine (π-MeH).

The methylation of His73 (MeH73) in actin is important as it plays a role in actin polymerisation, which is important for cell motility.[171,172] As mentioned previously, SAM is generally known as the methyl group donor involved in methylation of mainly arginine and lysins residues, however for methylation of histidine, the SET domain containing 3 (SETD3) was discovered as the first human methyltransferase enzyme, which targeted histidine and was responsible for the methylation of Actin-H73.[173-177]

The urinary methylhistidine metabolites of actin and myosin have been qualitatively and quantitatively analysed in the past using synthesised standards and hyphenated mass spectrometry techniques such as GC-MS[178] and LC-MS. [179,180] As these metabolites are often present in plasma and urine, chromatographic separation is often required to reduce sample complexity. Methylation of actin at H73 can result in either tele-methylhistidine (τ-MeH) or pros-methylhistidine (π-MeH), which is difficult to distinguish by mass spectrometry as the metabolites and the tryptic peptides often used for these studies are isomeric hence there is no mass difference observed. However, in this work, the use of tandem mass spectrometry techniques will be thoroughly investigated for the differentiation of the methylated histidine isomers of actin.

1.2.3. Differentiation of Vitamin D Metabolite Isomers

Vitamin D compounds are a group of fat-soluble hormones. They are specifically referred to as secosteroids, which means that in terms of structure, they are organic compounds with a broken ring, exhibiting some biological activity. Vitamin D_3 (cholecalciferol) is generally referred to as vitamin D, which is taken as a supplement and

prescribed to patients with vitamin D deficiency. This can be distinguished from vitamin D_2, which is found and formed naturally in plants, whereas vitamin D_3 is formed in the skin of mammals via photosynthesis (Figure 1.32).

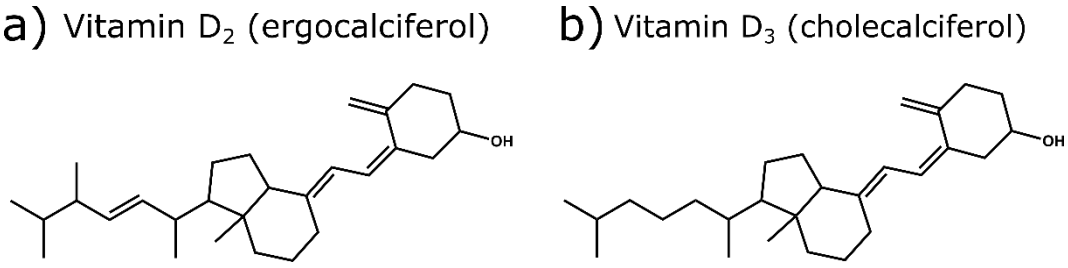


Figure 1. 32 Chemical structures of a) vitamin D_2 and b) vitamin D_3.

Vitamin D_3 is known to help regulate the amount of important minerals such as phosphate and calcium in the body, which is necessary to keep bones, teeth, and muscles healthy. Globally, it has been estimated that approximately one billion people across all age groups and ethnicities have low vitamin D levels.[181] A lack of vitamin D_3 can lead to bone deformities such as rickets in young children and bone pain in adults resulting in osteoporosis, a condition, where the bone weakens and becomes brittle.[181-183] Vitamin D deficiency has also been linked to other diseases such as diabetes, heart disease, Alzheimer's disease and schizophrenia.

Vitamin D_3 (cholecalciferol) is made in the skin from 7-dehydrocholesterol (7-DHC) under the influence of UV light (290-315 nm, UV_B) from the sun. The breakdown of vitamin D_3 via formation of 25-hydroxyvitamin D_3 ($25(OH)D_3$) in the liver and oxidation of $25(OH)D_3$ further generates the biologically active compound, 1,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$) and the isomer 24,25-dihydroxyvitamin D_3 ($24,25(OH)_2D_3$), which is generally known to be inactive as a hormone. However, some early experiments have shown that $24,25(OH)_2D_3$ may play a role in the development of skeletal tissues and healing of fractures in animal studies that were carried out. $1,25(OH)_2D_3$ is known to increase the amount of calcium that the gut can absorb and prevents calcium loss from the kidneys.

It is generally preferred that the active metabolite ($1,25(OH)_2D_3$) is used as a biomarker for vitamin D_3 sufficiency but this is difficult as its half-life is only a few hours and its concentration levels are very low in the blood. Therefore, the abundant yet inactive metabolite $25(OH)D_3$ is used clinically as a marker for vitamin D levels.

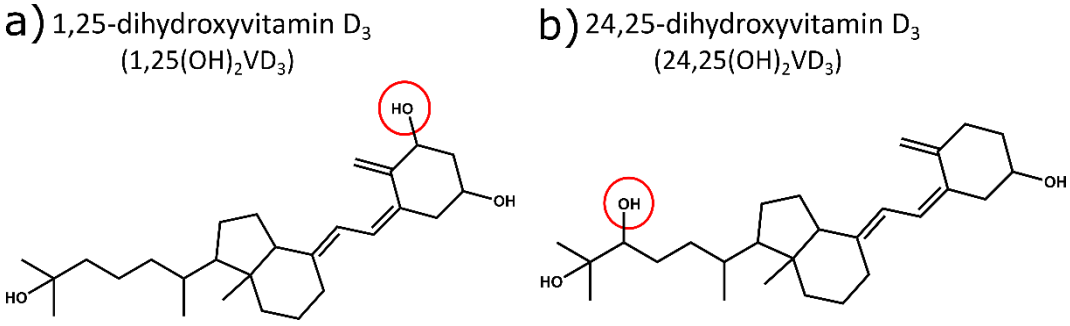


Figure 1. 33 Chemical structures of the dihydroxylated vitamin D_3 isomers a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with the hydroxyl groups responsible for the differences in the structures encircled in red.

Immunoassays such as chemiluminescence immunoassays (CLIA) and radioimmunoassays (RIA) are techniques often used for routine clinical analysis of vitamin D metabolites.[184-188] Generally, immunoassays involve the use of antibodies or the vitamin D binding protein (DBP) against the metabolite $25(OH)D_3$. Although the benefits of immunoassays include ease of use and high throughput, reproducibility remains to be one of the main issues as it is difficult to fully separate or recover $25(OH)D_3$ from the DBP as the target vitamin D metabolite is susceptible to matrix effects from the serum.[189]

More recently, the use of separation methods combined with mass spectrometry or MS techniques alone such as GC-MS,[190,191] LC-MS/MS,[192,193] MALDI-MS/MS,[194,195] and IMS-MS[196,197] have predominantly been used for vitamin D analysis, particularly LC-MS/MS due to improvements in sensitivity and specificity. The separation of critical vitamin D biomarkers from endogenous materials in matrices such as blood or human serum is aided by hyphenated chromatography methods. Further information including the benefits and shortcomings of these techniques for vitamin D analysis is provided in detail in chapter 4.2.

Differentiation of isomeric and epimeric vitamin D metabolites have also been carried out in previous studies with the use of MALDI-MS/MS and IMS-MS. Chapter 4 demonstrates the use of tandem MS techniques, enabling the differentiation of the dihydroxylated vitamin D isomers mentioned above via generation of characteristic fragments, which can then be utilised for quantification.

1.2.4. Quantification of Small Molecules, Peptides, and Proteins

Quantification can be defined as the determination of the absolute or relative amount (often expressed in concentration) of the analyte of interest or several target species present in a sample. For all types of molecules, the ability to quantify them has important implications. For example, for small molecules such as the vitamin D compounds that circulate in the blood, quantification of the biomarker $25(OH)D_3$ in human serum is essential to determine vitamin D sufficiency levels. If the biomarker falls below a concentration level of < 20 ng/ml in the serum, then it is vitamin D deficiency.[181-183,198] This information is essential when clinicians recommend treatments and helps towards disease diagnosis. Peptide and protein quantification becomes critical when they are expressed in cells or tissues, which may be diseased therefore the ability to determine the concentration of the target proteins in these cells can also aid studies in drug development. There are many different quantification methods that have been developed over the years by researchers, particularly in the field of proteomics and the most frequently used methods and concepts are discussed below.

1.2.4.1. Relative and Absolute Quantification

In proteomics, relative quantification is often described as comparing the relative amounts of peptides/proteins in samples. In relative quantitation, the samples that are to be compared are differentially labelled with stable isotopes. They are then combined and subjected to quantitative MS analysis. The peak intensity ratio between the heavy and light peptides is normally measured by MS to determine the relative change in the peptide/protein abundances in the samples.[199,200] For example, a common relative quantification method which involves chemical labelling is known as the isotope-coded affinity tag (ICAT) method, where a compound that has the stable isotope is coupled the cysteine residues in proteins.[201-203] Then follows the general relative quantification method mentioned with the differential labelling of the samples, which are mixed and they undergo protease digestion followed by affinity-purification of the cysteine-containing peptides.

Absolute quantification as the name suggests is not dependent on other quantities hence it can be used to determine the absolute concentration of distinct peptides or proteins within a sample. For absolute quantification, a known amount of isotope-labelled standard (synthetic peptides or proteins) is mixed with the analyte and the absolute amount (concentration) of the analyte is calculated from the ratio of the ion intensity

between the analyte and the standard.[199,204] Isobaric labelling such as tandem mass tags (TMT) and isobaric tags for both relative and absolute quantitation (iTRAQ) methods will be discussed in further detail.

1.2.4.2. Label-Free Quantification

Label-free quantification in MS is used to achieve a relative abundance of peptides or proteins between two or more samples without the use of stable isotopes for chemical binding and labelling of the target analytes. The main assumption for this type of quantification method is that the same peptides across different experiments can be directly compared if the conditions of the instrument, sample preparation and data acquisition parameters are closely controlled. The quantification method is often based on different categories of measurements such as measuring the ion abundance, ion abundance ratios or spectral counting.

Measurements of ion abundance are based on the direct comparison of the peak height or peak area between two identical peaks as the number of ions are generally reflected by the peak height/area.

Quantification using the ion abundance ratios involves calculating the peak area ratio of the analyte ion in the extracted ion chromatogram (EIC) and the EIC peak area of the internal control in sample 1 and then that is compared to sample 2.

The spectral count for a protein usually refers to the number of MS/MS spectra obtained from the digested peptides for that protein during the LC-MS/MS run. The more abundant the peptide, the more likely it will be selected for MS/MS.

1.2.4.3. Quantification using Isobaric Labels

The two types of labels commonly used are tandem mass tags (TMT) and isobaric tags for relative and absolute quantitation (iTRAQ).

Quantification of peptide pairs using TMTs was first demonstrated by Thompson *et. al.* in 2003.[205] The chemical tag is made up of the reporter group, a balance group and a primary amine specific reactive group. Labelled peptides release the mass reporter group during MS/MS fragmentation (normally CAD is applied), and the ratio of these reporter groups is used for relative quantification of peptides.

Peptides are covalently labelled with stable isotope molecules with tags of varying mass introduced from iTRAQ specific reagents in a protein digest via free amines at the peptide N-terminus and on the side chain of lysine residues.[206] Different samples with different tags induced into will be mixed together and are analysed at the same time. The peptide of interest from all samples will appear as a single peak because the masses of all tags are the same. Instead, the MS signal from the same peptide from all samples is summed. When the iTRAQ-peptides are fragmented by e.g., by CAD, the mass balancing group is released, liberating the isotope incorporated reporter ions, which can be used for relative quantification.

1.2.4.4. Quantification using SILAC

Stable isotope labelling with amino acids in cell culture (SILAC) involves growing cells and organisms in media that contain stable isotope labelled amino acids, such as 13C or 15N-labelled arginine or lysine residues.[207,208] Two populations of cells are grown in two different culture media, "light" or "heavy", with the "light" medium containing amino acids with natural isotopes, and the "heavy" medium containing the stable isotope-labelled amino acids and then both populations are mixed after proteins are grown. The cell cultures are then subjected to proteolytic digestion and analysis by LC-MS. It is expected that the each of the digested peptides contain at least one of the isotopically labelled amino acid, which would cause small *m/z* shifts, allowing for the relative quantification of the isotope-labelled peptides compared to the unlabelled peptides.

1.2.5. Detection of Clinical Biomarkers of Disease

Biomarkers (shortened for biological markers) are molecules, which act as indicators of biological processes that take place in the body.[209] These biomolecules e.g., proteins or nucleic acid, are typically used to measure the presence or progression of disease. They also serve as critical quantifiable characteristics in the development of drug treatments. To monitor diseases in patients, clinical biomarkers need to be identified and quantified in human samples such as urine, saliva, or blood in real-time.[210] Therefore, it is important that the clinical biomarkers in these types of samples are abundant, sensitive, and specific enough to predict the progress of disease states. It is also necessary to have instrumentation that can handle complex samples and provide sufficient sensitivity to detect the biomarkers in a time-efficient manner.

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a large, enveloped ribonucleic acid (RNA) virus responsible for causing the global outbreak

of the coronavirus disease 2019 (COVID-19). SARS-CoV-2 belongs to the Coronaviridae family and is categorised into the genus known as β coronavirus, showing similarities to known coronaviruses in the same category such as the severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV). SARS-CoV-2 has four main structural proteins, which are the spike (S) glycoprotein, envelope (E) glycoprotein, membrane (M) glycoprotein, and the nucleocapsid (N) protein. The main structural proteins play a crucial role in the infection of host cells,[211] fusion between viral and host cell membranes,[212] assembly of the virus[213] and release of the viral particles.[214]

Some of the current diagnostic tests for SARS-CoV-2 infection uses nucleic acid (e.g., RNA),[226] immune-based assays (serological tests),[227] and protein-based (enzyme-linked immunosorbent assay (ELISA)[228] detection methods. However, the main recommended diagnostic test involves detection and amplification of the viral RNA using methods such as reverse transcription polymerase chain reaction (RT-PCR), which is generally carried out on symptomatic patients during the acute phase of SARS-CoV-2 infection.

Matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) has emerged as a promising analytical tool for the rapid and sensitive detection of microorganisms. Over recent years, MALDI-TOF MS has found routine use in clinical microbiology laboratories as it is an easy to use, fast and high throughput technique. It has increasingly been used for microbial identification including detection of harmful bacteria in contaminated water and food, as well as for the detection of pathogens in blood and urine samples.[229]

Recent studies have demonstrated the potential for implementing MS and MALDI-TOF MS as rapid detection methods for SARS-CoV-2. Iles *et. al.* used a Shimadzu MALDI-TOF 8020 on gargle solutions spiked with cultures of SARS-CoV-2 to test for detection of the viral proteins.[233] Nikolaev *et. al.* successfully developed an LC-MS/MS method utilising a nano-HPLC coupled to a tims-TOF Pro (Bruker Daltonics) for the detection of tryptic peptides of the viral N protein from nasal epithelial swabs.[235] The use of chromatography in combination with MS also provides an extra dimension of separation of the species, particularly for the digested viral proteins. The detection methods and analysis of SARS-CoV-2 viral proteins are provided in further detail in chapter 5.2.

1.3. Overview of the Thesis

This thesis focuses on the implementation of advanced fragmentation techniques on a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) to distinguish between isomeric species of biological importance, including small metabolites and peptides. Application of MALDI-TOF MS and FT-ICR MS for the detection of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins is also investigated in this thesis.

In Chapter 2, ECD and UVPD MS/MS are used to differentiate between the isomeric isoD and D peptides, based on the generation of diagnostic fragments for ECD and the significant differences in the relative abundance of specific fragments for the isoD and D peptides with UVPD. An improved relative quantification method was developed and applied to determine the isoD content in deamidated peptides in tryptic digested bovine serum album (BSA).

Chapter 3 focuses on the application of various fragmentation methods, for the differentiation and relative quantification of isomeric N-methylated histidine containing peptides in the cytoskeletal protein, actin. MS/MS analysis resulted in the detection of diagnostic fragments for one isomeric peptide. Mixtures of the isomeric synthetic peptides containing τ-MeH and π-MeH were prepared and used to generate a calibration curve, which were the used for the relative quantification of τ-MeH in rabbit, chicken, bovine, and human actin samples.

Chapter 4 explores the application of various fragmentation methods such as CAD, IRMPD, UVPD and EID MS/MS for the distinguishing between isomeric dihydroxylated vitamin D_3 compounds. With all MS/MS methods, multiple isomer specific fragments were observed for the bioactive 1,25-dihydroxyvitamin D_3, which were absent in the MS/MS spectra for the inactive isomer, 24,25-dihydroxyvitamin D_3.

In Chapter 5, the use of linear MALDI-TOF MS and FT-ICR MS is investigated for the detection of SARS-CoV-2 biomarker proteins including the nucleoprotein (N-protein) and spike-protein (S-protein). Top-down and bottom-up optimisation experiments were carried out for the standard proteins and then applied to COVID-19 negative and positive patient swab samples.

Conclusions of this thesis and the future outlooks of the projects mentioned herein are summarised in Chapter 6.

1.4. References

(1) Thomson, J. J. Rays of positive electricity and their application to chemical analyses, ed.; Longmans, Green and Co.: London, New York [etc.], 1921; Vol., p x, 237 p.

(2) Aston, F. W. LXXIV. A positive ray spectrograph. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 1919, 38 (228), 707-714.

(3) Dempster, A. J. A new Method of Positive Ray Analysis. Phys. Rev. 1918, 11 (4), 316-325.

(4) de Hoffmann, E.; Stroobant, V. Mass Spectrometry: Principles and Applications, ed.; Wiley, 2013; Vol., p.

(5) Tal'roze, V. L.; Ljubimova, A. K. Secondary processes in the ion source of a mass spectrometer (presented by academician N.N. Semenov 27 viii 1952)—reprinted from report of the Soviet Academy of Sciences, Volume LXXXVI, -n5 (1952). J. Mass Spectrom. 1998, 33 (6), 502-504.

(6) Munson, M. S. B.; Field, F. H. Chemical Ionization Mass Spectrometry. I. General Introduction. J. Am. Chem. Soc. 1966, 88 (12), 2621-2630.

(7) Harrison, A. G. In Encyclopedia of Spectroscopy and Spectrometry (Third Edition), Lindon, J. C.; Tranter, G. E.; Koppenaal, D. W., Ed.^Eds.; Academic Press: Oxford, 2017, p^pp 188-194.

(8) Yamashita, M.; Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. J. Phys. Chem. 1984, 88 (20), 4451-4459.

(9) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. Science 1989, 246 (4926), 64.

(10) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int. J. Anal. Chem. 2012, 2012, 282574.

(11) Meng, C. K.; Mann, M.; Fenn, J. B. Of protons or proteins. Zeitschrift für Physik D Atoms, Molecules and Clusters 1988, 10 (2), 361-368.

(12) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization–principles and practice. Mass Spectrom. Rev. 1990, 9 (1), 37-70.

(13) Taylor, G. I. Disintegration of water drops in an electric field. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 1964, 280 (1382), 383-397.

(14) Melcher, J. R.; Taylor, G. I. Electrohydrodynamics: A Review of the Role of Interfacial Shear Stresses. Annu. Rev. Fluid Mech. 1969, 1 (1), 111-146.

(15) Reznik, S. N.; Yarin, A. L.; Theron, A.; Zussman, E. Transient and steady shapes of droplets attached to a surface in a strong electric field. J. Fluid Mech. 2004, 516, 349-377.

(16) Rayleigh, L. XX. On the equilibrium of liquid conducting masses charged with electricity. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 1882, 14 (87), 184-186.

(17) Davis, E. J.; Bridges, M. A. The rayleigh limit of charge revisited: light scattering from exploding droplets. J. Aerosol Sci 1994, 25 (6), 1179-1199.

(18) Konermann, L. A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionization. J. Am. Soc. Mass Spectrom. 2009, 20 (3), 496-506.

(19) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. Anal. Chem. 2013, 85 (1), 2-9.

(20) Wilm, M.; Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. Anal. Chem. 1996, 68 (1), 1-8.

(21) El-Faramawy, A.; Siu, K. W. M.; Thomson, B. A. Efficiency of Nano-Electrospray Ionization. J. Am. Soc. Mass Spectrom. 2005, 16 (10), 1702-1707.

(22) Karas, M.; Bachmann, D.; Hillenkamp, F. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. Anal. Chem. 1985, 57 (14), 2935-2939.

(23) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. Int. J. Mass Spectrom. Ion Process. 1987, 78, 53-68.

(24) Bahr, U.; Hillenkamp, F.; Karas, M. Analysis of biopolymers by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. Fresenius J. Anal. Chem. 1994, 348 (12), 783-791.

(25) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 1988, 2 (8), 151-153.

(26) Dreisewerd, K. The Desorption Process in MALDI. Chem. Rev. 2003, 103 (2), 395-426.

(27) Paul, W.; Steinwedel, H. A NEW MASS SPECTROMETER WITHOUT A MAGNETIC FIELD. 1953.

(28) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. Appl. Spectrosc. Rev. 2009, 44 (3), 210-230.

(29) March, R. E.; McMahon, A. W.; Londry, F. A.; Alfred, R. L.; Todd, J. F. J.; Vedel, F. Resonance excitation of ions stored in a quadrupole ion trap. Part 1. A simulation study. Int. J. Mass Spectrom. Ion Process. 1989, 95 (2), 119-156.

(30) Wolff, M. M.; Stephens, W. E. A Pulsed Mass Spectrometer with Time Dispersion. Rev. Sci. Instrum. 1953, 24 (8), 616-617.

(31) Wiley, W. C.; McLaren, I. H. Time-of-Flight Mass Spectrometer with Improved Resolution. Rev. Sci. Instrum. 1955, 26 (12), 1150-1157.

(32) Mamyrin, B. A.; Karataev, V. I.; Shmikk, D. V.; Zagulin, V. A. The mass-reflectron A new nonmagnetic time-of-flight high resolution mass-spectrometer. Zh. Eksp. Teor. Fiz. 1973, 64 (1), 82-89.

(33) Lawrence, E. O. In, Edlefsen, N. E.; Lewis, G. N., Ed.^Eds.; s.n.]: [New York :, 1930, p^pp.

(34) Lawrence, E. O.; Livingston, M. S. The Production of High Speed Light Ions Without the Use of High Voltages. Phys. Rev. 1932, 40 (1), 19-35.

(35) Lawrence, E. O.; Sloan, D. H. The Production of High Speed Canal Rays without the Use of High Voltages. Proc. Natl. Acad. Sci. U.S.A. 1931, 17 (1), 64-70.

(36) Hipple, J. A.; Sommer, H.; Thomas, H. A. A Precise Method of Determining the Faraday by Magnetic Resonance. Phys. Rev. 1949, 76 (12), 1877-1878.

(37) Sommer, H.; Thomas, H. A.; Hipple, J. A. The Measurement of e/M by Cyclotron Resonance. Phys. Rev. 1951, 82 (5), 697-702.

(38) Comisarow, M. B.; Marshall, A. G. Fourier transform ion cyclotron resonance spectroscopy. Chem. Phys. Lett. 1974, 25 (2), 282-283.

(39) Comisarow, M. B.; Marshall, A. G. Frequency-sweep fourier transform ion cyclotron resonance spectroscopy. Chem. Phys. Lett. 1974, 26 (4), 489-490.

(40) Smith, D. F.; Kiss, A.; Leach, F. E.; Robinson, E. W.; Paša-Tolić, L.; Heeren, R. M. A. High mass accuracy and high mass resolving power FT-ICR secondary ion mass spectrometry for biological tissue imaging. Anal. Bioanal. Chem. 2013, 405 (18), 6069-6076.

(41) Barrow, M. P.; Burkitt, W. I.; Derrick, P. J. Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. Analyst 2005, 130 (1), 18-28.

(42) Lam, Y. P. Y.; Chiu, C. K. C.; Wootton, C. A.; Hands-Portman, I.; Li, M.; Barrow, M. P.; O'Connor, P. B. Does deamidation affect inhibitory mechanisms towards amyloid protein aggregation? Chem. Commun. 2020, 56 (68), 9787-9790.

(43) Lam, Y. P. Y.; Wootton, C. A.; Hands-Portman, I.; Wei, J.; Chiu, C. K. C.; Romero-Canelon, I.; Lermyte, F.; Barrow, M. P.; O'Connor, P. B. Does deamidation of islet amyloid polypeptide accelerate amyloid fibril formation? Chem. Commun. 2018, 54 (98), 13853-13856.

(44) Wootton, C. A.; Sanchez-Cano, C.; Lopez-Clavijo, A. F.; Shaili, E.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Sequence-dependent attack on peptides by photoactivated platinum anticancer complexes. Chem. Sci. 2018, 9 (10), 2733-2739.

(45) Jebanathirajah, J. A.; Pittman, J. L.; Thomson, B. A.; Budnik, B. A.; Kaur, P.; Rape, M.; Kirschner, M.; Costello, C. E.; O'Connor, P. B. Characterization of a New qQq-FT-ICR Mass Spectrometer for Post-Translational Modification Analysis and Top-Down Tandem Mass Spectrometry of Whole Proteins. J. Am. Soc. Mass Spectrom. 2005, 16 (12), 1985-1999.

(46) Cho, Y.; Ahmed, A.; Islam, A.; Kim, S. Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics. Mass Spectrom. Rev. 2015, 34 (2), 248-263.

(47) Palacio Lozano, D. C.; Thomas, M. J.; Jones, H. E.; Barrow, M. P. Petroleomics: Tools, Challenges, and Developments. Annu Rev Anal Chem 2020, 13 (1), 405-430.

(48) Ghaste, M.; Mistrik, R.; Shulaev, V. Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics. Int. J. Mol. Sci. 2016, 17 (6), 816.

(49) Maia, M.; Monteiro, F.; Sebastiana, M.; Marques, A. P.; Ferreira, A. E. N.; Freire, A. P.; Cordeiro, C.; Figueiredo, A.; Sousa Silva, M. Metabolite extraction for high-throughput FT-ICR-MS-based metabolomics of grapevine leaves. EuPA Open Proteomics 2016, 12, 4-9.

(50) Laremore, T. N.; Leach, F. E., 3rd; Solakyildirim, K.; Amster, I. J.; Linhardt, R. J. Glycosaminoglycan characterization by electrospray ionization mass spectrometry including fourier transform mass spectrometry. Methods Enzymol. 2010, 478, 79-108.

(51) Marshall Alan, G.; Hendrickson Christopher, L.; Jackson George, S. Fourier transform ion cyclotron resonance mass spectrometry: A primer. Mass Spectrom. Rev. 1998, 17 (1), 1-35.

(52) Marshall, A. G.; Hendrickson, C. L. Fourier transform ion cyclotron resonance detection: principles and experimental configurations. Int. J. Mass spectrom. 2002, 215 (1), 59-75.

(53) Amster, I. J. Fourier Transform Mass Spectrometry. J. Mass Spectrom. 1998, 31 (12), 1325-1337.

(54) Marshall, A. G.; Roe, D. C. Theory of Fourier transform ion cyclotron resonance mass spectroscopy: Response to frequency-sweep excitation. J. Chem. Phys. 1980, 73 (4), 1581-1590.

(55) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. Tailored excitation for Fourier transform ion cyclotron mass spectrometry. J. Am. Chem. Soc. 1985, 107 (26), 7893-7897.

(56) Nikolaev, E.; Lioznov, A. Evaluation of major historical ICR cell designs using electric field simulations. Mass Spectrom. Rev. 2020, n/a (n/a).

(57) Marshall, A. G. Milestones in fourier transform ion cyclotron resonance mass spectrometry technique development. Int. J. Mass spectrom. 2000, 200 (1), 331-356.

(58) Caravatti, P.; Allemann, M. The 'infinity cell': A new trapped-ion cell with radiofrequency covered trapping electrodes for fourier transform ion cyclotron resonance mass spectrometry. Org. Mass Spectrom. 1991, 26 (5), 514-518.

(59) Beu, S. C.; Laude, D. A. Elimination of axial ejection during excitation with a capacitively coupled open trapped-ion cell for Fourier transform ion cyclotron resonance mass spectrometry. Anal. Chem. 1992, 64 (2), 177-180.

(60) Gabrielse, G. The true cyclotron frequency for particles and ions in a Penning trap. Int. J. Mass spectrom. 2009, 279 (2), 107-112.

(61) Roepstorff, P.; Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed. Mass Spectrom. 1984, 11 (11), 601.

(62) Jennings, K. R. Collision-induced decompositions of aromatic molecular ions. International Journal of Mass Spectrometry and Ion Physics 1968, 1 (3), 227-235.

(63) McLafferty, F. W. Collisional Activation Mass Spectra. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 1979, 293 (1400), 93-102.

(64) Seidler, J.; Zinn, N.; Boehm, M. E.; Lehmann, W. D. De novo sequencing of peptides by MS/MS. Proteomics 2010, 10 (4), 634-649.

(65) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. Anal. Chem. 1994, 66 (18), 2809-2815.

(66) Goolsby, B. J.; Brodbelt, J. S. Tandem Infrared Multiphoton Dissociation and Collisionally Activated Dissociation Techniques in a Quadrupole Ion Trap. Anal. Chem. 2001, 73 (6), 1270-1276.

(67) Gabryelski, W.; Li, L. Photoinduced dissociation of electrospray-generated ions in an ion trap/time-of-flight mass spectrometer using a pulsed CO2 laser. Rapid Commun. Mass Spectrom. 2002, 16 (19), 1805-1811.

(68) Hashimoto, Y.; Hasegawa, H.; Waki, I. High sensitivity and broad dynamic range infrared multiphoton dissociation for a quadrupole ion trap. Rapid Commun. Mass Spectrom. 2004, 18 (19), 2255-2259.

(69) Keller, K. M.; Brodbelt, J. S. Collisionally activated dissociation and infrared multiphoton dissociation of oligonucleotides in a quadrupole ion trap. Anal. Biochem. 2004, 326 (2), 200-210.

(70) Brodbelt, J. S.; Wilson, J. J. Infrared multiphoton dissociation in quadrupole ion traps. Mass Spectrom. Rev. 2009, 28 (3), 390-424.

(71) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. Implementing Photodissociation in an Orbitrap Mass Spectrometer. J. Am. Soc. Mass Spectrom. 2011, 22 (6), 1105-1108.

(72) Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. Identification of Intact Proteins in Mixtures by Alternated Capillary Liquid Chromatography Electrospray Ionization and LC ESI Infrared Multiphoton Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Anal. Chem. 1999, 71 (19), 4397-4402.

(73) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions. Rev. Sci. Instrum. 2005, 76 (2), 023103.

(74) Lancaster, K. S.; An, H. J.; Li, B.; Lebrilla, C. B. Interrogation of N-Linked Oligosaccharides Using Infrared Multiphoton Dissociation in FT-ICR Mass Spectrometry. Anal. Chem. 2006, 78 (14), 4990-4997.

(75) Schäfer, M.; Schmuck, C.; Geiger, L.; Chalmers, M. J.; Hendrickson, C. L.; Marshall, A. G. Structurally related non-covalent complexes examined by quadrupole ion trap (QIT) MS2 and infrared multiphoton dissociation Fourier transform ion cyclotron resonance mass spectrometry IRMPD-FT-ICR MS: evidence for salt-bridge structures in the gas phase. Int. J. Mass spectrom. 2004, 237 (1), 33-45.

(76) Marzullo, B. P.; Morgan, T. E.; Wootton, C. A.; Perry, S. J.; Saeed, M.; Barrow, M. P.; O'Connor, P. B. Advantages of Two-Dimensional Electron-Induced Dissociation and Infrared Multiphoton Dissociation Mass Spectrometry for the Analysis of Agrochemicals. Anal. Chem. 2020, 92 (17), 11687-11695.

(77) Wei, J.; O'Connor, P. B. Extensive fragmentation of pheophytin-a by infrared multiphoton dissociation tandem mass spectrometry. Rapid Commun. Mass Spectrom. 2015, 29 (24), 2411-2418.

(78) Stannard, P. R.; Gelbart, W. M. Intramolecular vibrational energy redistribution. J. Phys. Chem. 1981, 85 (24), 3592-3599.

(79) R Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. J. Am. Soc. Mass Spectrom. 2017, 28 (9), 1823-1826.

(80) Brodbelt, J. S. Photodissociation mass spectrometry: new tools for characterization of biological molecules. Chem. Soc. Rev. 2014, 43 (8), 2757-2783.

(81) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. Chem. Rev. 2020, 120 (7), 3328-3380.

(82) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. Ultrafast Ultraviolet Photodissociation at 193 nm and its Applicability to Proteomic Workflows. J. Proteome Res. 2010, 9 (8), 4205-4214.

(83) Shaw, J. B.; Madsen, J. A.; Xu, H.; Brodbelt, J. S. Systematic Comparison of Ultraviolet Photodissociation and Electron Transfer Dissociation for Peptide Anion Characterization. J. Am. Soc. Mass Spectrom. 2012, 23 (10), 1707-1715.

(84) Fornelli, L.; Srzentić, K.; Toby, T. K.; Doubleday, P. F.; Huguet, R.; Mullen, C.; Melani, R. D.; dos Santos Seckler, H.; DeHart, C. J.; Weisbrod, C. R.; Durbin, K. R.; Greer, J. B.; Early, B. P.; Fellers, R. T.; Zabrouskov, V.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L. Thorough Performance Evaluation of 213 nm Ultraviolet Photodissociation for Top-down Proteomics. Mol. Cell. Proteomics 2020, 19 (2), 405.

(85) Greisch, J.-F.; Tamara, S.; Scheltema, R. A.; Maxwell, H. W. R.; Fagerlund, R. D.; Fineran, P. C.; Tetter, S.; Hilvert, D.; Heck, A. J. R. Expanding the mass range for UVPD-based native top-down mass spectrometry. Chem. Sci. 2019, 10 (30), 7163-7171.

(86) Morrison, L. J.; Brodbelt, J. S. Charge site assignment in native proteins by ultraviolet photodissociation (UVPD) mass spectrometry. Analyst 2016, 141 (1), 166-176.

(87) Klein, D. R.; Brodbelt, J. S. Structural Characterization of Phosphatidylcholines Using 193 nm Ultraviolet Photodissociation Mass Spectrometry. Anal. Chem. 2017, 89 (3), 1516-1522.

(88) Macias, L. A.; Feider, C. L.; Eberlin, L. S.; Brodbelt, J. S. Hybrid 193 nm Ultraviolet Photodissociation Mass Spectrometry Localizes Cardiolipin Unsaturations. Anal. Chem. 2019, 91 (19), 12509-12516.

(89) Williams, P. E.; Klein, D. R.; Greer, S. M.; Brodbelt, J. S. Pinpointing Double Bond and sn-Positions in Glycerophospholipids via Hybrid 193 nm Ultraviolet Photodissociation (UVPD) Mass Spectrometry. J. Am. Chem. Soc. 2017, 139 (44), 15681-15690.

(90) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. J. Am. Chem. Soc. 1998, 120 (13), 3265-3266.

(91) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. Electron capture dissociation of multiply charged peptide cations. Int. J. Mass spectrom. 1999, 185-187, 787-793.

(92) Sawicka, A.; Skurski, P.; Hudgins, R. R.; Simons, J. Model Calculations Relevant to Disulfide Bond Cleavage via Electron Capture Influenced by Positively Charged Groups. J. Phys. Chem. B 2003, 107 (48), 13505-13511.

(93) Syrstad, E. A.; Tureček, F. Toward a general mechanism of electron capture dissociation. J. Am. Soc. Mass Spectrom. 2005, 16 (2), 208-224.

(94) Leymarie, N.; Costello, C. E.; O'Connor, P. B. Electron Capture Dissociation Initiates a Free Radical Reaction Cascade. J. Am. Chem. Soc. 2003, 125 (29), 8949-8958.

(95) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. Anal. Chem. 2000, 72 (3), 563-573.

(96) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W. Top Down Characterization of Larger Proteins (45 kDa) by Electron Capture Dissociation Mass Spectrometry. J. Am. Chem. Soc. 2002, 124 (4), 672-678.

(97) Cooper, H. J.; Håkansson, K.; Marshall, A. G. The role of electron capture dissociation in biomolecular analysis. Mass Spectrom. Rev. 2005, 24 (2), 201-222.

(98) Li, H.; O'Connor, P. B. Electron Capture Dissociation of Disulfide, Sulfur–Selenium, and Diselenide Bound Peptides. J. Am. Soc. Mass Spectrom. 2012, 23 (11), 2001-2010.

(99) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation of Gaseous Multiply-Charged Proteins Is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity. J. Am. Chem. Soc. 1999, 121 (12), 2857-2862.

(100) Robinson, A. B.; Rudd, C. J. In Curr. Top. Cell. Regul., Horecker, B. L.; Stadtman, E. R., Ed.^Eds.; Academic Press, 1974, p^pp 247-295.

(101) Cournoyer, J. J.; Lin, C.; O'Connor, P. B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. Anal. Chem. 2006, 78 (4), 1264-1271.

(102) Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. Protein Sci. 2005, 14 (2), 452-463.

(103) O'Connor, P. B.; Cournoyer, J. J.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. Differentiation of Aspartic and Isoaspartic Acids Using Electron Transfer Dissociation. J. Am. Soc. Mass Spectrom. 2006, 17 (1), 15-19.

(104) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. Identification of aspartic and isoaspartic acid residues in amyloid beta peptides, including Abeta1-42, using electron-ion reactions. Anal. Chem. 2009, 81 (23), 9778-9786.

(105) Li, X.; Lin, C.; O'Connor, P. B. Glutamine deamidation: Differentiation of Glutamic acid and γ-Glutamic acid in peptides by electron capture dissociation. Anal. Chem. 2010, 82 (9), 3606-3615.

(106) Haselmann, K. F.; Budnik, B. A.; Zubarev, R. A. Electron capture dissociation of b 2+ peptide fragments reveals the presence of the acylium ion structure. Rapid Commun. Mass Spectrom. 2000, 14 (23), 2242-2246.

(107) Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. Electron Ionization Dissociation of Singly and Multiply Charged Peptides. J. Am. Chem. Soc. 2009, 131 (29), 9977-9985.

(108) Cody, R. B.; Freiser, B. S. Electron impact excitation of ions from organics: an alternative to collision induced dissociation. Anal. Chem. 1979, 51 (4), 547-551.

(109) Jones, J. W.; Thompson, C. J.; Carter, C. L.; Kane, M. A. Electron-induced dissociation (EID) for structure characterization of glycerophosphatidylcholine: determination of double-bond positions and localization of acyl chains. J. Mass Spectrom. 2015, 50 (12), 1327-1339.

(110) Lopez-Clavijo, A. F.; Griffiths, R. L.; Goodwin, R. J. A.; Cooper, H. J. Liquid Extraction Surface Analysis (LESA) Electron-Induced Dissociation and Collision-Induced Dissociation Mass Spectrometry of Small Molecule Drug Compounds. J. Am. Soc. Mass Spectrom. 2018, 29 (11), 2218-2226.

(111) Mosely, J. A.; Smith, M. J. P.; Prakash, A. S.; Sims, M.; Bristow, A. W. T. Electron-Induced Dissociation of Singly Charged Organic Cations as a Tool for Structural Characterization of Pharmaceutical Type Molecules. Anal. Chem. 2011, 83 (11), 4068-4075.

(112) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl. Acad. Sci. U.S.A. 2004, 101 (26), 9528.

(113) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. Dissociative capture of hot (3–13 eV) electrons by polypeptide polycations: an efficient

process accompanied by secondary fragmentation. Chem. Phys. Lett. 2002, 356 (3), 201-206.

(114) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron detachment dissociation of peptide di-anions: an electron–hole recombination phenomenon. Chem. Phys. Lett. 2001, 342 (3), 299-302.

(115) Ganisl, B.; Valovka, T.; Hartl, M.; Taucher, M.; Bister, K.; Breuker, K. Electron Detachment Dissociation for Top-Down Mass Spectrometry of Acidic Proteins. Chem. Eur. J. 2011, 17 (16), 4460-4469.

(116) Song, H.; Håkansson, K. Electron Detachment Dissociation and Negative Ion Infrared Multiphoton Dissociation of Electrosprayed Intact Proteins. Anal. Chem. 2012, 84 (2), 871-876.

(117) Adamson, J. T.; Håkansson, K. Electron detachment dissociation of neutral and sialylated oligosaccharides. J. Am. Soc. Mass Spectrom. 2007, 18 (12), 2162-2172.

(118) Kornacki, J. R.; Adamson, J. T.; Håkansson, K. Electron Detachment Dissociation of Underivatized Chloride-Adducted Oligosaccharides. J. Am. Soc. Mass Spectrom. 2012, 23 (11), 2031-2042.

(119) Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. Electron Detachment Dissociation of Glycosaminoglycan Tetrasaccharides. J. Am. Soc. Mass Spectrom. 2007, 18 (2), 234-244.

(120) Antoine, R.; Lemoine, J.; Dugourd, P. Electron photodetachment dissociation for structural characterization of synthetic and bio-polymer anions. Mass Spectrom. Rev. 2014, 33 (6), 501-522.

(121) Guan, Z.; Kelleher, N. L.; O'Connor, P. B.; Aaserud, D. J.; Little, D. P.; McLafferty, F. W. 193 nm photodissociation of larger multiply-charged biomolecules. Int. J. Mass Spectrom. Ion Process. 1996, 157-158, 357-364.

(122) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. Cα□C Backbone Fragmentation Dominates in Electron Detachment Dissociation of Gas-Phase Polypeptide Polyanions. Chem. Eur. J. 2005, 11 (6), 1803-1812.

(123) Robinson, N. E.; Robinson, A. B. Molecular clocks. 2001, 98 (3), 944-949.

(124) Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.; Zanni, M. T. Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. J. Am. Chem. Soc. 2012, 134 (30), 12658-12667.

(125) Robinson, N. E.; Robinson, M. L.; Schulze, S. E. S.; Lai, B. T.; Gray, H. B. Deamidation of α-synuclein. Protein Sci. 2009, 18 (8), 1766-1773.

(126) Hains, P. G.; Truscott, R. J. W. Age-dependent deamidation of lifelong proteins in the human lens. Invest. Ophthalmol. Vis. Sci. 2010, 51 (6), 3107-3114.

(127) Michiel, M.; Duprat, E.; Skouri-Panet, F.; Lampi, J. A.; Tardieu, A.; Lampi, K. J.; Finet, S. Aggregation of deamidated human βB2-crystallin and incomplete rescue by α-crystallin chaperone. Exp. Eye Res. 2010, 90 (6), 688-698.

(128) Takata, T.; Oxford, J. T.; Demeler, B.; Lampi, K. J. Deamidation destabilizes and triggers aggregation of a lens protein, βA3-crystallin. Protein Sci. 2008, 17 (9), 1565-1575.

(129) Geiger, T.; Clarke, S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J. Biol. Chem. 1987, 262 (2), 785-794.

(130) Clarke, S. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. Int. J. Pept. Protein Res. 1987, 30 (6), 808-821.

(131) Robinson, N. E.; Robinson, A. B. Deamidation of human proteins. Proc. Natl. Acad. Sci. U.S.A. 2001, 98 (22), 12409.

(132) Wright, H. T. Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit. Rev. Biochem. Mol. Biol. 1991, 26 (1), 1-52.

(133) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. Kinetics and mechanism of succinimide ring formation in the deamidation process of asparagine residues. J. Chem. Soc., Perkin trans. 2 1993, (4), 679-682.

(134) Pace, A. L.; Wong, R. L.; Zhang, Y. T.; Kao, Y.-H.; Wang, Y. J. Asparagine Deamidation Dependence on Buffer Type, pH, and Temperature. J. Pharm. Sci. 2013, 102 (6), 1712-1723.

(135) Bhanuramanand, K.; Ahmad, S.; Rao, N. M. Engineering deamidation-susceptible asparagines leads to improved stability to thermal cycling in a lipase. Protein Sci. 2014, 23 (10), 1479-1490.

(136) Wakankar, A. A.; Borchardt, R. T. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. J. Pharm. Sci. 2006, 95 (11), 2321-2336.

(137) Capasso, S. Thermodynamic parameters of the reversible isomerization of aspartic residues via a succinimide derivative. Thermochim. Acta 1996, 286 (1), 41-50.

(138) Gary, J. D.; Clarke, S. Purification and Characterization of an Isoaspartyl Dipeptidase from Escherichia coli. J. Biol. Chem. 1995, 270 (8), 4076-4087.

(139) Sandmeier, E.; Hunziker, P.; Kunz, B.; Sack, R.; Christen, P. Spontaneous Deamidation and Isomerization of Asn108 in Prion Peptide 106–126 and in Full-Length Prion Protein. Biochem. Biophys. Res. Commun. 1999, 261 (3), 578-583.

(140) Vlasak, J.; Bussat, M. C.; Wang, S.; Wagner-Rousset, E.; Schaefer, M.; Klinguer-Hamour, C.; Kirchmeier, M.; Corvaïa, N.; Ionescu, R.; Beck, A. Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. Anal. Biochem. 2009, 392 (2), 145-154.

(141) Kameoka, D.; Ueda, T.; Imoto, T. A Method for the Detection of Asparagine Deamidation and Aspartate Isomerization of Proteins by MALDI/TOF–Mass Spectrometry Using Endoproteinase Asp-N. J. Biochem. 2003, 134 (1), 129-135.

(142) Aswad, D. W. Stoichiometric methylation of porcine adrenocorticotropin by protein carboxyl methyltransferase requires deamidation of asparagine 25. Evidence for methylation at the alpha-carboxyl group of atypical L-isoaspartyl residues. J. Biol. Chem. 1984, 259 (17), 10714-10721.

(143) Shimizu, T.; Watanabe, A.; Ogawara, M.; Mori, H.; Shirasawa, T. Isoaspartate Formation and Neurodegeneration in Alzheimer's Disease. Arch. Biochem. Biophys. 2000, 381 (2), 225-234.

(144) Johnson, B. A.; Aswad, D. W. Optimal conditions for the use of protein l-isoaspartyl methyltransferase in assessing the isoaspartate content of peptides and proteins. Anal. Biochem. 1991, 192 (2), 384-391.

(145) Chazin, W. J.; Koerdel, J.; Thulin, E.; Hofmann, T.; Drakenberg, T.; Forsen, S. Identification of an isoaspartyl linkage formed upon deamidation of bovine calbindin D9k and structural characterization by 2D proton NMR. Biochemistry 1989, 28 (21), 8646-8653.

(146) Takata, T.; Woodbury, L. G.; Lampi, K. J. Deamidation alters interactions of beta-crystallins in hetero-oligomers. Mol. Vis. 2009, 15, 241-249.

(147) Karlsson, G.; Gellerfors, P.; Persson, A.; Norén, B.; Edlund, P. O.; Sandberg, C.; Birnbaum, S. Separation of oxidized and deamidated human growth hormone variants by isocratic reversed-phase high-performance liquid chromatography. J. Chromatogr. A 1999, 855 (1), 147-155.

(148) Carlson, A. D.; Riggin, R. M. Development of Improved High-Performance Liquid Chromatography Conditions for Nonisotopic Detection of Isoaspartic Acid to Determine the Extent of Protein Deamidation. Anal. Biochem. 2000, 278 (2), 150-155.

(149) Badgett, M. J.; Boyes, B.; Orlando, R. The Separation and Quantitation of Peptides with and without Oxidation of Methionine and Deamidation of Asparagine Using Hydrophilic Interaction Liquid Chromatography with Mass Spectrometry (HILIC-MS). J. Am. Soc. Mass Spectrom. 2017, 28 (5), 818-826.

(150) Robinson, N. E.; Zabrouskov, V.; Zhang, J.; Lampi, K. J.; Robinson, A. B. Measurement of deamidation of intact proteins by isotopic envelope and mass defect with ion cyclotron resonance Fourier transform mass spectrometry. 2006, 20 (23), 3535-3541.

(151) Perez Hurtado, P.; O'Connor, P. B. Deamidation of Collagen. Anal. Chem. 2012, 84 (6), 3017-3025.

(152) Jin, Y.; Yi, Y.; Yeung, B. Mass spectrometric analysis of protein deamidation – A focus on top-down and middle-down mass spectrometry. Methods 2020.

(153) Lehmann, W. D.; Schlosser, A.; Erben, G.; Pipkorn, R.; Bossemeyer, D.; Kinzel, V. Analysis of isoaspartate in peptides by electrospray tandem mass spectrometry. Protein Sci. 2000, 9 (11), 2260-2268.

(154) Gonzalez, L. J.; Shimizu, T.; Satomi, Y.; Betancourt, L.; Besada, V.; Padron, G.; Orlando, R.; Shirasawa, T.; Shimonishi, Y.; Takao, T. Differentiating alpha- and beta-aspartic acids by electrospray ionization and low-energy tandem mass spectrometry. Rapid Commun. Mass Spectrom. 2000, 14 (22), 2092-2102.

(155) Clarke, S. Protein methylation. Curr. Opin. Cell Biol. 1993, 5 (6), 977-983.

(156) Lee, D. Y.; Teyssier, C.; Strahl, B. D.; Stallcup, M. R. Role of Protein Methylation in Regulation of Transcription. Endocr. Rev. 2005, 26 (2), 147-170.

(157) Alam, H.; Gu, B.; Lee, M. G. Histone methylation modifiers in cellular signaling pathways. Cell. Mol. Life Sci. 2015, 72 (23), 4577-4592.

(158) Wei, W.; Ji, X.; Guo, X.; Ji, S. Regulatory Role of N6-methyladenosine (m6A) Methylation in RNA Processing and Human Diseases. J. Cell. Biochem. 2017, 118 (9), 2534-2543.

(159) Clarke, S. Aging as war between chemical and biochemical processes: Protein methylation and the recognition of age-damaged proteins for repair. Ageing Res. Rev. 2003, 2 (3), 263-285.

(160) Shimizu, T.; Matsuoka, Y.; Shirasawa, T. Biological Significance of Isoaspartate and Its Repair System. Biol. Pharm. Bull. 2005, 28 (9), 1590-1596.

(161) Bedford, M. T.; Frankel, A.; Yaffe, M. B.; Clarke, S.; Leder, P.; Richard, S. Arginine Methylation Inhibits the Binding of Proline-rich Ligands to Src Homology 3, but Not WW, Domains *. J. Biol. Chem. 2000, 275 (21), 16030-16036.

(162) Han, D.; Huang, M.; Wang, T.; Li, Z.; Chen, Y.; Liu, C.; Lei, Z.; Chu, X. Lysine methylation of transcription factors in cancer. Cell Death Dis. 2019, 10 (4), 290-290.

(163) Johnson, P.; Harris, C. I.; Perry, S. V. 3-methylhistidine in actin and other muscle proteins. Biochem. J. 1967, 105 (1), 361-370.

(164) Huszar, G.; Elzinga, M. Amino acid sequence around the single 3-methylhistidine residue in rabbit skeletal muscle myosin. Biochemistry 1971, 10 (2), 229-236.

(165) Huszar, G.; Elzinga, M. Homologous Methylated and Nonmethylated Histidine Peptides in Skeletal and Cardiac Myosins. J. Biol. Chem. 1972, 247 (3), 745-753.

(166) Asatoor, A. M.; Armstrong, M. D. 3-Methylhistidine, a component of actin. Biochem. Biophys. Res. Commun. 1967, 26 (2), 168-174.

(167) Trayer, I. P.; Harris, C. I.; Perry, S. V. 3-Methyl Histidine and Adult and Foetal Forms of Skeletal Muscle Myosin. Nature 1968, 217 (5127), 452-453.

(168) Ballard, F. J.; Tomas, F. M. 3-Methylhistidine as a Measure of Skeletal Muscle Protein Breakdown in Human Subjects: The Case for Its Continued Use. Clin. Sci. 1983, 65 (3), 209-215.

(169) Young, V. R.; Munro, H. N. Ntau-methylhistidine (3-methylhistidine) and muscle protein turnover: an overview. Fed. Proc. 1978, 37 (9), 2291-2300.

(170) Butt, J. H.; Fleshler, B. Anserine, A Source of 1-Methylhistidine in Urine of Man. Proc. Soc. Exp. Biol. Med. 1965, 118 (3), 722-725.

(171) Yao, X.; Grade, S.; Wriggers, W.; Rubenstein, P. A. His(73), often methylated, is an important structural determinant for actin. A mutagenic analysis of HIS(73) of yeast actin. J. Biol. Chem. 1999, 274 (52), 37443-37449.

(172) Nyman, T.; Schüler, H.; Korenbaum, E.; Schutt, C. E.; Karlsson, R.; Lindberg, U. The role of MeH73 in actin polymerization and ATP hydrolysis11Edited by R. Huber. J. Mol. Biol. 2002, 317 (4), 577-589.

(173) Guo, Q.; Liao, S.; Kwiatkowski, S.; Tomaka, W.; Yu, H.; Wu, G.; Tu, X.; Min, J.; Drozak, J.; Xu, C. Structural insights into SETD3-mediated histidine methylation on β-actin. eLife 2019, 8, e43676.

(174) Kwiatkowski, S.; Seliga, A. K.; Vertommen, D.; Terreri, M.; Ishikawa, T.; Grabowska, I.; Tiebe, M.; Teleman, A. A.; Jagielski, A. K.; Veiga-da-Cunha, M.; Drozak, J. SETD3 protein is the actin-specific histidine N-methyltransferase. eLife 2018, 7, e37921.

(175) Shu, W.-J.; Du, H.-N. The methyltransferase SETD3-mediated histidine methylation: Biological functions and potential implications in cancers. Biochim. Biophys. Acta 2021, 1875 (1), 188465.

(176) Wilkinson, A. W.; Diep, J.; Dai, S.; Liu, S.; Ooi, Y. S.; Song, D.; Li, T. M.; Horton, J. R.; Zhang, X.; Liu, C.; Trivedi, D. V.; Ruppel, K. M.; Vilches-Moure, J. G.; Casey, K.

M.; Mak, J.; Cowan, T.; Elias, J. E.; Nagamine, C. M.; Spudich, J. A.; Cheng, X., *et. al.* SETD3 is an actin histidine methyltransferase that prevents primary dystocia. Nature 2019, 565 (7739), 372-376.

(177) Zheng, Y.; Zhang, X.; Li, H. Molecular basis for histidine N3-specific methylation of actin H73 by SETD3. Cell Discov. 2020, 6 (1), 3.

(178) Rathmacher, J. A.; Link, G. A.; Flakoll, P. J.; Nissen, S. L. Gas chromatographic/mass spectrometric analysis of stable isotopes of 3-methylhistidine in biological fluids: Application to plasma kinetics in vivo. Biol. Mass Spectrom. 1992, 21 (11), 560-566.

(179) Ali Qureshi, G.; Van den berg, S.; Gutierrez, A.; Bergström, J. Determination of histidine and 3-methylhistidine in physiological fluids by high-performance liquid chromatography. J. Chromatogr. A 1984, 297, 83-89.

(180) Houweling, M.; van der Drift, S. G. A.; Jorritsma, R.; Tielens, A. G. M. Quantification of plasma 1- and 3-methylhistidine in dairy cows by high-performance liquid chromatography- tandem mass spectrometry. J. Dairy Sci. 2012, 95 (6), 3125-3130.

(181) Holick, M. F.; Chen, T. C. Vitamin D deficiency: a worldwide problem with health consequences. Am. J. Clin. Nutr. 2008, 87 (4), 1080S-1086S.

(182) Holick, M. F. Vitamin D Deficiency. N. Engl. J. Med. 2007, 357 (3), 266-281.

(183) Holick, M. F. Resurrection of vitamin D deficiency and rickets. J. Clin. Invest. 2006, 116 (8), 2062-2072.

(184) Altieri, B.; Cavalier, E.; Bhattoa, H. P.; Pérez-López, F. R.; López-Baena, M. T.; Pérez-Roncero, G. R.; Chedraui, P.; Annweiler, C.; Della Casa, S.; Zelzer, S.; Herrmann, M.; Faggiano, A.; Colao, A.; Holick, M. F. Vitamin D testing: advantages and limits of the current assays. Eur. J. Clin. Nutr. 2020, 74 (2), 231-247.

(185) Cavalier, E.; Lukas, P.; Crine, Y.; Peeters, S.; Carlisi, A.; Le Goff, C.; Gadisseur, R.; Delanaye, P.; Souberbielle, J.-C. Evaluation of automated immunoassays for 25(OH)-vitamin D determination in different critical populations before and after standardization of the assays. Clin. Chim. Acta 2014, 431, 60-65.

(186) Haddad, J. G., Jr.; Walgate, J. Radioimmunoassay of the binding protein for vitamin D and its metabolites in human serum: concentrations in normal subjects and patients with disorders of mineral homeostasis. J. Clin. Invest. 1976, 58 (5), 1217-1222.

(187) Li, L.; Zeng, Q.; Yuan, J.; Xie, Z. Performance evaluation of two immunoassays for 25-hydroxyvitamin D. J. Clin. Biochem. Nutr. 2016, 58 (3), 186-192.

(188) Snellman, G.; Melhus, H.; Gedeborg, R.; Byberg, L.; Berglund, L.; Wernroth, L.; Michaëlsson, K. Determining vitamin D status: a comparison between commercially available assays. PLoS One 2010, 5 (7), e11555-e11555.

(189) Graham, D. C. Accuracy of 25-Hydroxyvitamin D Assays: Confronting the Issues. Curr. Drug Targets 2011, 12 (1), 19-28.

(190) Coldwell, R. D.; Porteous, C. E.; Trafford, D. J. H.; Makin, H. L. J. Gas chromatography—mass spectrometry and the measurement of vitamin D metabolites in human serum or plasma. Steroids 1987, 49 (1), 155-196.

(191) Lehner, A.; Johnson, M.; Zimmerman, A.; Zyskowski, J.; Buchweitz, J. Vitamin D analyses in veterinary feeds by gas chromatography-tandem mass spectrometry. Eur. J. Mass Spectrom. 2021, 27 (1), 48-62.

(192) Ogawa, S.; Ooki, S.; Shinoda, K.; Higashi, T. Analysis of urinary vitamin D3 metabolites by liquid chromatography/tandem mass spectrometry with ESI-enhancing and stable isotope-coded derivatization. Anal. Bioanal. Chem. 2014, 406 (26), 6647-6654.

(193) Qi, Y.; Geib, T.; Schorr, P.; Meier, F.; Volmer, D. A. On the isobaric space of 25-hydroxyvitamin D in human serum: potential for interferences in liquid chromatography/tandem mass spectrometry, systematic errors and accuracy issues. Rapid Commun. Mass Spectrom. 2015, 29 (1), 1-9.

(194) Qi, Y.; Müller, M.; Stokes, C. S.; Volmer, D. A. Rapid Quantification of 25-Hydroxyvitamin D3 in Human Serum by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. J. Am. Soc. Mass Spectrom. 2018, 29 (7), 1456-1462.

(195) Qi, Y.; Müller, M. J.; Volmer, D. A. Activation of Reactive MALDI Adduct Ions Enables Differentiation of Dihydroxylated Vitamin D Isomers. J. Am. Soc. Mass Spectrom. 2017, 28 (12), 2532-2537.

(196) Chouinard, C. D.; Cruzeiro, V. W. D.; Beekman, C. R.; Roitberg, A. E.; Yost, R. A. Investigating Differences in Gas-Phase Conformations of 25-Hydroxyvitamin D3 Sodiated Epimers using Ion Mobility-Mass Spectrometry and Theoretical Modeling. J. Am. Soc. Mass Spectrom. 2017, 28 (8), 1497-1505.

(197) Chouinard, C. D.; Cruzeiro, V. W. D.; Kemperman, R. H. J.; Oranzi, N. R.; Roitberg, A. E.; Yost, R. A. Cation-dependent conformations in 25-hydroxyvitamin D3-cation adducts measured by ion mobility-mass spectrometry and theoretical modeling. Int. J. Mass spectrom. 2018, 432, 1-8.

(198) Holick, M. F. Vitamin D status: measurement, interpretation, and clinical application. Ann. Epidemiol. 2009, 19 (2), 73-78.

(199) Ankney, J. A.; Muneer, A.; Chen, X. Relative and Absolute Quantitation in Mass Spectrometry–Based Proteomics. Annu Rev Anal Chem 2016, 11 (1), 49-77.

(200) Unwin, R. D.; Evans, C. A.; Whetton, A. D. Relative quantification in proteomics: new approaches for biochemistry. Trends Biochem. Sci. 2006, 31 (8), 473-484.

(201) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. 1999, 17 (10), 994-999.

(202) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat. Biotechnol. 2001, 19 (10), 946-951.

(203) Schmidt, A.; Kellermann, J.; Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 2005, 5 (1), 4-15.

(204) Heudi, O.; Barteau, S.; Zimmer, D.; Schmidt, J.; Bill, K.; Lehmann, N.; Bauer, C.; Kretz, O. Towards Absolute Quantification of Therapeutic Monoclonal Antibody in Serum by LC−MS/MS Using Isotope-Labeled Antibody Standard and Protein Cleavage Isotope Dilution Mass Spectrometry. Anal. Chem. 2008, 80 (11), 4200-4207.

(205) Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. Anal. Chem. 2003, 75 (8), 1895-1904.

(206) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. Proteomics 2007, 7 (3), 340-350.

(207) Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. Proteomics 2015, 15 (18), 3175-3192.

(208) Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics *. Mol. Cell. Proteomics 2002, 1 (5), 376-386.

(209) Selleck, M. J.; Senthil, M.; Wall, N. R. Making Meaningful Clinical Use of Biomarkers. Biomark. Insights 2017, 12, 1177271917715236-1177271917715236.

(210) Davis, K. D.; Aghaeepour, N.; Ahn, A. H.; Angst, M. S.; Borsook, D.; Brenton, A.; Burczynski, M. E.; Crean, C.; Edwards, R.; Gaudilliere, B.; Hergenroeder, G. W.; Iadarola, M. J.; Iyengar, S.; Jiang, Y.; Kong, J.-T.; Mackey, S.; Saab, C. Y.; Sang, C. N.; Scholz, J.; Segerdahl, M., *et. al.* Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. Nat. Rev. Neurol. 2020, 16 (7), 381-400.

(211) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. Proc. Natl. Acad. Sci. U.S.A. 2020, 117 (21), 11727.

(212) Benton, D. J.; Wrobel, A. G.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. Nature 2020, 588 (7837), 327-330.

(213) McBride, R.; van Zyl, M.; Fielding, B. C. The coronavirus nucleocapsid is a multifunctional protein. Viruses 2014, 6 (8), 2991-3018.

(214) Siu, Y. L.; Teoh, K. T.; Lo, J.; Chan, C. M.; Kien, F.; Escriou, N.; Tsao, S. W.; Nicholls, J. M.; Altmeyer, R.; Peiris, J. S. M.; Bruzzone, R.; Nal, B. The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. J. Virol. 2008, 82 (22), 11318.

(215) Snijder, E. J.; Decroly, E.; Ziebuhr, J. In Adv. Virus Res., Ziebuhr, J., Ed.^Eds.; Academic Press, 2016, p^pp 59-126.

(216) Yadav, R.; Chaudhary, J. K.; Jain, N.; Chaudhary, P. K.; Khanra, S.; Dhamija, P.; Sharma, A.; Kumar, A.; Handu, S. Role of Structural and Non-Structural Proteins and Therapeutic Targets of SARS-CoV-2 for COVID-19. Cells 2021, 10 (4), 821.

(217) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu. Rev. Virol. 2016, 3 (1), 237-261.

(218) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 2020, 581 (7807), 215-220.

(219) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; Angyal, A.; Brown, R. L.; Carrilero, L.; Green, L. R., *et. al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell 2020, 182 (4), 812-827.e819.

(220) Volz, E.; Hill, V.; McCrone, J. T.; Price, A.; Jorgensen, D.; O'Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F. F.; Rey, S. M.; Nicholls, S. M.; Colquhoun, R. M.; da Silva Filipe, A.; Shepherd, J.; Pascall, D. J.; Shah, R.; Jesudason, N.; Li, K.; Jarrett, R., *et. al.* Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell 2021, 184 (1), 64-75.e11.

(221) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J. E.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P. C.; Kyratsous, C. A.; Dudkina, N. V., *et. al.* Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell 2020, 183 (3), 739-751.e738.

(222) Bangaru, S.; Ozorowski, G.; Turner, H. L.; Antanasijevic, A.; Huang, D.; Wang, X.; Torres, J. L.; Diedrich, J. K.; Tian, J.-H.; Portnoff, A. D.; Patel, N.; Massare, M. J.; Yates, J. R.; Nemazee, D.; Paulson, J. C.; Glenn, G.; Smith, G.; Ward, A. B. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science 2020, 370 (6520), 1089.

(223) Dai, L.; Gao, G. F. Viral targets for vaccines against COVID-19. Nat. Rev. Immunol. 2021, 21 (2), 73-82.

(224) Du, L.; He, Y.; Zhou, Y.; Liu, S.; Zheng, B.-J.; Jiang, S. The spike protein of SARS-CoV — a target for vaccine and therapeutic development. Nat. Rev. Microbiol. 2009, 7 (3), 226-236.

(225) Salvatori, G.; Luberto, L.; Maffei, M.; Aurisicchio, L.; Roscilli, G.; Palombo, F.; Marra, E. SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. J. Transl. Med. 2020, 18 (1), 222.

(226) Tombuloglu, H.; Sabit, H.; Al-Suhaimi, E.; Al Jindan, R.; Alkharsah, K. R. Development of multiplex real-time RT-PCR assay for the detection of SARS-CoV-2. PLoS One 2021, 16 (4), e0250942.

(227) Lisboa Bastos, M.; Tavaziva, G.; Abidi, S. K.; Campbell, J. R.; Haraoui, L.-P.; Johnston, J. C.; Lan, Z.; Law, S.; MacLean, E.; Trajman, A.; Menzies, D.; Benedetti, A.; Ahmad Khan, F. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ 2020, 370, m2516.

(228) MacMullan, M. A.; Ibrayeva, A.; Trettner, K.; Deming, L.; Das, S.; Tran, F.; Moreno, J. R.; Casian, J. G.; Chellamuthu, P.; Kraft, J.; Kozak, K.; Turner, F. E.; Slepnev, V. I.; Le Page, L. M. ELISA detection of SARS-CoV-2 antibodies in saliva. Sci. Rep. 2020, 10 (1), 20818.

(229) Singhal, N.; Kumar, M.; Kanaujia, P. K.; Virdi, J. S. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front. Microbiol. 2015, 6, 791-791.

(230) Cherkaoui, A.; Hibbs, J.; Emonet, S.; Tangomo, M.; Girard, M.; Francois, P.; Schrenzel, J. Comparison of Two Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry Methods with Conventional Phenotypic Identification for Routine Identification of Bacteria to the Species Level. J. Clin. Microbiol. 2010, 48 (4), 1169.

(231) La Scola, B.; Raoult, D. Direct Identification of Bacteria in Positive Blood Culture Bottles by Matrix-Assisted Laser Desorption Ionisation Time-of-Flight Mass Spectrometry. PLoS One 2009, 4 (11), e8041.

(232) Thiede, B.; Höhenwarter, W.; Krah, A.; Mattow, J.; Schmid, M.; Schmidt, F.; Jungblut, P. R. Peptide mass fingerprinting. Methods 2005, 35 (3), 237-247.

(233) Iles, R. K.; Zmuidinaite, R.; Iles, J. K.; Carnell, G.; Sampson, A.; Heeney, J. L. A clinical MALDI-ToF Mass spectrometry assay for SARS-CoV-2: Rational design and multi-disciplinary team work. medRxiv 2020, 2020.2008.2022.20176669.

(234) Ihling, C.; Tänzler, D.; Hagemann, S.; Kehlen, A.; Hüttelmaier, S.; Sinz, A. Mass Spectrometric Identification of SARS-CoV-2 Proteins from Gargle Solution Samples of COVID-19 Patients. bioRxiv 2020, 2020.2004.2018.047878.

(235) Nikolaev, E. N.; Indeykina, M. I.; Brzhozovskiy, A. G.; Bugrova, A. E.; Kononikhin, A. S.; Starodubtseva, N. L.; Petrotchenko, E. V.; Kovalev, G. I.; Borchers, C. H.; Sukhikh, G. T. Mass-Spectrometric Detection of SARS-CoV-2 Virus in Scrapings of the Epithelium of the Nasopharynx of Infected Patients via Nucleocapsid N Protein. J. Proteome Res. 2020, 19 (11), 4393-4397.

2. Differentiation and Relative Quantification of the Isomeric Products of Deamidation using ECD and UVPD Tandem Mass Spectrometry

This chapter demonstrates the applications of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) and the fragmentation methods available on the FT-ICR MS to study the detection and relative quantification of the isomeric products of asparagine deamidation.

Sample preparation, data acquisition and analysis results presented in this chapter were carried out by the thesis author with significant help provided by Dr. Yuko P. Y. Lam including setup, optimisation and running of the nano-LC experiments. Help was also provided by Dr. Christopher A. Wootton and Dr. Alina Theisen for laser alignment and setup of the UVPD experiments.

One manuscript entitled “Differentiation and Relative Quantification of the Isomeric Products of Deamidation using ECD and UVPD Tandem Mass Spectrometry” by Anisha Haris, Yuko P. Y. Lam, Christopher A. Wootton, Alina Theisen, Tomos E. Morgan, Mark P. Barrow, and Peter B. O’Connor, is being prepared for submission to the Journal of Analytical Chemistry based on the results presented in this chapter.

2.1. Abstract

Deamidation is an important post-translational modification (PTM) involved in age-related neurodegenerative diseases. Deamidation of an asparagine residue results in the formation of isomeric products; aspartic acid and isoaspartic acid, which are challenging to distinguish between using various analytical techniques due to the subtle changes in their structures. Electron capture dissociation tandem mass spectrometry (ECD MS/MS) has demonstrated the capability to differentiate and relatively quantify the deamidated isomeric products directly via the generation of a characteristic fragment (z_{i-n}-C_2O_2H); this method, however, is still limited to certain types of MS instrumentation. Herein, we demonstrate the use of ultraviolet photodissociation (UVPD) MS/MS, a laser system that can be easily integrated with various types of MS instrument, to differentiate and quantify the deamidated isomeric products. Furthermore, the quantification method for the deamidated products was improved, so that a good linearity relationship (R^2 > 0.99) was easily obtained for all target peptides. These results are important to promote the general implementation of MS instruments with MS/MS capabilities, which can provide a reliable and robust method to relatively quantify and differentiate between the deamidated isomeric products.

2.2. Introduction

Deamidation is a spontaneous, non-enzymatic yet highly common post-translational modification (PTM) observed in peptides and proteins.[1] It serves as a molecular clock for the regulation of biological processes and can be used as a marker for protein ageing, especially in long-lived proteins such as the α-, β- and γ-crystallin proteins, which are found in eye lens.[2] Over time and with the ageing of the eye lens, the crystallins undergo several modifications including deamidation, which has been shown to increase aggregation resulting in protein insolubilization, eventually leading to the formation of cataracts.[3-5]

Deamidation is also implicated in the progression of a wide range of diseases from neurological disorders such as Alzheimer's and Parkinson's diseases to the formation of senile plaques, as well as amyloid polypeptide aggregation in type II diabetes (T2D).[6-10] Deamidation is known to cause structurally and functionally important alterations in protein structures.[11,12] For example, in therapeutic antibodies, deamidation is one of the PTMs that needs to be carefully monitored as it has been shown to affect the stability and shelf-life of monoclonal antibodies.[13-16] Previous research has shown that the spontaneous formation of the isomeric products of asparagine deamidation in the complementary determining regions of a monoclonal antibody decreases antigen receptor binding and efficacy.[17,18]

All peptides/proteins are susceptible to deamidation if asparagine (Asn) and/or glutamine (Gln) are included in the amino acid sequence. Asn and Gln amino acid residues are prone to deamidation, generating aspartic acid (D)/isoaspartic acid (isoD) and glutamic acid (α-Glu)/isoglutamic acid (γ-Glu) respectively (Figure 2.1).[11,19-21] The amide group ($-NH_2$) at the side chain of Asn and Gln residues is readily converted to a hydroxyl group (-OH) during the process of deamidation; while an extra methylene group ($-CH_2-$) is inserted into the backbone of isoD and γ-Glu residues, making them β-amino acids, which can lead to a dramatic change in the tertiary structure, folding, stability, and the function of a protein.[12] Although the mechanism of deamidation is the same for both residues, the rate of deamidation is faster for Asn than Gln because the distance between the main chain amide group (-NH-) and the side chain amide group ($-NH_2$) is shorter in Asn.[1,22] The ratio of formation of isoD to D for unfolded, random coils of peptides or proteins is approximately 3:1 via deamidation of Asn; ratio deviations, however, can be observed when varying amino acid sequence, protein structure, and experimental buffer

conditions.[19,23-26] Since deamidation in Gln is relatively slower than Asn, the more prevalent modification i.e., Asn deamidation and isoD formation will be the focus of this work.

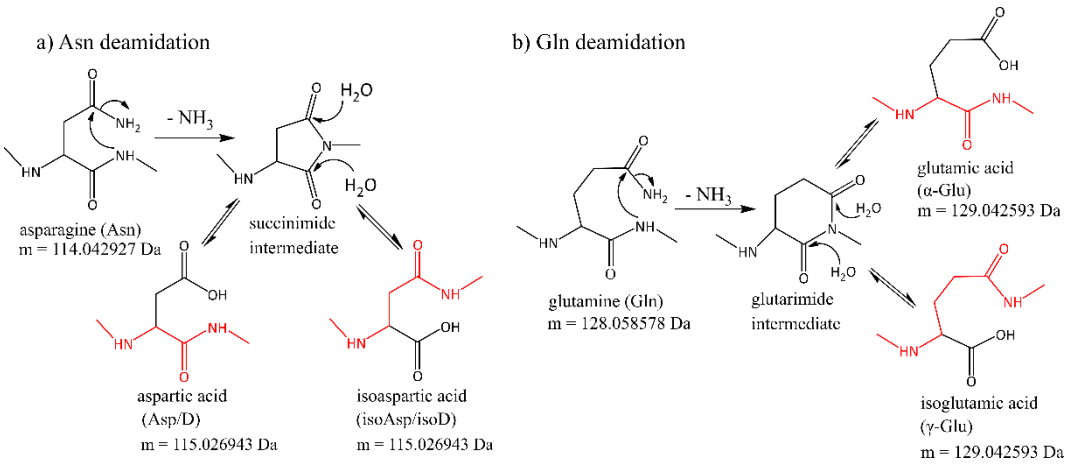


Figure 2. 1 Deamidation mechanisms of a) asparagine and b) glutamine.

For the products of deamidation, the backbone of amino acid residues is coloured in red, while the amino acid side chain groups are coloured in black.

The conversion of $–NH_2$ to –OH group during deamidation results in a mass shift of +0.984 Daltons (Da), which can be easily detected by high-resolution mass spectrometry (HRMS).[7,27,28] Differentiation of the deamidated isomeric products, however, is much more challenging due to the minor structural and zero mass differences between the two products. Currently, the most common methods for detection and quantitation of the deamidated products include liquid chromatography (LC) methods combined with MS such as reverse phase-high performance liquid chromatography (RP-HPLC) and hydrophilic interaction liquid chromatography (HILIC).[15,29-32] In LC-based techniques, particularly in RP-HPLC, the non-deamidated and deamidated peptides are often assigned based on their elution order; previous research has shown that isoD peptide elute earlier than the D peptide in general.[33,34] Using RP-HPLC to differentiate and quantify the isomeric deamidated products requires baseline separation of the variants;[35] furthermore, the assignment based on HPLC elution order may not always be reliable.[33,34,36] Previously, other analytical methods, including nuclear magnetic resonance (NMR) spectroscopy,[37,38] Edman degradation,[39] ^{18}O isotope labelling,[40,41] and affinity enrichment using chemo-enzymatic methods such as protein L-isoaspartyl methyltransferase (PIMT) and hydrazine trapping [42] have also been used to differentiate the isomeric deamidated products but these methods may require high concentration, large quantities, or even modification and

manipulation of the sample to some extent, which may be a big challenge to most of the biological samples.

Tandem mass spectrometry (MS^2 or MS/MS) methods such as collisional activated dissociation (CAD), also referred to as collision induced dissociation (CID), and electron capture dissociation (ECD) have been successfully utilised for the differentiation of isoD and D peptides.[43-47] CAD is the most common fragmentation method available on most if not all MS platforms. In the late 1990's and beginning of the year 2000, various research groups demonstrated the use of CAD for isoD and D peptide differentiation, based on the differences in the abundances of the immonium and b/y ions depending on the peptide sequence (Figure 2.2).[43,44]

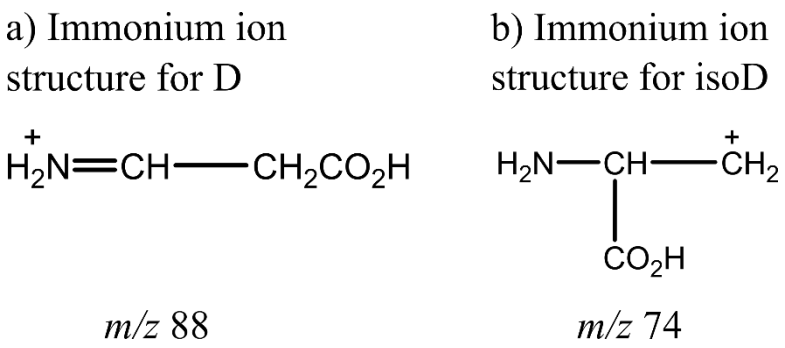
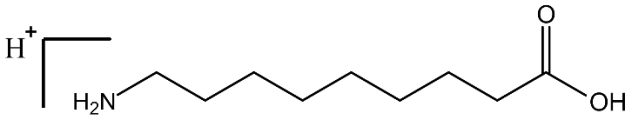


Figure 2. 2 Immonium ions for a) D and b) isoD (exact mass of 88.08545 Da) (adapted and redrawn from Castet *et. al.*)[43]

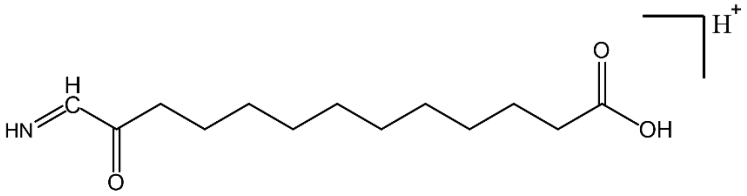
Low energy CAD was used and diagnostic fragment ions $b_{n-1}+H_2O$ and y_{l-n}-46 (loss of CO_2H_2) was observed in the isoD peptide CAD MS/MS spectrum when compared to D peptide (Figure 2.3).[45,48] The characteristic fragments, however, appeared to be low in abundance and due to the peptide sequence dependency, this method may not be suitable for all peptides.[45]

a) Generated from D and isoD amino acids:



$b_{n+1}+H_2O$

b) Genereated from only isoD amino acid:



y_{l-n+1}-46

Figure 2. 3 Results of the proposed internal rearrangements of D and isoD amino acids, generated from a) both D and isoD and b) isoD only. (Adapted and redrawn from Gonzalez *et. al.*)[45]

ECD is an electron-mediated fragmentation method, applied to multiply charged positive ions by submission to a beam of low energy electrons produced by a hollow dispenser cathode.[49,50] Cournoyer *et. al.*[46] found that the cleavage of the C_α - C_β backbone bond by ECD generated diagnostic fragment ions for isoD containing peptides, $c_n\cdot$ + 58 (c_n + C_2O_2H) and z_{i-n}-C_2O_2H (where n is the position of isoD and $_i$ is the total number of amino acids in the peptide), which was absent for D peptides (Figure 2.4).[46,51]

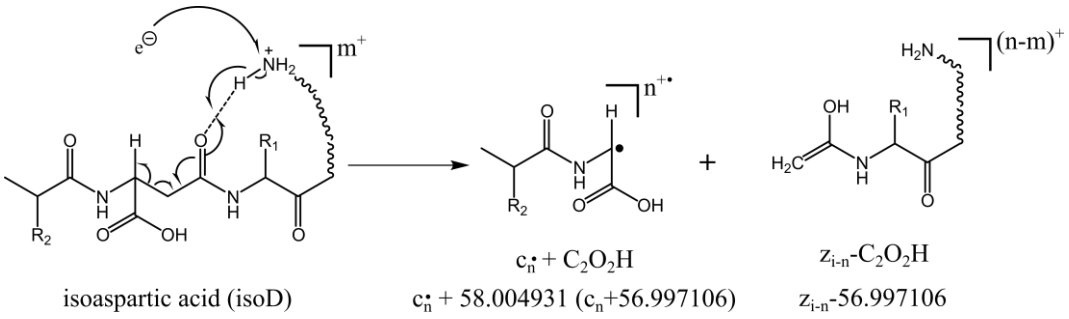


Figure 2. 4 ECD fragmentation mechanism at isoaspartic acid to generate a characteristic (c_n+58 Da) ion and a (z_{i-n}-57 Da) ion.

More recently, ultraviolet photodissociation (UVPD), has been investigated as an alternative fragmentation method as it has been shown to provide high sequence coverage and complementary information to other fragmentation methods for peptides and proteins.[52-55] Julian group has demonstrated that radical directed dissociation (RDD) can be useful in the differentiation of isoD and D peptides and even the l and d-enantiomers of both products.[55] This method involves, modifying the D and isoD peptides with photolabile chromophores and irradiation of the samples with a UV laser at 266 nm to generate a radical species followed by CAD MS/MS, which has shown to result in isomer and enantiomer specific fragments, enabling the differentiation of D and isoD peptides.[55-57] However, - in the current literature, no characteristic fragments for isoD and D peptides have been detected solely using UVPD, without any prior modification to the molecules.

Previous research has shown that the presence of isoD in proteins can accelerate diseases formation; quantification of isoD peptide/protein, therefore, is important for biological studies. Relative quantitation of isoD peptide is possible using the characteristic isoD fragment (z_{i-n}-C_2O_2H) generated by ECD MS/MS. Cournoyer *et. al.*[46] demonstrated a sequence independent linear relationship between the relative abundance of the specific (z_{i-n}-C_2O_2H) isoD fragment and the percentage isoD peptide in a mixture.[58] A calibration curve was generated using the mixture of known concentration, standard mixture of isoD and D peptides, which is applicable to both ECD and CAD MS/MS deamidation quantification. Peak intensity of the characteristic fragments generated by isoD peptide over the sum of all fragment peak intensities were commonly used for isoD peptide quantification.[58] Following the method of the Nikolaev group[59-61] assuming that bond breakage further from the modification site should be less affected by the changes in the isoform concentrations, than those located closer, a ratio was taken of the isoD

dependent fragments to the isoD independent fragments and calibration curves was generated with increasing isoD content in the isoD/D peptide mixtures.

Herein, we first demonstrate an improvement of the isoD relative quantitation methods using ECD MS/MS on mixtures of standard peptides, then further apply the quantification method to UVPD MS/MS on the same standard peptides to demonstrate a similar performance can be achieved using different fragmentation strategy. This method is then used to determine the isoD percentage in deamidated and tryptic digested BSA peptides to compare the performance of ECD and UVPD MS/MS in isomeric product quantification. Isomeric quantification can be generated either from ECD or UVPD MS/MS, indicating that MS instrumentation is no longer a limitation factor to deamidation products quantification.

2.3. Experimental Section

Sample preparation for synthetic peptides

Standard D and isoD peptides group number 1 – 3 (shown in Table 2.1. below) were synthesised by GenScript Biotech Corporation (Netherlands). RP-HPLC was carried out by the supplier to estimate the purity percentage of the synthetic peptides based on UV absorbance at a wavelength of 220 nm (Table 2.1).

Table 2. 1 Summary of synthetic peptides. Variable residue was underlined in the table.

Peptide group number	Peptide sequence	Peptide purity (%)
1	LVNELTEFAK	95.1
	LVDELTEFAK	99.8
	LV(isoD)ELTEFAK	98.5
2	TVMEDFVAFDK	97.5
	TVME(isoD)FVAFDK	96.0
3	LGEYGFQDALIVR	98.3
	LGEYGFQ(isoD)ALIVR	98.7

The synthetic peptides were dissolved in Milli-Q (Direct-Q® 3 UV System, Millipore Corporation, US) H_2O (~pH 7) at a concentration of 200 µM for storage in -80 °C freezer. The 200 µM stock solutions of each peptide were then diluted to 10 µM with 49.95:49.95:0.1 water/acetonitrile/formic acid prior to MS analysis. In this work, focus was placed on the isoD quantification by MS/MS of peptides 1-3, which were obtained by mixing the D and isoD containing peptides at 20 %, 40 %, 60 %, 80 %, and 100 % with concentration at 10 µM.

Sample preparation for Bovine Albumin Serum (BSA) and deamidation experiments

BSA powder (Sigma Aldrich Company Ltd, Dorset, England) was dissolved in 100 mM ammonium bicarbonate (ABC) solution to 2 mg/mL with a volume of 25 mL. Disulphide bonds were then reduced with 3 mL of 50 mM dithiothreitol (DTT, Sigma Aldrich Company Ltd) for 30 minutes at 60 °C, followed by alkylation with 3 mL of 100 mM of iodoacetamide (IAA, Sigma Aldrich Company Ltd) and stored the sample in dark at room temperature for 1 hour. The solution was then tryptic digested with 2 mL of 1 mg/mL trypsin (Sigma Aldrich Company Ltd) in 100 mM ABC solution at 37 °C for 16 hours. After the digestion, the sample was first purified with HyperSEP C18 solid phase extraction cartridges (Thermo Scientific, Waltham, MA, USA) to remove excess chemicals. Samples were then dried with a Savant SPD121P SpeedVac concentrator (Thermo Scientific, Waltham, MA, USA) to remove the organic solvent. Dried, digested BSA was then re-suspended in 10 mM ABC solution (~ pH 8) to a final concentration of 10 μM. All samples were then incubated at 60 °C for 5 days and one samples was taken out for MS analysis at day 0, 1, 3, and 5. Prior to direct infusion analysis on the FT-ICR MS, samples were desalted using SOLAμ SPE C18 cartridges (ThermoFisher, Waltham, MA, USA). The desalted BSA digest samples were diluted with 49.95:49.95:0.1 water/acetonitrile/formic acid into final concentrations of 1-10 μM for MS analysis.

Nano-LC separation

A homemade C18 RP nano capillary trap column (3cm, 150 μm I.D., 3 μm particle size, 300 Å pore size) and an in-house packed C18 RP nano capillary analytical column (20cm, 75 μm I.D., 3 μm particle size, 300 Å pore size) were connected to an EASY nLC II system (Proxeon, Hemel Hempstead, UK) for the online nLC separation of the tryptic digested BSA samples. The sample was separated with solvent A (99.9% H_2O with 0.1% formic acid) and solvent B (19.9% H_2O, 80% ACN, and 0.1% formic acid) on the nLC system.

0.2 μM tryptic digested BSA was loaded onto the C18 RP column. Sample was separated with the analytical column at a constant flow rate of 300 nL/min from 5% to 35 % solvent B for 30 minutes, maintaining at 35 % solvent B for 5 minutes, following by a short gradient from 35% to 80% solvent B for 10 minutes, maintaining at 80% solvent B for 30 minutes with constant flow rate increased to 500 nL/min, then reducing from 80% to 5% solvent B for 1 minute, and equilibrating at 5% solvent B for 15 minutes. The total run

time of one sample is around 90 minutes which included sample loading, nLC separation, and column equilibration time.

The nLC separation was automatically controlled by Bruker Daltonics Hyster automation software (Bruker Daltonics, Bremen, Germany). The nLC platform was coupled to a 12 Tesla Bruker SolariX FT-ICR-MS via a custom-made nanospray source using pre-cut conductively coated SilicaTipTM emitters of 5cm long, with a 360 μm tip O.D., 20 μm I.D., 10 μm tip I.D. (New Objective, MA, USA).

FT-ICR MS analysis

All the experiments were carried out using a 12 tesla (T) SolariX Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS; Bruker Daltonik GmbH, Bremen, Germany), equipped with a shielded superconducting magnet.

Direct infusion experiment

For the direct infusion experiments, the samples were loaded into borosilicate glass capillary tips (purchased from World Precision Instruments, Inc., Sarasota, FL, USA), which were pulled using a Sutter P-97 capillary Flaming/Brown micropipette puller instrument (Sutter instruments Co., Novato, CA, USA). The pulled tips were optimised for a low-flow nano-electrospray ionisation (nESI) experiments.

All samples were sprayed in positive ionisation mode. Mass spectra were acquired with a 4 mega-word (M) data-points (32 bits) over a mass range of *m/z* 147 – 3,000 to produce a 1.67 s transient and ~500,000 resolving power at *m/z* 400.

Positively charged ions were transmitted through a glass capillary to a quadrupole and then externally accumulated in a hexapole collision cell for 0.35 s before transferred to an infinity cell for MS excitation and detection.

For CAD MS/MS experiments, 2+ precursor ions of the synthetic peptides (Table 2.1) were first quadrupole isolated at *m/z* 582.8, *m/z* 700.8 and *m/z* 740.9 respectively with an isolation window ranging from 5-10 *m/z*. The ions then underwent collisions with argon gas in the collision cell. The optimised collision energies for each peptide were ranged from 10 V – 18 V. Fragments, together with the precursor ions, were then transferred to the infinity cell for mass detection.

For IRMPD MS/MS, precursor ions were first isolated with the quadrupole and then transmitted to the ICR cell. The ions were then fragmented using a continuous-wave CO_2 laser (Synrad, Washington, USA) with an output wavelength of 10.6 μm. The optimised pulse length of 0.1-0.3 s and 65% laser power from a 25 W laser was used for the fragmentation. The CO_2 laser was introduced from the back of the ICR cell through a BaF_2 window and precursor ions were fragmented inside the ICR cell before excitation and detection.

For ECD MS/MS experiments, after quadrupole isolation, ions were directly transferred to the ICR cell, where they were irradiated with low energy electrons emitted from an indirectly heated hollow dispenser cathode heated via 1.5 A continuous current for ion fragmentation. The optimised ECD parameters for the peptide fragmentation were electron irradiation time of 0.1-0.3 s, bias of 1.0-1.5 V, extraction lens at 4.0 V, and cathode potential of 1.5 V.

With the pre-existing IRMPD setup, a 193 nm ArF excimer laser beam (10 Hz; Coherent, UK) was also introduced from the back of the ICR cell through a BaF_2 window. Like the IRMPD MS/MS experiment, ions were first isolated in the quadrupole, transmitted to the ICR cell, and eventually irradiated with 1 laser shot (~5 mJ/pulse at the laser head) to generate fragments.

For 213 nm UVPD fragmentation, a stable telescopic compact high energy Q-switched pulsed Nd:YAG laser with an output wavelength of 213 nm (5^{th} harmonic of the Nd:YAG laser) (10 Hz; Litron Lasers, UK) was also used and ions were irradiated with 10 laser shots (~1.5 mJ/pulse at the laser head).

For all direct infusion MS/MS experiments (excluding 213 nm UVPD data), samples were prepared and ran in triplicate.

nLC MS/MS experiment

For the nLC MS/MS experiments, spectra were acquired with 1 M data-point and auto MS/MS was applied. Throughout the nLC experiment, one MS scan was followed with one auto MS/MS scan. 0.3 s accumulation time was used for MS scan. The highest intensity peptide in the MS scan with charge state between 2+ and 5+, as well as intensity threshold higher than $1x10^7$ was automatically selected for ECD MS/MS fragmentation. Peptide ions were then isolated with 5 *m/z* window in quadrupole and accumulated in the collision cell for 1 s (MS/MS boost) to achieve higher intensity before ECD fragmentation in the ICR cell. The ions were then fragmented by irradiation with 1.5 eV electrons produced from an indirectly heated hollow dispenser cathode (1.5 A) for 0.1 s, with an ECD bias of 1.5 V and the extraction lens at 4.0 V. The fragmented *m/z* was then excluded from auto MS/MS for 60 seconds. During the experiment, unknown and single charge state peptides as well as peptides presented between *m/z* 147 and 350 were excluded from the auto MS/MS experiment.

Data analysis

All spectra were analysed using DataAnalysis 4.3 (Bruker) and fragments were assigned both manually and using in-house developed software with an error <1.2 ppm (supplementary table S2.1 – S2.6).

All spectra were internally calibrated with known *m/z* fragmented peaks that contain minimum threshold of S/N >3 and intensities higher than $1x10^6$ according to the Bruker FTMS peak picking algorithm.

The deamidation percentage was determined by using peak area of the 1st isotopic peak from deamidated peptide, divided by the sum of peak area of the 1st isotopic peak from deamidated and non-deamidated peptides [Eqn. 2.1]:

$$\text{Ratio of deamidation to non-deamidation peak (\%)} = \frac{\text{deamidated } 1^{st} \text{ isotope peak area}}{(\text{non-deamidated } 1^{st} \text{ isotope peak area} + \text{deamidated } 1^{st} \text{ isotope peak area})} \quad \text{[Eqn. 2.1]}$$

5-point calibration curves were generated by mixing the ratio of deamidated and non-deamidated peptides according to the Table 2.2 (peptide 1) for the deamidation quantification MS experiments. 6-point calibration curves were generated by mixing the ratio of isoD and D containing peptides for the MS/MS quantification experiments according to Table 2.4 (peptide 1) and Supplementary table S2.1 and 2.3 (peptide 2 and 3). Peptides 2 and 3 were mixed in the same manner as peptide 1 for the MS/MS quantification experiments.

Table 2. 2 Volume percentage of synthetic peptide 1 in mixtures for MS experiments.

Calibration Point	**LVNELTEFAK (%)**	**LVDELTEFAK (%)**	**LVisoDELTEFAK (%)**
1	20	40	40
2	40	30	30
3	50	25	25
4	60	20	20
5	80	10	10

To achieve the best accuracy for the quantification experiment, the purity of each peptide (Table 2.1) was used to correct the percentage of each component in the mixture solutions. The corrected non-deamidated and deamidated peptide percentages are presented in Table 2.3 (peptide 1) and the corrected isoD and D-containing peptide percentages for the MS/MS experiments are shown in Table 2.5 (peptide 1). The corrected isoD and D-containing peptide percentages for the MS/MS experiments for peptides 2 and 3 can be found in the Supplementary Table 2.2 and 2.4 respectively.

Table 2. 3 Scaled percentage of each synthetic peptide in mixtures with the account of peptide purity for MS experiments.

Calibration Point	**LVNELTEFAK (%)**	**LVDELTEFAK (%)**	**LVisoDELTEFAK (%)**
1	19.02	39.92	39.40
2	38.04	29.94	29.55
3	47.55	24.95	24.63
4	57.06	19.96	19.70
5	76.08	9.98	9.85

Table 2. 4 Volume percentage of synthetic peptide 1 in mixtures for MS/MS experiments.

Calibration Point	**LVisoDELTEFAK (%)**	**LVDELTEFAK (%)**
1	0	100
2	20	80
3	40	60
4	60	40
5	80	20
6	100	0

Table 2. 5 Scaled percentage of each synthetic peptide in mixtures with the account of peptide purity for MS/MS experiments.

Calibration Point	**LVisoDELTEFAK (%)**	**LVDELTEFAK (%)**
1	0	99.8
2	19.7	79.8
3	39.4	59.9
4	59.1	39.9
5	78.8	20.0
6	98.5	0.0

The relative abundance of the isoD peptides were estimated using the following equation [Eqn. 2.2], as mentioned previously by Cournoyer *et. al.*[58] with the change made to use peak areas rather than peak intensities:

$$Relative\ percentage\ ratio\ of\ isoD\ (\%) = \frac{z_{n-i}-C_2O_2\mathrm{H}\ \ peak\ area}{sum\ of\ all\ fragment\ peak\ areas} \times 100 \qquad \text{[Eqn. 2.2]}$$

z_{n-i} - C_2O_2H is the characteristic peak generated by isoD peptide compared to D peptide. A calibration curve was then generated to correlate the isoD percentage in the mixed standard peptide solution with the estimated relative percentage of isoD peptides observed in the MS spectra. Each calibration point was calculated by averaging the peak area or intensity (previous publication method)[58] from the triplicate experiments. The R^2 of the calibration curve was then calculated by linear fitting of the curve. Although, the relative isoD abundance calculation method for the ECD MS/MS quantification

experiment is similar in principle for the UVPD MS/MS quantification experiment, due to the absence of diagnostic z_{n-i} - C_2O_2H fragment in the isoD peptide UVPD spectra, a different characteristic fragment (y fragment generated at the deamidated modification site) was utilised, hence the following equation [Eqn. 2.3] was applied for the calculation of relative percentage of isoD from the UVPD MS/MS spectra:

$$Relative\ percentage\ ratio\ of\ isoD\ (\%) = \frac{y \text{ fragment peak area at deamidation modification site}}{sum\ of\ all\ fragment\ peak\ areas} \times 100 \quad \text{[Eqn. 2.3]}$$

The relative isoD peptide ratios of the tryptic digested BSA observed in the MS spectra was also estimated via the above equations. The obtained percentage values were then applied into the calibration curve which was used to determine the actual percentage of isoD peptide in the BSA samples.

For the nLC experiment, the percentage of D and isoD peptides were calculated using the peak area obtained from the extracted ion chromatogram (XIC) with the target *m/z* of peptide 1, 2, and 3 (± 0.1 *m/z*) in the BSA digested sample. The obtained peak area was then calculated with the following equation [Eqn. 2.4] to convert to percentage of the isoD peptide in the mixture solution:

$$isoD\ \% = \frac{isoD\ peak\ area}{isoD+N+D\ peak\ area} \times 100 \quad \text{[Eqn. 2.4]}$$

All calibration curves were plotted and a linear fit was applied using the software package of Origin 2019 (OriginLab Corporation, USA).

2.4. Results and Discussion

MS detection of deamidation

31 tryptic digested BSA peptides were observed in the direct infusion MS spectrum (Figure 2.5a). 14 potential Asn deamidation sites were found in the BSA sequence. Peptide deamidation results in a mass shift of + 0.984 Da in the mass spectrum and 7 BSA peptides (Figure 2.5b) were clearly observed with a mass shift of +0.984 Da in the MS spectrum which were also pattern matched with the theoretical peak stimulation as shown in Figure S2.2. High resolution MS is required for the differentiation of deamidated peptides from non-deamidated peptides directly in a MS scan without prior separation due to the subtle mass difference between the first isotope of a deamidated peptide and the second isotope of a non-deamidated peptide, i.e., 0.019 Da for a 1+ charge state peptide. Three target deamidated peptides (Figure 2.5) out of 7 deamidated peptides from the digested BSA sample (Table 2.1) were selected for further detail analysis (highlighted in red in Figure 2.5b) as these three peptides contained the highest deamidation rate which facilitated the following quantification experiments regarding the comparison of the performance of various MS/MS fragmentation techniques.

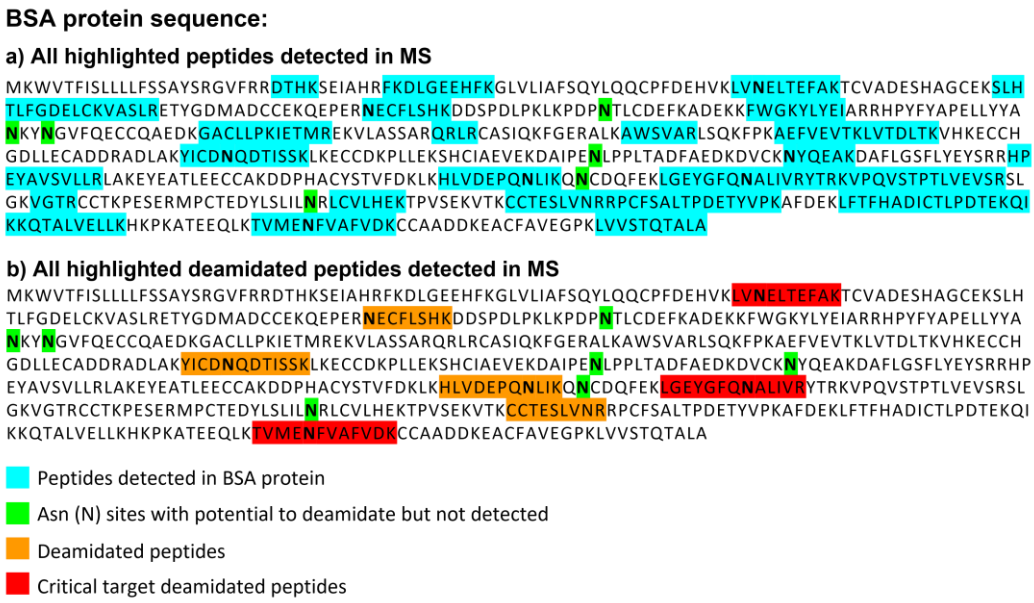


Figure 2. 5 BSA sequence highlighted a) all detected peptides and b) all deamidated peptides observed in the experiment of 10 µM tryptic digested BSA direct infusion into the 12 T FT-ICR-MS.

The mass difference between the first isotopic peak of the non-deamidated and deamidated peptide 1, $[LVNELTEFAK+2H]^{2+}$ was 0.492 *m/z* at the *m/z* 582.5 range, which can be resolved by most commercial MS instruments (Figure 2.6a). However, the mass difference between the second isotopic peak of non-deamidated peptide 1 and the first isotopic peak of deamidated peptide 1 was only different by 0.009 *m/z* (Figure 2.8a bottom trace zoomed in), a high-resolution mass spectrometer is therefore required to baseline resolve the peaks. The fully resolved deamidated peak is critical for direct infusion as the peak intensity or area is useful for determining the relative deamidation percentage of sample mixtures.[62] To fully resolve the deamidated peaks at 582.5 *m/z*, theoretical resolving power was required at ~194 k (Figure 2.7). Peptide 2 and peptide 3 were at 700 *m/z* and 740 *m/z* range respectively, the mass difference between the second non-deamidated peak and the first deamidated peak was still maintained at 0.009 *m/z*; however, higher resolving power was required to achieve the baseline separation in a higher mass range, indicating the importance of utilising high-resolution MS for this type of experiment. In the experiment, the experimental resolving power at *m/z* 582 (peptide 1), *m/z* 700 (peptide 2), and *m/z* 740 (peptide 3) were 346 k, 337 k, and 330 k respectively, which were around 1.2- to 2-fold higher than the minimum required resolving power for a full resolution of the deamidated peak (Figure 2.6 & 2.7), suggesting the instrument and acquisition method applied in here were suitable for the following deamidation experiment.

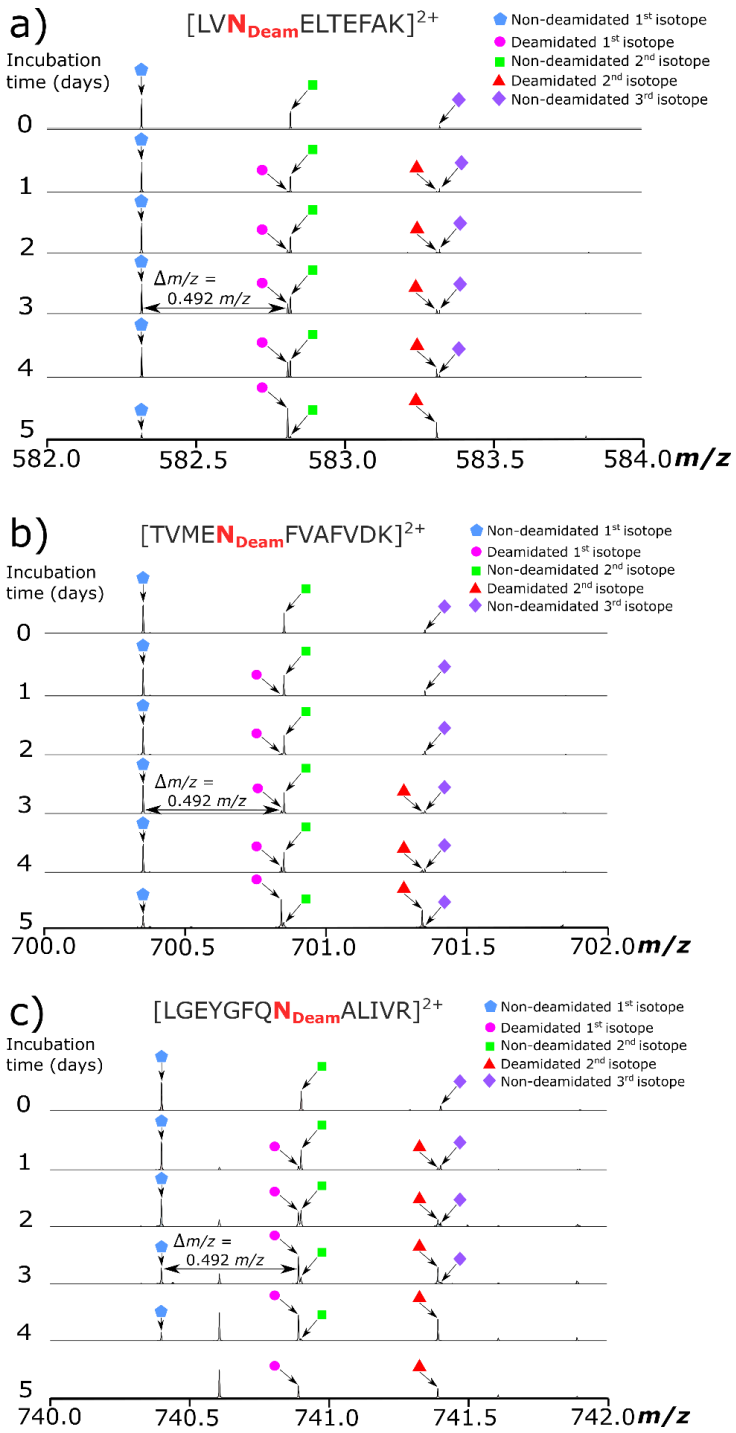


Figure 2. 6 Zoom in of the *m/z* regions focusing on the doubly charged precursors in the mass spectra of the deamidated tryptic digested BSA peptides of a) $[LVN_{Deam}ELTEFAK+2H]^{2+}$ b) $[TVMEN_{Deam}FVAFVDK+2H]^{2+}$, and c) $[LGEYGFQN_{Deam}ALIVR+2H]^{2+}$.

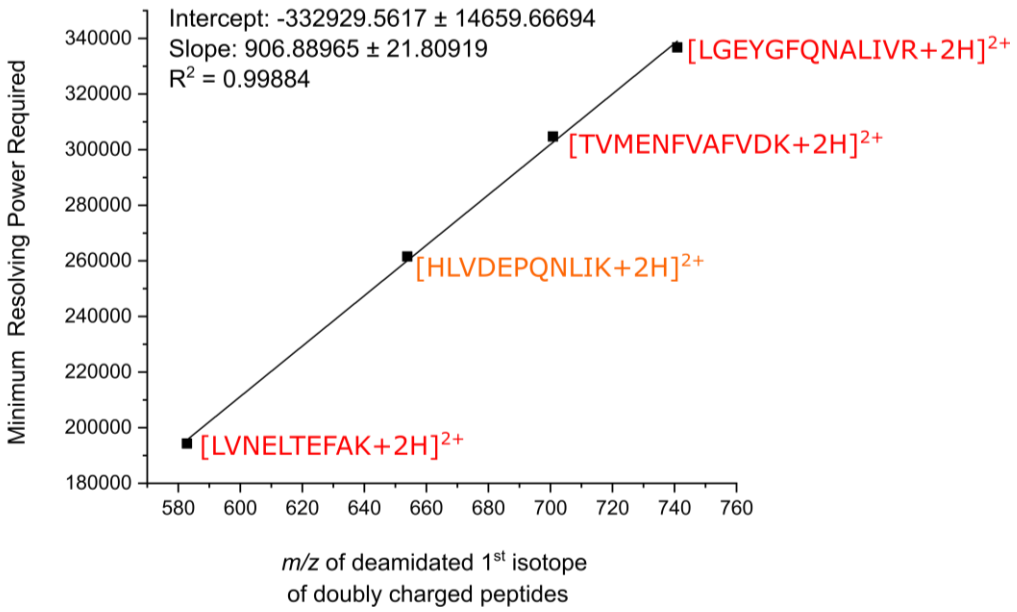


Figure 2. 7 Minimum theoretical resolving power against *m/z* for baseline resolving 2nd non-deamidated isotopic peak and 1st deamidation isotopic peak (peptide that did deamidate was coloured in orange and the critical target deamidated peptides are coloured in red). Data was obtained via simulation of the theoretical isotopic pattern spectra for the molecular formulae of the non-deamidated and deamidated forms of the peptides.

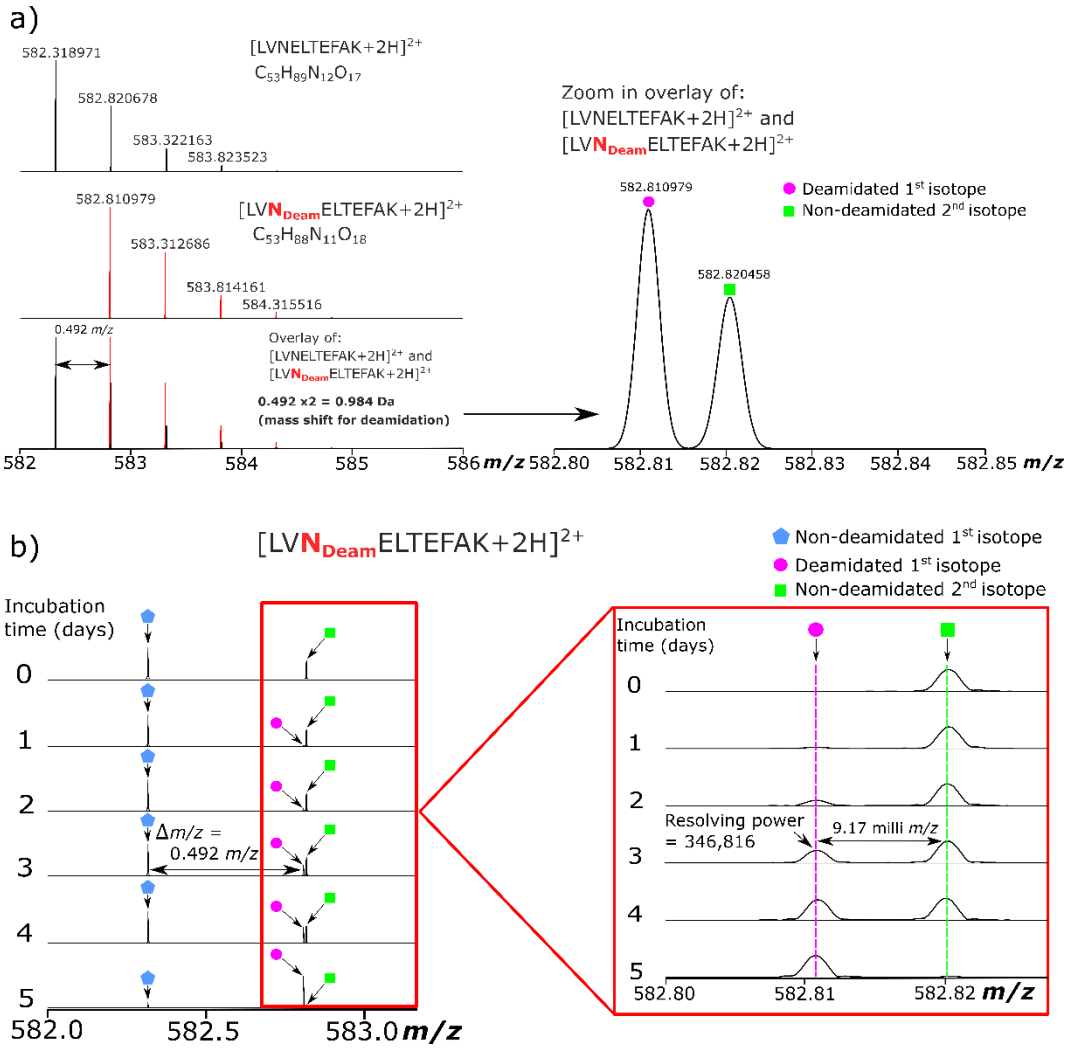


Figure 2. 8 a) Theoretical isotopic distributions of BSA peptide ([LVNELTEFAK+2H]$^{2+}$) with 0 % deamidation (top trace), 100 % deamidation (middle trace), and overlay of the non-deamidated and fully deamidated peptide (bottom trace) with a zoom in on the bottom trace for the *m/z* region 582.80-582.85 to demonstrate the baseline resolution needed for deamidation quantification in MS. b) Deamidated tryptic digested BSA peptides of [LVNELTEFAK+2H]$^{2+}$ clearly observed within 5 incubation days.

The deamidation percentage for peptide 1 ([LVNELTEFAK+2H]$^{2+}$), peptide 2 ([TVMENFVAFDK+2H]$^{2+}$), and peptide 3 ([LGEYGFQNALIVR+2H]$^{2+}$) were generated by measuring the peak area ratio of the 1st isotopic peak between the non-deamidated and deamidated peptides according to the method mentioned in the experimental section. As an example, the R-square of the calibration curve for deamidation quantification of peptide 1 was > 0.99 (Supplementary Figure S2.1), indicating that MS is a suitable analytical tool for measuring the deamidated percentage in a mixture solution, which is similar to previous studies.[7,15,63] The relative deamidation percentages after one day of

incubation were at 13.2 %, 5.1 %, and 35.0 % for peptide 1, peptide 2, and peptide 3 respectively (Figure 2.9), suggesting there should be sufficient levels of deamidated peptides for D and isoD differentiation for the following experiment. The deamidation rates of peptides (peptide 3 > peptide 1 > peptide 2) were also aligned with the predicted deamidation order at 37 °C according to the Robinson *et. al*. study (Supplementary Table S2.5),[21] suggesting the influence of deamidation rate in the experiment was dominated by the neighbouring amino acid, and the temperature effect does not significantly affect the order of deamidation rate.

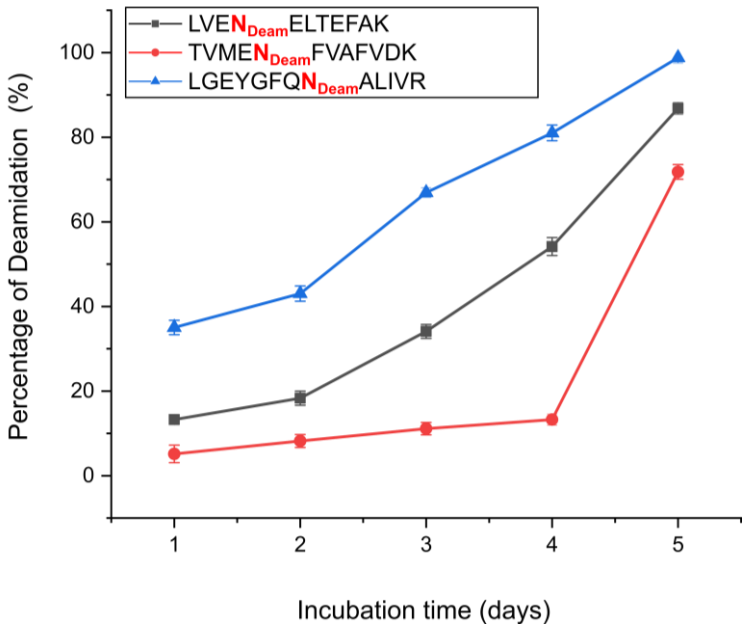


Figure 2. 9 Daily deamidation percentage of peptide 1 ($[LVN_{Deam}ELTEFAK+2H]^{2+}$), peptide 2 ($[TVMEN_{Deam}FVAFDK+2H]^{2+}$), and peptide 3 ($[LGEYGFQN_{Deam}ALIVR+2H]^{2+}$) from BSA tryptic digested samples which were incubated at 60 °C for 5 days.

Comparison and optimisation of isoD peptide quantification using ECD MS/MS

ECD MS/MS has been shown to generate the characteristic z_{n-i}-C_2O_2H (z_{n-i}-57) fragment ion reliably from isoD peptides, which has also been generally applied to quantify the relative percentage of D and isoD peptides in a mixture sample.[15,28,35,46,51,58,64-67] Various methodologies have been applied to quantify isoD peptide percentage in a sample; however, no previous study has included the comparison between different quantification methods. Herein, various quantification methods were used to generate the calibration curves for the deamidated synthetic peptides 1 to 3, in which the characteristic fragment was generated in all ECD MS/MS spectra (Supplementary Figure S2.4-2.5).

In general, methods used for isoD peptide quantification can be classified into 3 groups. The first method (Method A) was proposed by Ni *et. al.*[36] which used the peak area ratio of the characteristic fragment over the z ion fragment at the deamidation site. The R-squares obtained for peptide 1 to 3 were ranging from 0.87 to 0.92, with an average R-square at 0.90 (Figure 2.10A). The second method (Modified Method B) was proposed by Pekov et. al.[60] which applied the peak area ratios of the b/y fragments closer to the deamidation site (isoD dependent fragment) to b/y fragments that do not contain the deamidation site (isoD independent fragment). The principle of this method was utilised and modified for the ECD quantification data therefore it was referred to as the modified method B. The R-squares generated from peptide 1 to 3 were between 0.93 and 0.96, the average R-square was 0.95 (Figure IIB). The last method (Method C) was applied by Cournoyer et. al[58] which used the peak intensity of the characteristic fragment of isoD peptide over the sum of all fragment peak intensities. The R-squares resulted from peptide 1 to 3 were ranged between 0.96 to 0.99 with an average R-square at 0.98 (Figure 2.10IIC).

Method A obtained the lowest average R-square due to the nominator (characteristic fragment) and denominator (z ion at deamidation site) were significantly affected by the percentage of isoD peptide, resulting in a low stability of calculating the percentage isoD peptide throughout a calibration curve. Method B obtained a higher average R-square than method A as the dominator (independent z ion) was not significantly affected by the percentage change of isoD peptide. Method C obtained the highest average R-square value compared to the other two methods as the sum of all fragment intensities were used which significantly decreased the fluctuation in the

denominator value across experiments, resulting in a higher accuracy in determining the percentage of isoD peptides in a sample mixture.

Even though Method C is the best method among all the other methods for determining the percentage of isoD peptide in a mixture solution, it still has a limitation when comes to low abundance characteristic fragments. Method C obtained a very good R-square with Peptide 1 and 3 as the normalised intensities of the characteristic fragments at 20 % isoD peptide in the mixtures were 2.95 % and 5.42 % respectively. For peptide 2, the normalised intensity of characteristic fragment was only 0.94 %, which was 3.1-fold and 5.8-fold lower than the fragments observed in peptide 1 and 3. Peak distortion was easily observed in low abundance ion peaks[36,68] which would remarkably affect the peak intensity and influence the quantification result. To truly reflect the quantity of low abundance ion peak; peak area, instead of intensity, should be applied. Herein, we adopted the methodology from Cournoyer *et. al.*[58] and further improved it with using peak area for quantification. With the new proposed method (Method D), the R-squares of peptide 1 to 3 were now greater than 0.99 (Figure 2.10IID), indicating this method is suitable for high as well as low abundance peak quantification. A summarised table for peptide 1 to 3 calibration curves using various quantification methods was shown in Table 2.6 (I).

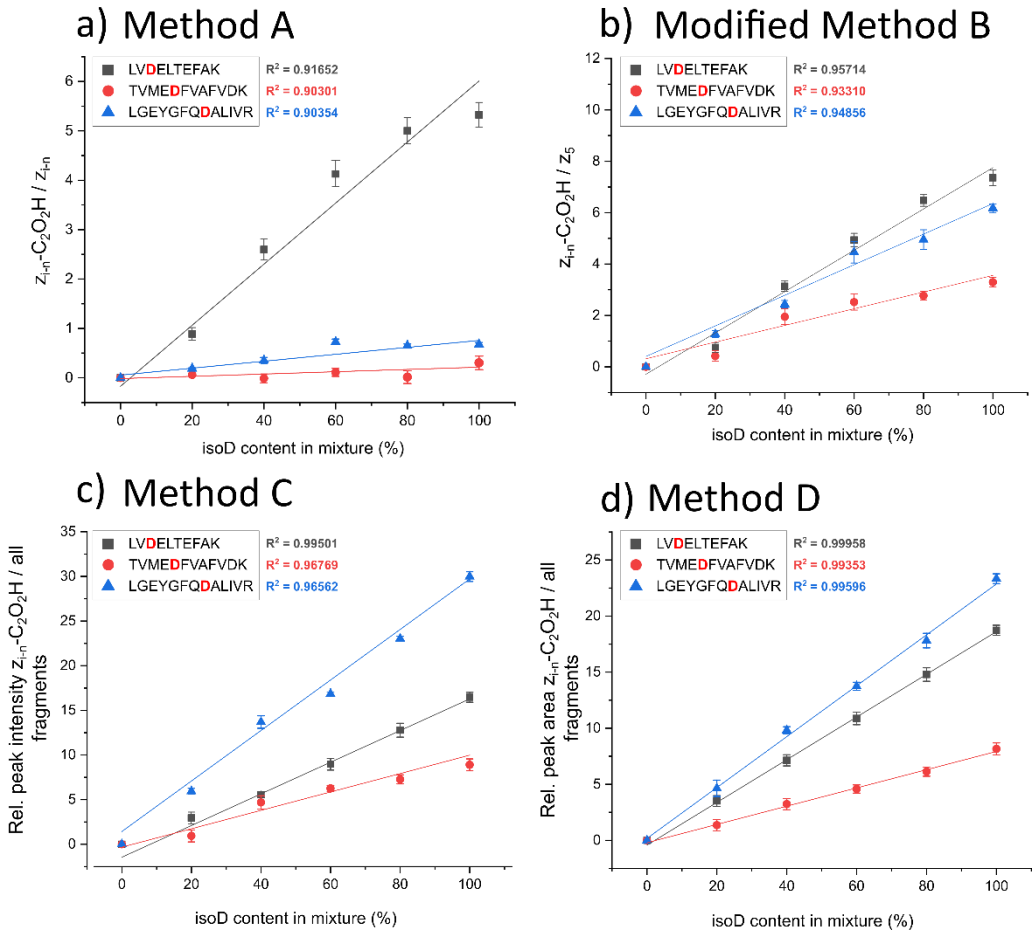


Figure 2. 10 (II) Calibration curves of peptide 1, 2, and 3 with a) Method A (Ni *et. al.*[36] method), and b) Modified method B (Pekov *et. al.*[60] method), c) Method C (Cournoyer *et. al.*[58] method), and d) Method D (proposed method) with corresponding R-square for each peptide calibration curve.

Table 2. 6 A summarised table for all quantification methods used for the isoD peptide determination in a mixture sample.

Method	Method reference	Brief explanation of method	Peptide group number	R^2	Average R^2	Comments
A	Ni *et al.*[36]	Peak area ratio of $z_{i-n}-C_2O_2H$ fragment ion to the z_{i-n} fragment ion.	1	0.91652	0.90769	Overall weak R^2 values (R^2 <0.92) therefore not appropriate for generation of linear calibration curves and sensitive to low % isoD.
			2	0.90301		
			3	0.90354		
B	Pekov *et al.*[60]	Modified method uses the peak area ratio of $z_{i-n}-C_2O_2H$ (isoD dependent fragment) to the z_5 fragment (isoD independent fragment).	1	0.95714	0.94626	Improvement in R^2 values as the z_5 fragment used for quantification is less affected by the z_{i-n} fragment ion generated at the modification site.
			2	0.93310		
			3	0.94856		
C	Cournoyer *et al.*[58]	Relative intensity ratio of the $z_{i-n}-C_2O_2H$ fragment ion with respect to the sum of all fragments peak intensities.	1	0.99501	0.976107	Further improvement in the linearity of the calibration curve observed (R^2> 0.96). However, intensity-based measurements suffer at lower limit of detection.
			2	0.96769		
			3	0.96562		
D	Proposed	Peak area ratio of the $z_{i-n}-C_2O_2H$ fragment ion to the sum of all fragment peak areas.	1	0.99958	0.99636	Less deviation, more robust, improvement in R^2 values observed. Peak area added into noise whereas the peak intensities are relative to noise.
			2	0.99353		
			3	0.99596		

The percentages of isoD peptide in the incubated tryptic digested BSA samples were determined using ECD MS/MS. The characteristic z_{n-i}-C_2O_2H fragment peak was clearly observed in target peptide 1 to 3 (Figure 2.11 III A-C), indicating isoD peptides were formed during the deamidation process of peptide 1 to 3. The percentages of isoD peptides were then calculated by using the calibration curves obtained from above (Figure 2.10 II Method D). The percentages of the isoD peptides increased with the increasing incubation time at 60 °C (Figure 2.11 III d), which were also in direct proportion to the deamidation percentage obtained for each peptide. The non-deamidated peptide, deamidated D-peptide, and deamidated isoD-peptide percentages of each target BSA

peptide at various incubation were calculated and summarised in Table 2.7, which demonstrated the ratio of D- to isoD-peptide was not always equal to 1:3.

Table 2. 7 A table of the non-deamidated peptide, deamidated D-peptide, and deamidated isoD-peptide percentages at different incubation times.

Incubation time (day)	Peptide 1			Peptide 2			Peptide 3		
	Non-deamidated (N) (%)	isoD (%)	D (%)	Non-deamidated (N) (%)	isoD (%)	D (%)	Non-deamidated (N) (%)	isoD (%)	D (%)
1	86.75	9.20	4.05	94.86	3.54	1.60	64.97	31.91	3.12
3	65.91	24.73	9.36	88.89	6.60	4.51	33.08	58.56	8.36
5	13.18	36.45	50.37	28.19	13.44	58.36	1.22	72.16	26.62

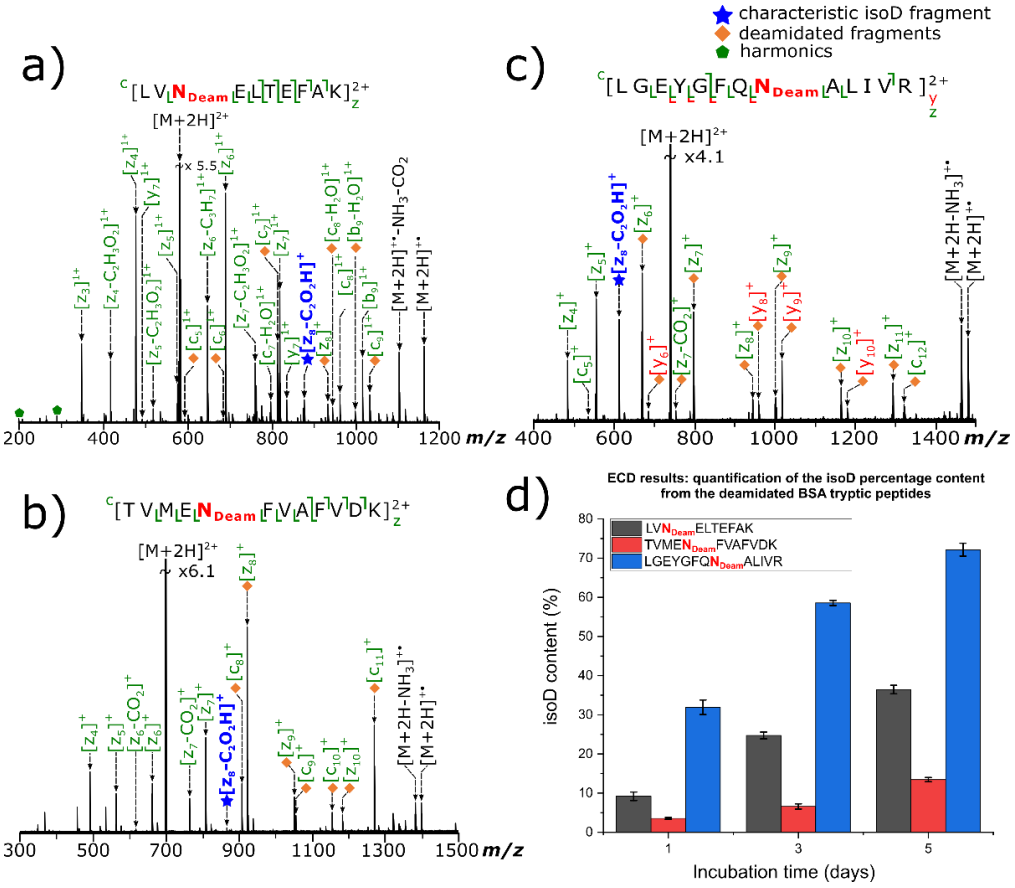


Figure 2. 11 (III) The ECD MS/MS fragmentation spectra of the target BSA a) peptide 1 ($[LVN_{Deam}ELTEFAK+2H]^{2+}$), b) peptide 2 ($[TVMEN_{Deam}FVAFDK+2H]^{2+}$), and c) peptide 3 ($[LGEYGFQN_{Deam}ALIVR+2H]^{2+}$) at day-5 incubation and d) isoD quantification plot for the deamidated BSA peptides. Peak assignment tables for assigned MS/MS spectra with absolute average mass errors approximately $< 0.65 \pm 0.62$ ppm for all peptides (Figure S2.12 – S2.14).

Applying 193 nm UVPD laser to differentiate and quantify D and isoD peptides without modifier

ECD MS/MS has been applied for identification and quantification of isoD peptides in a mixture solution[35,46,58,66]; the electron cathode used in ECD MS/MS, however, is still mainly utilised with FT-ICR MS, although cathodes have been developed (e.g. ExD Cells by e-MSion) for implementation on other mass spectrometers such as Q-ToF instruments. 193 nm UVPD is a re-emerging MS/MS technique for protein and peptide assignment[53,54,69], which can easily co-operate with various types of MS instruments.[70-73] It is also known that a peptide bonds has a strong absorption around 190 nm (Figure 2.12) therefore access to direct dissociation pathways is feasible resulting in the formation of all types of fragment ion series.[74-76]

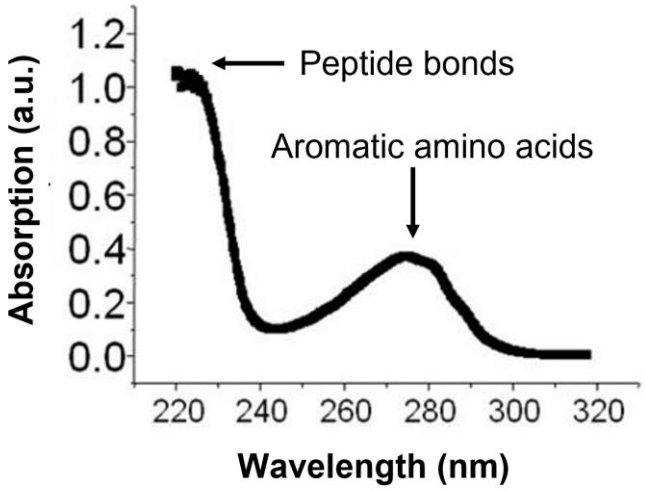


Figure 2. 12 Absorption spectrum of a protein (adapted from Introduction to Practical Biochemistry by Hegyi *et. al.*).[77]

Herein, 193 nm UVPD MS/MS was directly applied to differentiate and quantify the percentage of isoD in the mixture solutions, the quantification results would then compare to the ECD MS/MS results to determine the performance of 193 nm UVPD MS/MS in isoD peptide quantification.

a, b, c, x, y, and z fragments were generated when the synthetic peptide 1 ($[LVNELTEFAK+2H]^{2+}$), peptide 2 ($[TVMENFVAFDK+2H]^{2+}$), and peptide 3 ($[LGEYGFQNALIVR+2H]^{2+}$) were dissociated with 193 nm UVPD MS/MS (Figure 2.13 IV), resulting in more complicated fragmentation spectra than ECD MS/MS. The peptide

sequence coverages of synthetic peptide 1, 2, and 3 using ECD MS/MS fragmentation were 90 %, 100 %, and 100% respectively (Figure S2.4); while the peptide sequence coverage of synthetic peptide 1, 2, and 3 using UVPD MS/MS fragmentation were 100 %, 92 %, and 100 % correspondingly (Figure 2.13 IV and Figure S2.8), indicating UVPD MS/MS can achieve similar or even better sequence coverage than ECD MS/MS.

The characteristic fragment (z_{n-i}-C_2O_2H) was observed in isoD-containing peptide compared to D-containing peptides when fragmented using ECD,[46] and it was commonly used for isoD peptide quantification in mixture solutions. In contrast, no characteristic fragment was observed between D- and isoD-containing peptide when 193 nm UVPD was applied to fragment the synthetic peptide 1, 2, and 3 (Figure 2.14 IV); the y ion at the deamidated site, however, demonstrated a significant intensity difference between D- and isoD-peptides (Figure 2.14 V a-c), which may potentially be applied to quantify the percentage of isoD-peptide in the solution. In general, the y ion at isoD deamidated site contained a higher intensity than the y ion at D deamidated site. With the previous developed quantification method, the peak area of y ion at the deamidated sites was calculated and divided by the sum of the peak areas generated from all assigned fragments. The R-squares in the calibration curves of the synthetic peptide 1, 2, and 3 using 193 nm UVPD MS/MS could also achieve higher than 0.99 (Figure 2.14 V d), suggesting the y ion intensities at deamidation site is directly proportional to the percentage of isoD; 193 nm UVPD MS/MS fragmentation; therefore, can also be used to differentiate and quantify the percentage of isoD peptide in a mixture solution.

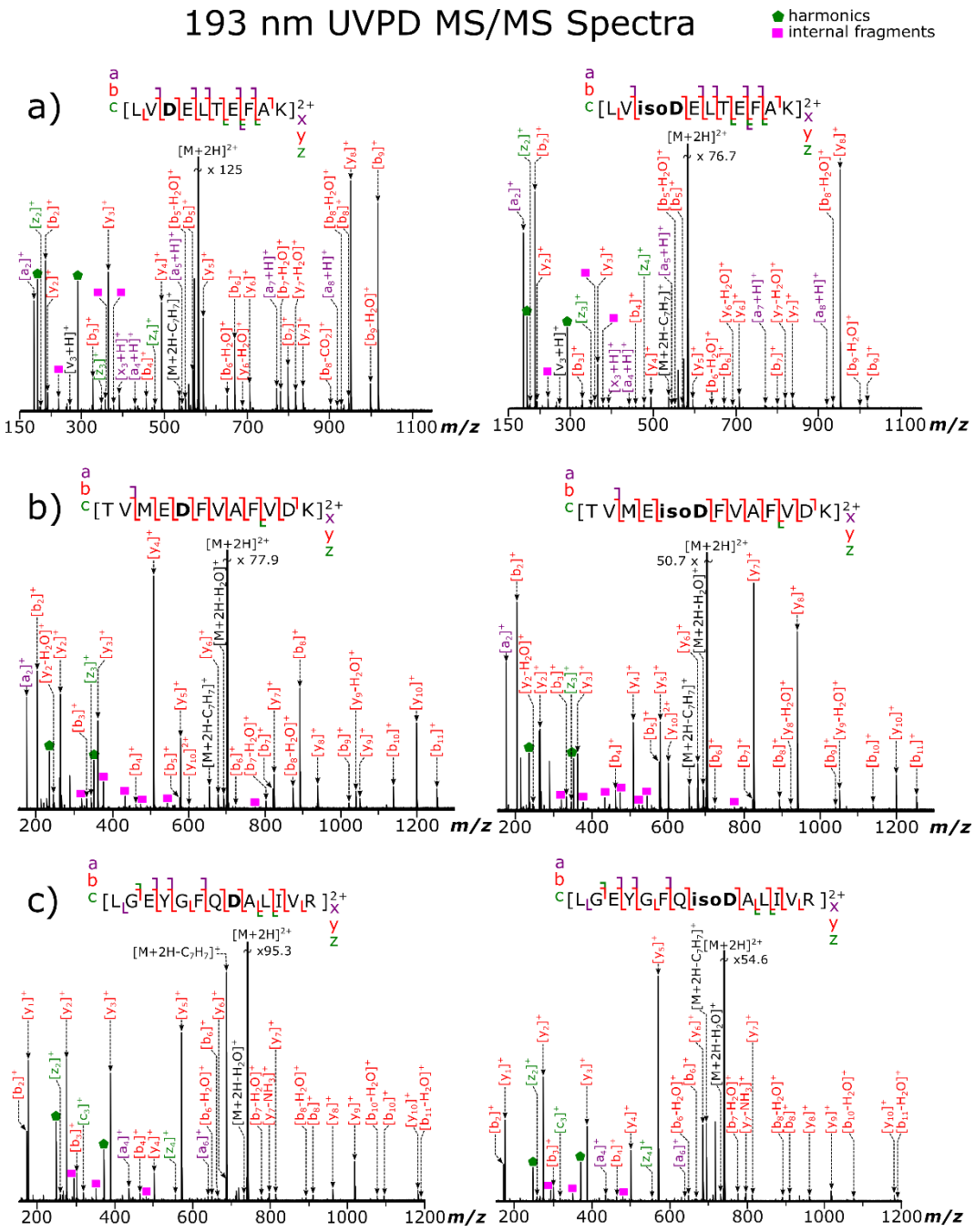


Figure 2. 13 (IV) Experimental 193 nm UVPD MS/MS spectra of synthetic isoD and D peptides a) [LVisoD/DELTEFAK]$^{2+}$ (peptide 1) b) [TVMEisoD/DFVAFVDK+2H]$^{2+}$ (peptide 2) and c) [LGEYGFQisoD/DALIVR+2H]$^{2+}$ (peptide 3). Peak assignment tables for assigned MS/MS spectra with absolute average mass errors approximately $< 0.5 \pm 0.38$ ppm for all peptides (Figure S2.15 – S2.20).

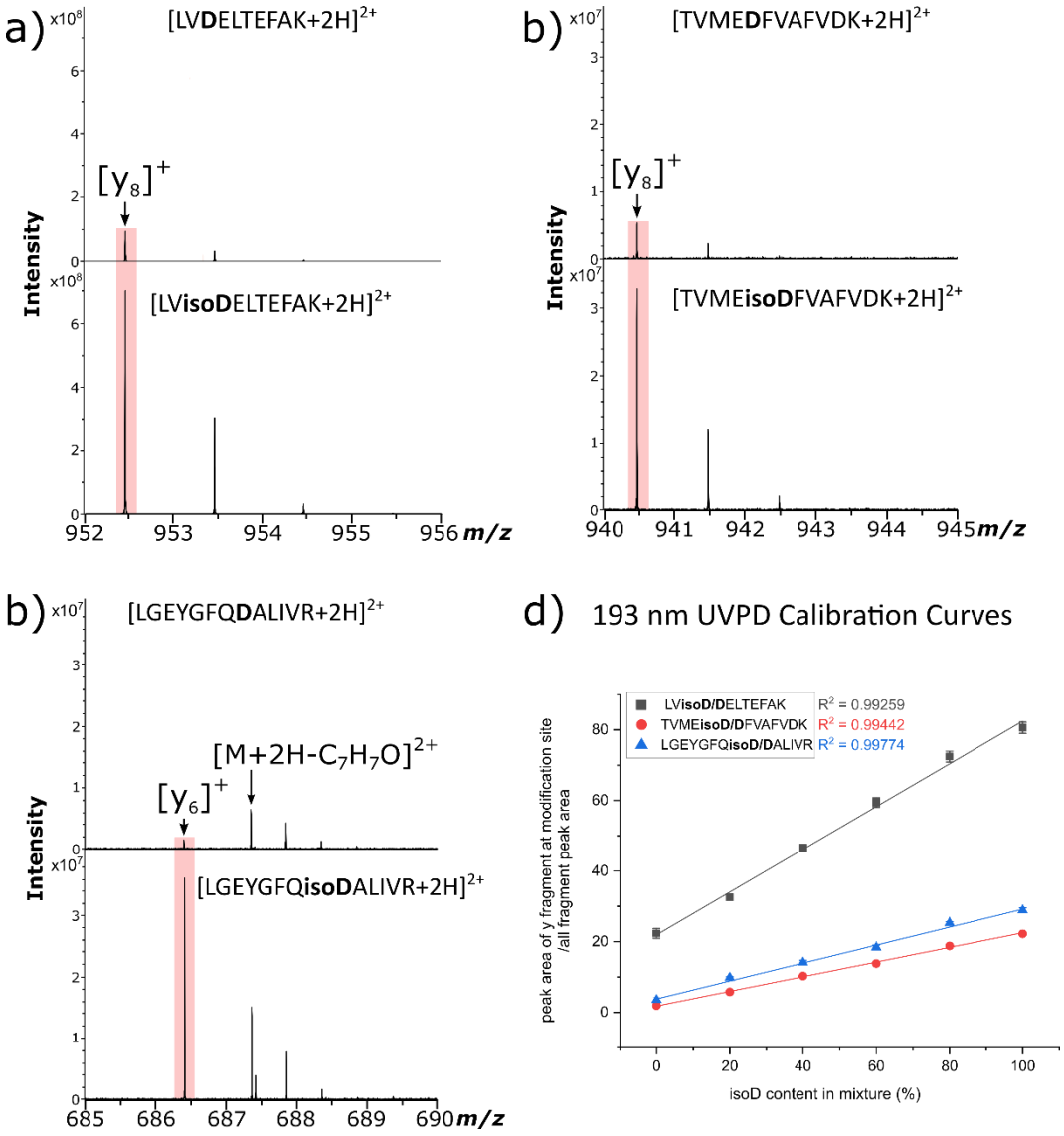


Figure 2. 14 (V) a) to c) Zoom in *m/z* region of the UVPD 193 nm MS/MS spectra for synthetic isoD and D BSA peptides y fragment intensity used for isoD quantification and d) Calibration curves for synthetic BSA peptide mixtures using peak area of y fragment generated at isoD and D cleavage position /sum of all the peak areas of the fragments.

193 nm UVPD was then used to fragment the BSA tryptic digested samples and compared the results against ECD fragmentation. UVPD fragmentation efficiency was ~ 5 times lower than ECD but a similar cleavage percentage was achieved in both fragmentation methods (Figure S2.9 and S2.10). The y ion fragments at the deamidated sites were used for the isoD percentage calculation by applying the calibration generated from Figure 2.14(D), and the isoD percentages of each peptide at different time-points were shown in Figure 2.15(D). The isoD percentages obtained from UVPD MS/MS were

then compared to the results obtained from ECD MS/MS, the result was summarised in Table 2.8.

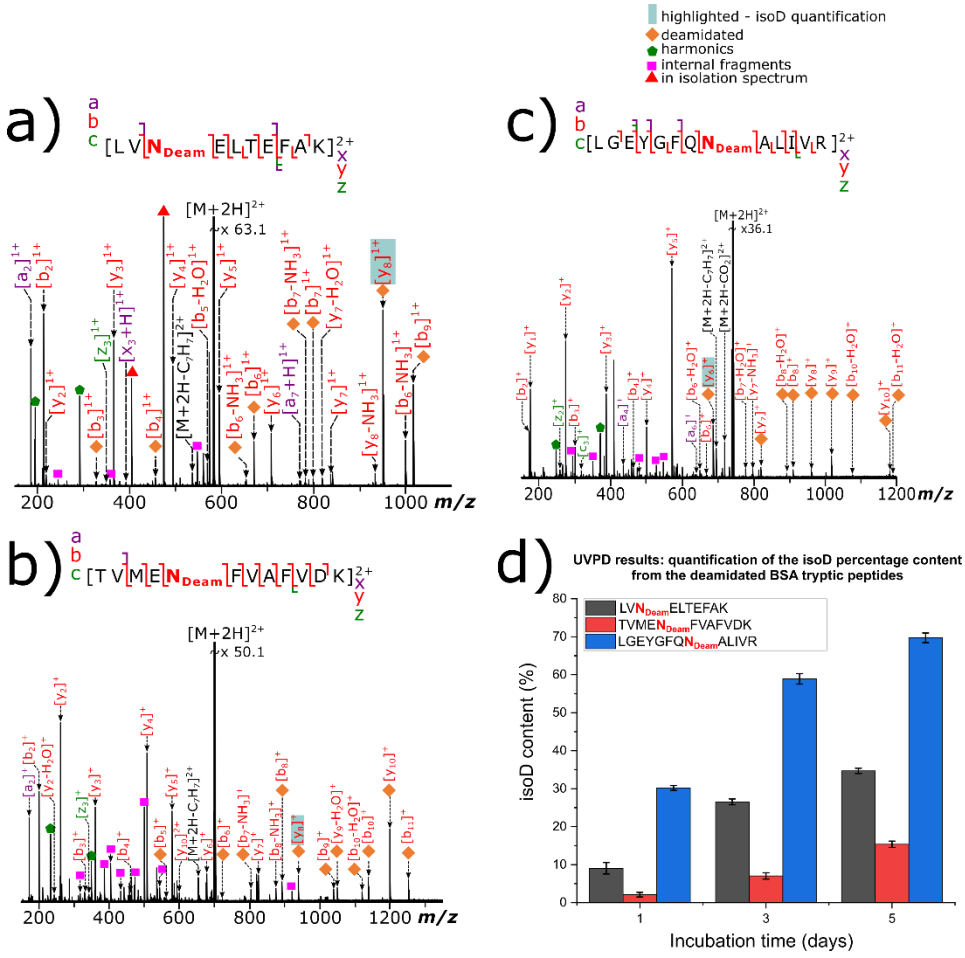


Figure 2. 15 (VI) UVPD 193 nm MS/MS spectra of deamidated BSA peptides 1-3 a) to c) with the isoD quantification results using the 193 nm UVPD method shown in d). Peak assignment tables for assigned MS/MS spectra with absolute average mass errors approximately $< 0.46 \pm 0.32$ ppm for all peptides (Figure S2.21 – S2.23).

The general trend observed with the application of ECD and UVPD MS/MS for isoD quantification is shown in Table 2.8, where an increase in the percentage isoD content in the deamidated peptides is detected with increasing incubation time. Using both MS/MS methods, as confirmed with the deamidated rates of peptides 1-3, the detected percentage of isoD in the peptides also follows the same order, with the lowest percentage of isoD detected in peptide 2, which then increases for peptide 1, and the highest percentage of isoD and D detected over the incubation period was for peptide 3. The ECD and UVPD MS/MS isoD quantification results are comparable with an overall 2.5 % difference observed in the percentage of isoD determined with both fragmentation

methods. However, comments on the underestimation or overestimation of the isoD percentage in the peptides are difficult to make due to the close similarities in the values obtained.

Table 2. 8 A summary table of the deamidated isoD-peptide percentages determined at different incubation times using direct infusion ECD and UVPD MS/MS.

Incubation time (day)	direct infusion ECD			direct infusion 193 nm UVPD		
	Peptide 1	Peptide 2	Peptide 3	Peptide 1	Peptide 2	Peptide 3
	isoD (%)					
1	9.20 ± 1.1	3.54 ± 1.0	31.9 ± 1.9	9.06 ± 1.5	2.11 ± 0.6	30.2 ± 0.6
3	24.7 ± 0.8	6.60 ± 0.7	58.6 ± 0.6	26.5 ± 0.8	7.03 ± 0.8	58.9 ± 1.4
5	36.5 ± 1.1	13.4 ± 0.5	72.2 ± 1.6	34.7 ± 0.7	15.4 ± 0.8	69.7 ± 1.3

Differentiation of isoD and D synthetic peptides using 213 nm UVPD MS/MS and CAD MS/MS

193 nm and 213 nm UVPD, which correspond to 6.4 eV per photon and 6.1 eV per photon respectively, excitation of the peptide back bone and access to direct dissociation pathways are possible.[74] Herein, UVPD using the 5th harmonic of a pulsed Nd:YAG laser with a wavelength of 213 nm was also applied to differentiate between the isomeric isoD and D peptides.

a, b, c, x, y, and z fragments were also generated when the synthetic peptide 1 ([LVD/isoDELTEFAK+2H]$^{2+}$), peptide 2 ([TVMED/isoDFVAFDK+2H]$^{2+}$), and peptide 3 ([LGEYGFQD/isoDALIVR+2H]$^{2+}$) when dissociating using 213 nm UVPD MS/MS (Figure 2.15 VII), which is similar to the results obtained with 193 nm MS/MS (Figure 2.13). The fragment types generated with both 193 nm and 213 nm UVPD were also accompanied with side chain and small neutral losses, resulting in complex spectra compared to ECD MS/MS.

The cleavage coverages of peptide 1, 2, and 3 using 213 nm UVPD were 100 %, 92 % and 100 %, which is comparable with the 193 nm UVPD fragmentation (Figure S2.8). The fragmentation efficiencies of 213 nm UVPD were 6.5 %, 11.7 %, and 9.6 % for the isoD peptides 1, 2, and 3 respectively. The fragmentation efficiencies of 213 nm UVPD were 3.9 %, 9.9 %, and 7.3 % for the D peptides 1, 2, and 3 respectively, indicating that slightly higher fragmentation efficiency is observed with 213 nm UVPD than 193 nm UVPD (Figure S2.7). However, this may be explained by a difference in the

UVPD parameters as a higher number of laser pulses were used during the 213 nm UVPD experiments (10 shots at 1.5 mJ/pulse), whereas 1 laser shot was applied for the 193 nm UVPD experiments at 5 mJ/pulse.

Like the observation with 193 nm UVPD MS/MS spectra, no characteristic fragments were observed in peptide 1, 2, and 3 using 213 nm UVPD fragmentation. Not surprisingly, distinct intensity differences at deamidated y ion sites were still observed between D and isoD peptides which could potentially be used for isoD peptide quantification (Figure 2.17A).

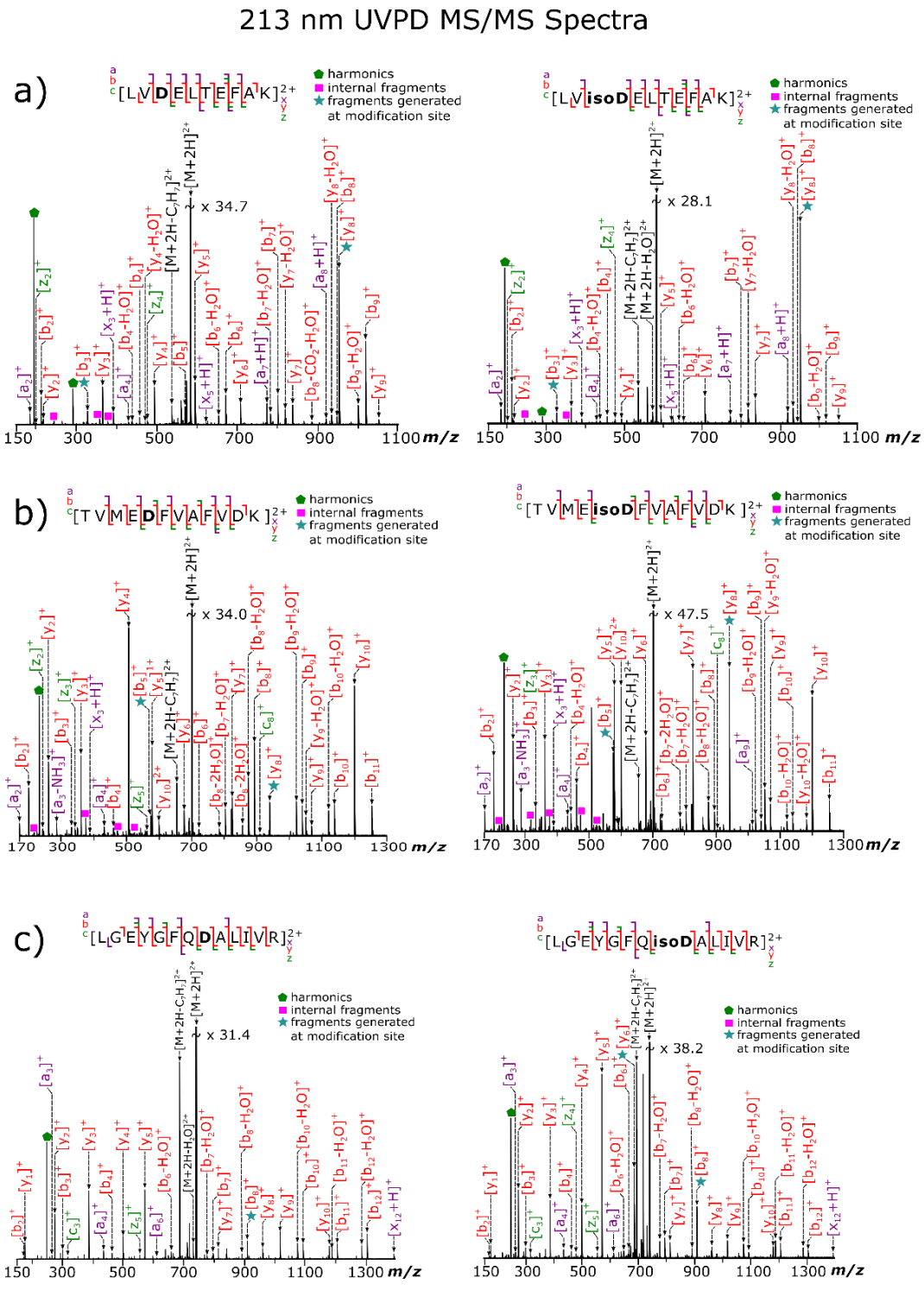


Figure 2. 16 (VII) Experimental 213 nm UVPD MS/MS spectra of synthetic isoD and D peptides a) [LVisoD/DELTEFAK]$^{2+}$ (peptide 1) b) [TVMEisoD/DFVAFVDK+2H]$^{2+}$ (peptide 2) and c) [LGEYGFQisoD/DALIVR+2H]$^{2+}$ (peptide 3).

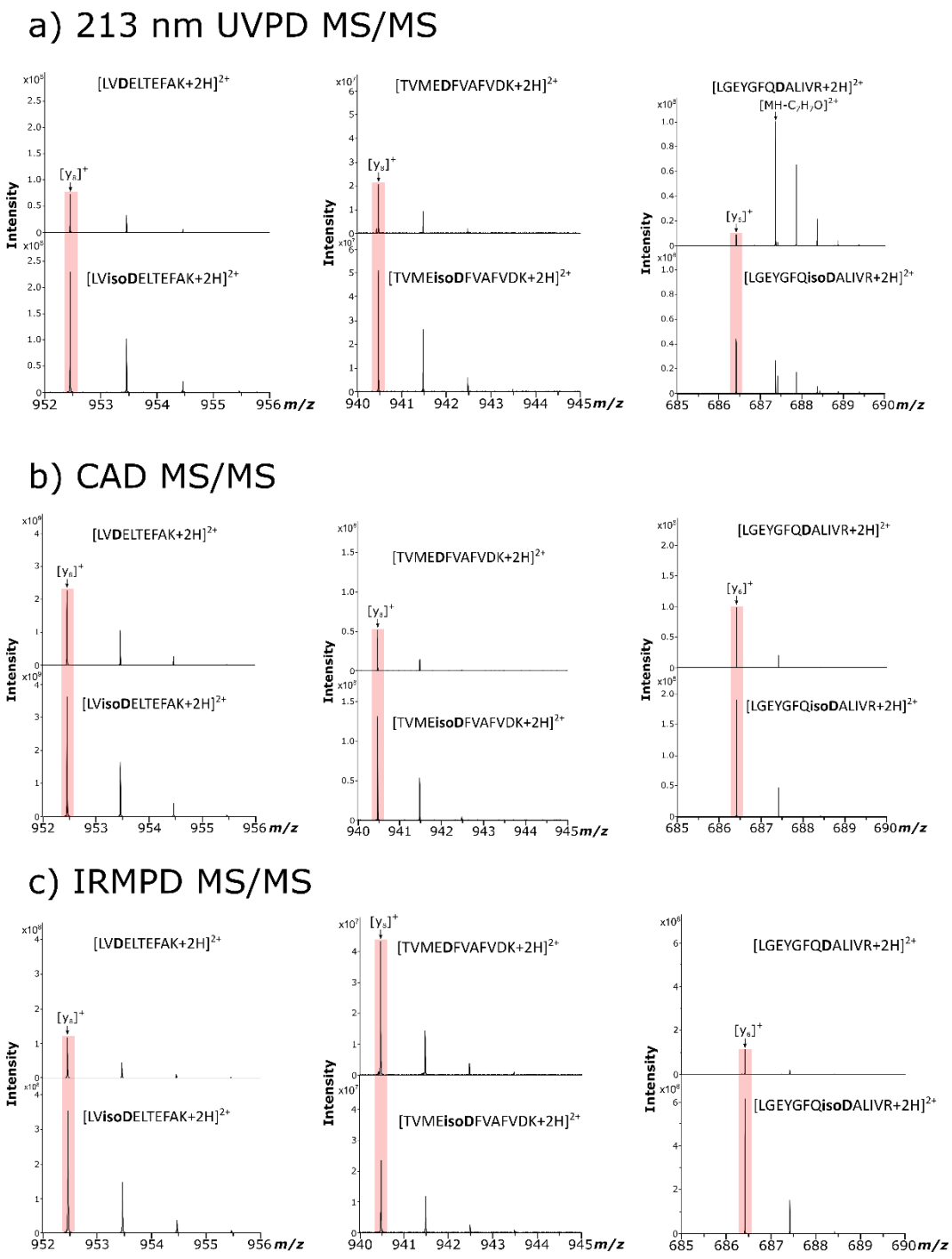


Figure 2. 17 Zoom in of a) 213 nm UVPD b) CAD and c) IRMPD MS/MS spectra for synthetic isoD and D peptides 1-3. The y fragment generated at the deamidation modification site is highlighted in red to show the difference in peak intensities between each isoD and D peptide.

CAD MS/MS could also generate higher y ion fragment at isoD deamidated site compared to D peptides (Figure 2.17B); the y ion intensities difference between D and isoD peptides obtained using CAD MS/MS, however, were ranged from 6- to 27-fold lower than the difference obtained from UVPD MS/MS, suggesting CAD may not be sensitive enough to provide the precise isoD percentage in a mixture solution compared to UVPD method. On the other hand, IRMPD MS/MS could not even generate a consistent intensity pattern to differentiate and quantify the D and isoD peptides (Figure 2.17C), indicating IRMPD is not suitable for D and isoD peptide differentiation.

Separation of deamidation products using nano-LC coupled to FT-ICR MS and quantification of isoD

Rather than various fragmentation methods, liquid chromatography, both HPLC as well as nLC, is commonly used for D and isoD differentiation and quantification.[31,32,78,79] The quantification results obtained from LC methods, however, has never compared with the results generated from various fragmentation methods. Herein, we compared the results of relative percentage of isoD in the tryptic digested BSA samples obtained using ECD/UVPD fragmentation and nLC separation.

To correctly assign each XIC peak obtained from the nLC experiment, auto ECD MS/MS was used to differentiate D and isoD peptides using the characteristic isoD peptide fragment (z_{n-i}-C_2O_2H); while high mass accuracy was used to differentiate deamidated and non-deamidated peptides (Figure 2.18). The elution order of non-deamidated, deamidated-D, and deamidated-isoD peptides did not always following the same trend (Figure 2.19), indicating nLC effective gradient was also optimised from 15 minutes to 90 minutes to achieve fully resolved XIC peak for the best D and isoD quantification (Figure 2.20). Based on the results, 30-minute effective gradient achieved the best separation results with the most time-cost effective, and this gradient would further be applied to the following experiments.

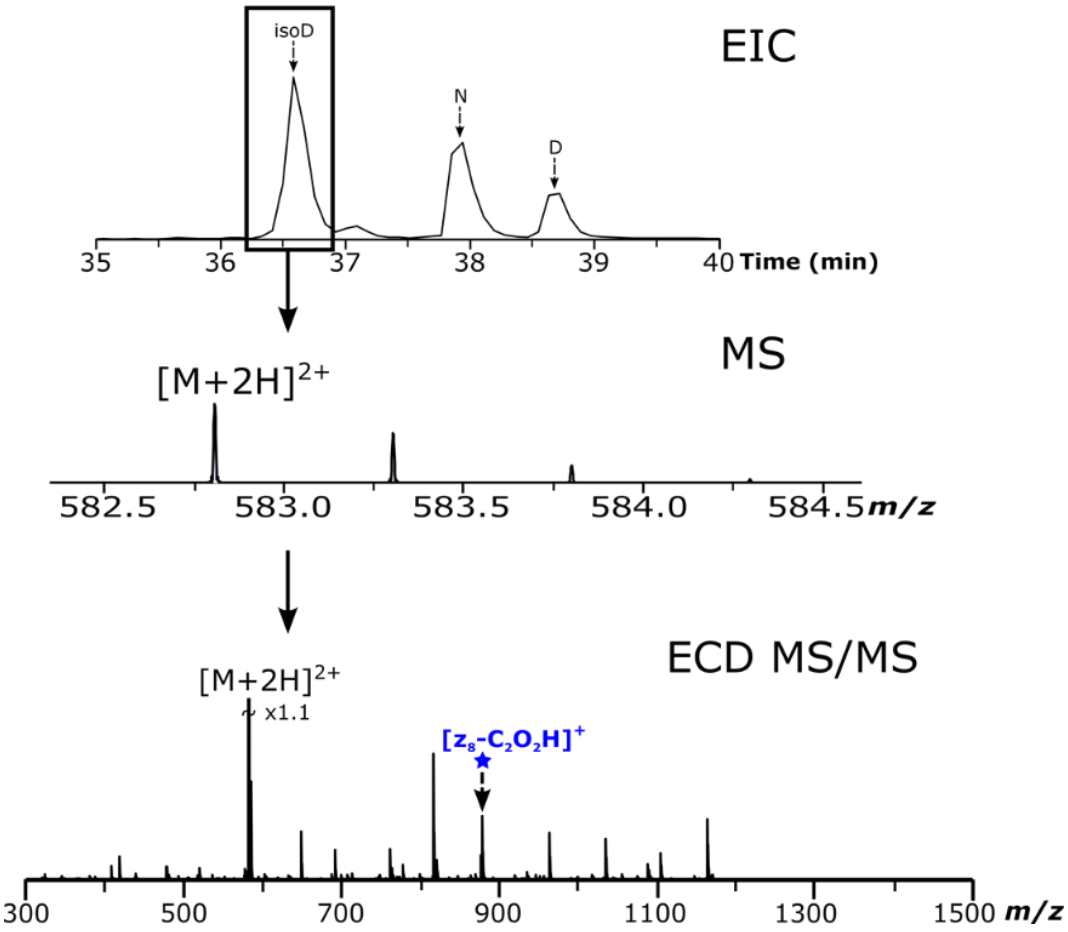


Figure 2. 18 Experimental workflow showing the extracted ion chromatogram for the deamidated BSA peptide $[LVN_{Deam}ELTEFAK]^{2+}$ following MS and ECD MS/MS analysis.

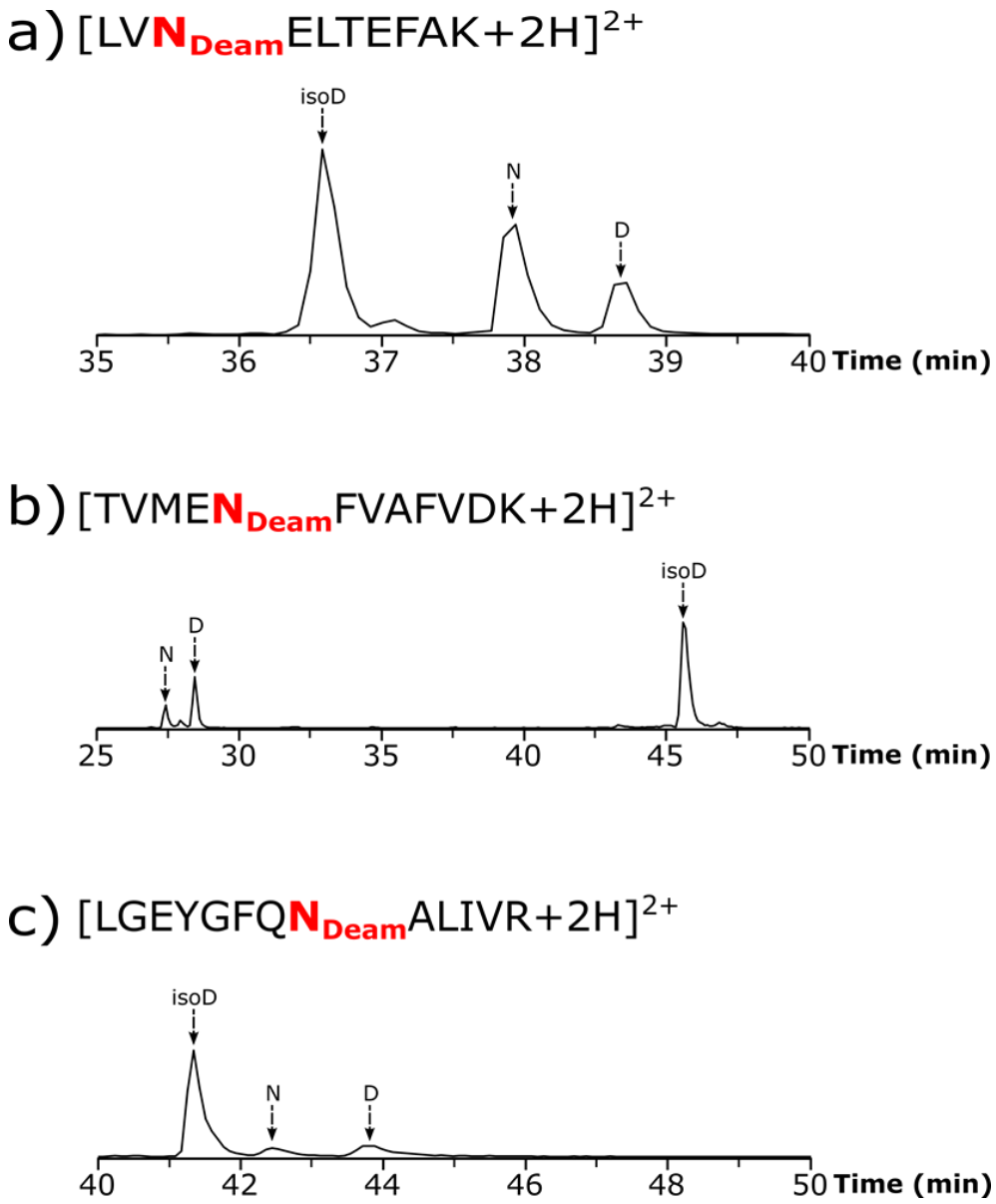


Figure 2. 19 nLC separation of the deamidation products for the three critical target peptides of BSA digest. nLC separation of the unmodified peptide (N) form and the deamidation products (D and isoD) of the three deamidated BSA tryptic peptides.

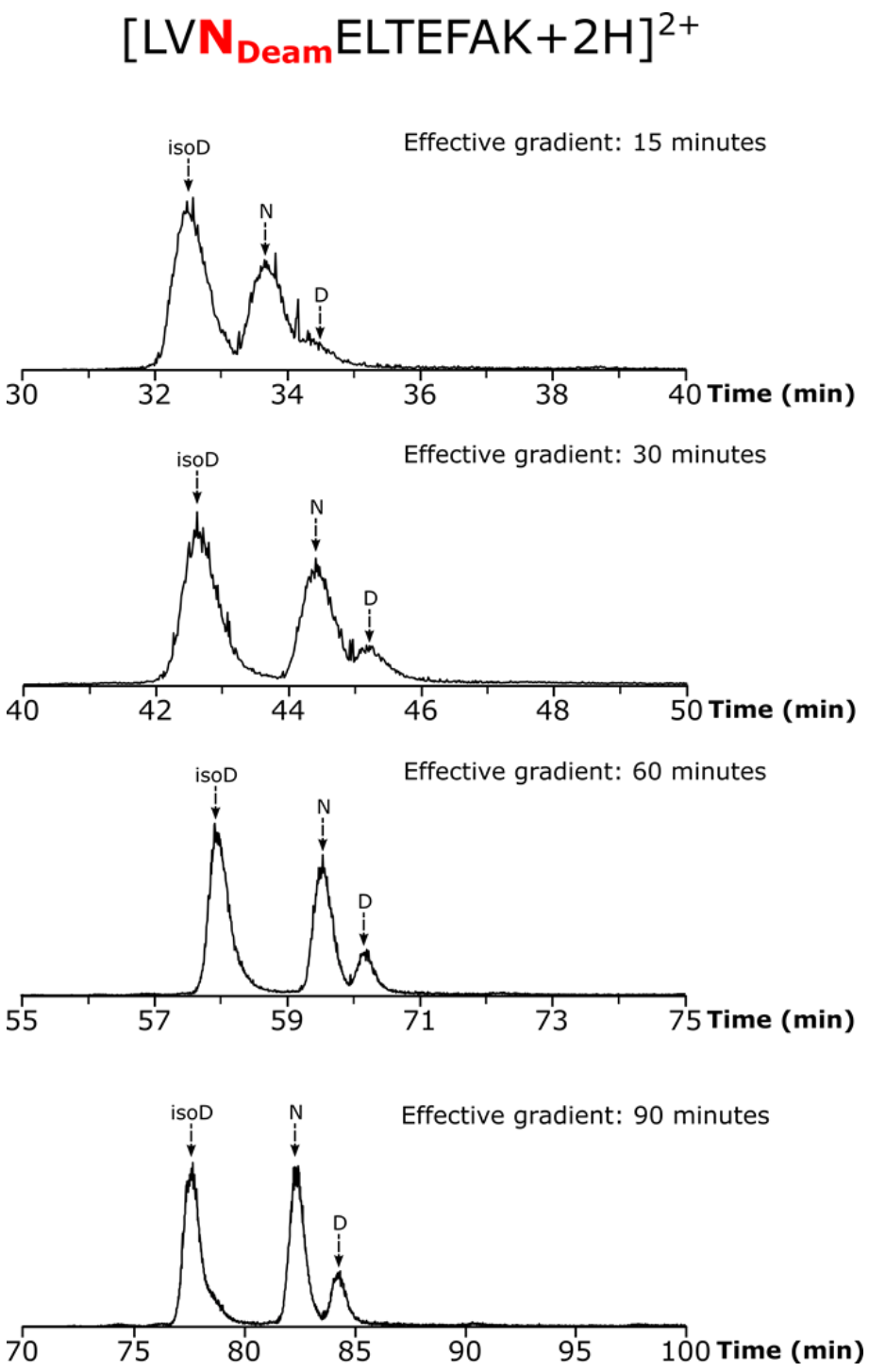


Figure 2. 20 The effect of the changes in the effective gradient time on the nLC separation of the unmodified peptide (N) form and the deamidation products (D and isoD) for the deamidated tryptic BSA peptide $[LVN_{Deam}ELTEFAK+2H]^{2+.}$

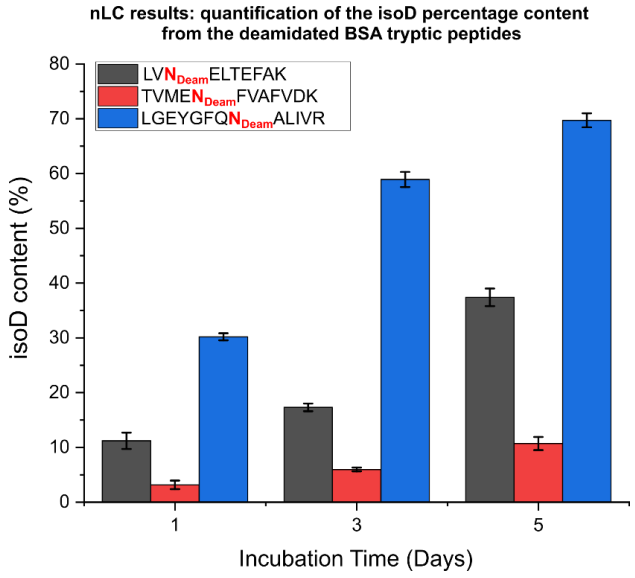


Figure 2. 21 nLC isoD percentage content quantification results for the target deamidated peptides.

The nanoLC isoD quantification results shown in Figure 2.21 and detailed in Table 2.9 follow the same trend as observed with ECD and UVPD MS/MS for isoD quantification, where an increase in the percentage isoD content in the deamidated peptides is detected with increasing incubation time. The nanoLC results are generally comparable with the ECD and UVPD MS/MS isoD quantification results. However, for some peptides such as peptide 1 and 2, on incubation day 3 and day 5, respectively, an underestimation of the isoD content in the deamidated peptides is observed with an approximately 9 % difference for peptide 1 and 4.7 % difference for peptide 2 when comparing the ECD MS/MS and nLC quantification results. This suggests that only using the XIC peak area as a measurement for the isoD percentage may not always be suitable as in this work some significant differences in the isoD quantification for the deamidated peptides 1-3 using this method are observed.

Table 2. 9 A summary table of the deamidated isoD-peptide percentages determined at different incubation times using nLC.

Incubation time (day)	direct infusion ECD			direct infusion 193 nm UVPD			online nano-LC		
	Peptide 1	Peptide 2	Peptide 3	Peptide 1	Peptide 2	Peptide 3	Peptide 1	Peptide 2	Peptide 3
	isoD (%)								
1	9.20 ± 1.1	3.54 ± 1.0	31.9 ± 1.9	9.06 ± 1.5	2.11 ± 0.6	30.2 ± 0.6	11.2 ± 1.5	3.14 ± 0.9	33.1 ± 1.1
3	24.7 ± 0.8	6.60 ± 0.7	58.6 ± 0.6	26.5 ± 0.8	7.03 ± 0.8	58.9 ± 1.4	17.3 ± 0.7	5.95 ± 1.2	61.8 ± 1.6
5	36.5 ± 1.1	13.4 ± 0.5	72.2 ± 1.6	34.7 ± 0.7	15.4 ± 0.8	69.7 ± 1.3	37.4 ± 1.6	10.7 ± 1.0	73.8 ± 0.9

2.5. Conclusions

In this chapter, a modified, improved quantification method for the quantification of deamidated isomeric products was demonstrated. A good linearity ($R^2 > 0.99$) was achieved in all calibration curves of synthetic peptides using ECD MS/MS regardless the magnitude order difference of the target fragment ions. We also showed UVPD MS/MS can be applied to differentiate and quantify the deamidated isomeric products without prior modification to the peptide sequence. It is a critical development for deamidated product quantification as a UVPD laser is easier to implement onto various types of instrument compared to the cathode used in ECD fragmentation, in the quantification of deamidation product is no longer limited by the instruments.

The relative quantification results obtained from ECD MS/MS, UVPD MS/MS, as well as nLC were also compared; and the result is that direct infusion ECD MS/MS provides the best results. This can be attributed to the fact that although UVPD MS/MS provide similar results for the isoD quantification of peptides 1-3, ECD MS/MS provides the confirmatory characteristic fragment for isoD peptides that is definitively absent for D peptides. This aids discrimination between isomeric peptides but also provides confidence in the quantification results as it is based on the presence of a characteristic fragment that is useful for quantification as it is unique only to the isoD peptides, whereas with the UVPD MS/MS quantification, dependence is placed on the significant intensity differences between the isoD and D peptides, which may be affected by something as simple as the condition of the instrument. Nano-LC on the other hand, has previously been used for isoD quantification. Based on the results of this work, similarities between the direct infusion approaches with tandem mass spectrometry are observed. However, some differences such as an underestimation of isoD % in peptides 1 and 2 were observed, when solely based on the XIC peak area for determining the isoD content in the deamidated peptides. Factors such as adequate baseline separation of the different isomeric products and peak broadening may affect the nano-LC quantification.

Overall, although synthetic standards are necessary but in comparison to other methods mentioned herein, direct infusion ECD MS/MS provides a fast and reliable approach for stringent quantification of isoD in the tryptic peptides of BSA.

2.6. References

(1) Robinson, A. B.; Rudd, C. J. In Curr. Top. Cell. Regul., Horecker, B. L.; Stadtman, E. R., Ed.^Eds.; Academic Press, 1974, p^pp 247-295.

(2) Groenen, P. J. T. A.; van Dongen, M. J. P.; Voorter, C. E. M.; Bloemendal, H.; de Jong, W. W. Age-dependent deamidation of αB-crystallin. FEBS Lett. 1993, 322 (1), 69-72.

(3) Harrington, V.; McCall, S.; Huynh, S.; Srivastava, K.; Srivastava, O. P. Crystallins in water soluble-high molecular weight protein fractions and water insoluble protein fractions in aging and cataractous human lenses. Mol. Vis. 2004, 10, 476-489.

(4) Wilmarth, P. A.; Tanner, S.; Dasari, S.; Nagalla, S. R.; Riviere, M. A.; Bafna, V.; Pevzner, P. A.; David, L. L. Age-Related Changes in Human Crystallins Determined from Comparative Analysis of Post-translational Modifications in Young and Aged Lens: Does Deamidation Contribute to Crystallin Insolubility? J. Proteome Res. 2006, 5 (10), 2554-2566.

(5) Lampi, K. J.; Wilmarth, P. A.; Murray, M. R.; David, L. L. Lens β-crystallins: The role of deamidation and related modifications in aging and cataract. Prog. Biophys. Mol. Biol. 2014, 115 (1), 21-31.

(6) Stadtman, E. R. Protein Modification in Aging. J. Gerontol. 1988, 43 (5), B112-B120.

(7) Lam, Y. P. Y.; Wootton, C. A.; Hands-Portman, I.; Wei, J.; Chiu, C. K. C.; Romero-Canelon, I.; Lermyte, F.; Barrow, M. P.; O'Connor, P. B. Does deamidation of islet amyloid polypeptide accelerate amyloid fibril formation? Chem. Commun. 2018, 54 (98), 13853-13856.

(8) Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.; Zanni, M. T. Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. J. Am. Chem. Soc. 2012, 134 (30), 12658-12667.

(9) Bagad, D. M.; Chowdhury, D.; Khan, Z. Towards understanding Alzheimer's Disease: An Overview. Res. J. Pharm. Biol. Chem. Sci. 2013, 4, 286-298.

(10) Nilsson, M. R.; Driscoll, M.; Raleigh, D. P. Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: implications for the study of amyloid formation. Protein Sci. 2002, 11 (2), 342-349.

(11) Robinson, N. E.; Robinson, A. B. Deamidation of human proteins. Proc. Natl. Acad. Sci. U.S.A. 2001, 98 (22), 12409.

(12) Robinson, N. E. Protein deamidation. Proc. Natl. Acad. Sci. U.S.A. 2002, 99 (8), 5283.

(13) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S. Characterization of Therapeutic Antibodies and Related Products. Anal. Chem. 2013, 85 (2), 715-736.

(14) Diepold, K.; Bomans, K.; Wiedmann, M.; Zimmermann, B.; Petzold, A.; Schlothauer, T.; Mueller, R.; Moritz, B.; Stracke, J. O.; Mølhøj, M.; Reusch, D.; Bulau, P. Simultaneous Assessment of Asp Isomerization and Asn Deamidation in Recombinant Antibodies by LC-MS following Incubation at Elevated Temperatures. PLoS One 2012, 7 (1), e30295.

(15) Eakin, C. M.; Miller, A.; Kerr, J.; Kung, J.; Wallace, A. Assessing analytical methods to monitor isoAsp formation in monoclonal antibodies. Front. Pharmacol. 2014, 5, 87-87.

(16) Wang, W.; Meeler, A. R.; Bergerud, L. T.; Hesselberg, M.; Byrne, M.; Wu, Z. Quantification and characterization of antibody deamidation by peptide mapping with mass spectrometry. Int. J. Mass spectrom. 2012, 312, 107-113.

(17) Cacia, J.; Keck, R.; Presta, L. G.; Frenz, J. Isomerization of an aspartic acid residue in the complementarity-determining regions of a recombinant antibody to human IgE: identification and effect on binding affinity. Biochemistry 1996, 35 (6), 1897-1903.

(18) Harris, R. J.; Kabakoff, B.; Macchi, F. D.; Shen, F. J.; Kwong, M.; Andya, J. D.; Shire, S. J.; Bjork, N.; Totpal, K.; Chen, A. B. Identification of multiple sources of charge heterogeneity in a recombinant antibody. J. Chromatogr. B 2001, 752 (2), 233-245.

(19) Geiger, T.; Clarke, S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J. Biol. Chem. 1987, 262 (2), 785-794.

(20) Clarke, S. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. Int. J. Pept. Protein Res. 1987, 30 (6), 808-821.

(21) Robinson, N. E.; Robinson, A. B. Molecular clocks. 2001, 98 (3), 944-949.

(22) Wright, H. T. Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit. Rev. Biochem. Mol. Biol. 1991, 26 (1), 1-52.

(23) Johnson, B. A.; Shirokawa, J. M.; Hancock, W. S.; Spellman, M. W.; Basa, L. J.; Aswad, D. W. Formation of Isoaspartate at Two Distinct Sites during in vitro Aging of Human Growth Hormone. J. Biol. Chem. 1989, 264 (24), 14262-14271.

(24) Capasso, S.; Kirby, A. J.; Salvadori, S.; Sica, F.; Zagari, A. Kinetics and mechanism of the reversible isomerization of aspartic acid residues in tetrapeptides. J. Chem. Soc., Perkin trans. 2 1995, (3), 437-442.

(25) Tyler-Cross, R.; Schirch, V. Effects of amino acid sequence, buffers, and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides. J. Biol. Chem. 1991, 266 (33), 22549-22556.

(26) Meinwald, Y. C.; Stimson, E. R.; Scheraga, H. A. Deamidation of the asparaginyl-glycyl sequence. Int. J. Pept. Protein Res. 1986, 28 (1), 79-84.

(27) Zabrouskov, V.; Han, X.; Welker, E.; Zhai, H.; Lin, C.; van Wijk, K. J.; Scheraga, H. A.; McLafferty, F. W. Stepwise deamidation of ribonuclease A at five sites determined by top down mass spectrometry. Biochemistry 2006, 45 (3), 987-992.

(28) Hurtado, P. P.; O'Connor, P. B. Differentiation of isomeric amino acid residues in proteins and peptides using mass spectrometry. Mass Spectrom. Rev. 2012, 31 (6), 609-625.

(29) Carlson, A. D.; Riggin, R. M. Development of Improved High-Performance Liquid Chromatography Conditions for Nonisotopic Detection of Isoaspartic Acid to Determine the Extent of Protein Deamidation. Anal. Biochem. 2000, 278 (2), 150-155.

(30) Hao, P.; Qian, J.; Dutta, B.; Cheow, E. S. H.; Sim, K. H.; Meng, W.; Adav, S. S.; Alpert, A.; Sze, S. K. Enhanced Separation and Characterization of Deamidated Peptides with RP-ERLIC-Based Multidimensional Chromatography Coupled with Tandem Mass Spectrometry. J. Proteome Res. 2012, 11 (3), 1804-1811.

(31) Badgett, M. J.; Boyes, B.; Orlando, R. The Separation and Quantitation of Peptides with and without Oxidation of Methionine and Deamidation of Asparagine Using Hydrophilic Interaction Liquid Chromatography with Mass Spectrometry (HILIC-MS). J. Am. Soc. Mass Spectrom. 2017, 28 (5), 818-826.

(32) De Boni, S.; Oberthür, C.; Hamburger, M.; Scriba, G. K. E. Analysis of aspartyl peptide degradation products by high-performance liquid chromatography and high-performance liquid chromatography-mass spectrometry. J. Chromatogr. A 2004, 1022 (1), 95-102.

(33) Winter, D.; Pipkorn, R.; Lehmann, W. D. Separation of peptide isomers and conformers by ultra performance liquid chromatography. J. Sep. Sci. 2009, 32 (8), 1111-1119.

(34) Krokhin, O. V.; Antonovici, M.; Ens, W.; Wilkins, J. A.; Standing, K. G. Deamidation of -Asn-Gly- Sequences during Sample Preparation for Proteomics: Consequences for MALDI and HPLC-MALDI Analysis. Anal. Chem. 2006, 78 (18), 6645-6650.

(35) Yang, H.; Fung, E. Y. M.; Zubarev, A. R.; Zubarev, R. A. Toward Proteome-Scale Identification and Quantification of Isoaspartyl Residues in Biological Samples. J. Proteome Res. 2009, 8 (10), 4615-4621.

(36) Ni, W.; Dai, S.; Karger, B. L.; Zhou, Z. S. Analysis of Isoaspartic Acid by Selective Proteolysis with Asp-N and Electron Transfer Dissociation Mass Spectrometry. Anal. Chem. 2010, 82 (17), 7485-7491.

(37) Chazin, W. J.; Koerdel, J.; Thulin, E.; Hofmann, T.; Drakenberg, T.; Forsen, S. Identification of an isoaspartyl linkage formed upon deamidation of bovine calbindin D9k and structural characterization by 2D proton NMR. Biochemistry 1989, 28 (21), 8646-8653.

(38) O'Connell, N. E.; Lelli, K.; Mann, R. S.; Palmer, A. G. Asparagine deamidation reduces DNA-binding affinity of the Drosophila melanogaster Scr homeodomain. FEBS Lett. 2015, 589 (21), 3237-3241.

(39) Di Donato, A.; Ciardiello, M. A.; de Nigris, M.; Piccoli, R.; Mazzarella, L.; D'Alessio, G. Selective deamidation of ribonuclease A. Isolation and characterization of the resulting isoaspartyl and aspartyl derivatives. J. Biol. Chem. 1993, 268 (7), 4745-4751.

(40) Liu, M.; Cheetham, J.; Cauchon, N.; Ostovic, J.; Ni, W.; Ren, D.; Zhou, Z. S. Protein Isoaspartate Methyltransferase-Mediated 18O-Labeling of Isoaspartic Acid for Mass Spectrometry Analysis. Anal. Chem. 2012, 84 (2), 1056-1062.

(41) Xiao, G.; Bondarenko, P. V.; Jacob, J.; Chu, G. C.; Chelius, D. 18O Labeling Method for Identification and Quantification of Succinimide in Proteins. Anal. Chem. 2007, 79 (7), 2714-2721.

(42) Alfaro, J. F.; Gillies, L. A.; Sun, H. G.; Dai, S.; Zang, T.; Klaene, J. J.; Kim, B. J.; Lowenson, J. D.; Clarke, S. G.; Karger, B. L.; Zhou, Z. S. Chemo-Enzymatic Detection of Protein Isoaspartate Using Protein Isoaspartate Methyltransferase and Hydrazine Trapping. Anal. Chem. 2008, 80 (10), 3882-3889.

(43) Castet, S.; Enjalbal, C.; Fulcrand, P.; Guichou, J. F.; Martinez, J.; Aubagnac, J. L. Characterization of Aspartic Acid and β-Aspartic Acid in Peptides by Fast-atom Bombardment Mass Spectrometry and Tandem Mass Spectrometry. Rapid Commun. Mass Spectrom. 1996, 10 (15), 1934-1938.

(44) Lehmann, W. D.; Schlosser, A.; Erben, G.; Pipkorn, R.; Bossemeyer, D.; Kinzel, V. Analysis of isoaspartate in peptides by electrospray tandem mass spectrometry. Protein Sci. 2000, 9 (11), 2260-2268.

(45) Gonzalez, L. J.; Shimizu, T.; Satomi, Y.; Betancourt, L.; Besada, V.; Padron, G.; Orlando, R.; Shirasawa, T.; Shimonishi, Y.; Takao, T. Differentiating alpha- and beta-aspartic acids by electrospray ionization and low-energy tandem mass spectrometry. Rapid Commun. Mass Spectrom. 2000, 14 (22), 2092-2102.

(46) Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. Protein Sci. 2005, 14 (2), 452-463.

(47) O'Connor, P. B.; Cournoyer, J. J.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. Differentiation of Aspartic and Isoaspartic Acids Using Electron Transfer Dissociation. J. Am. Soc. Mass Spectrom. 2006, 17 (1), 15-19.

(48) Schindler, P.; Müller, D.; Märki, W.; Grossenbacher, H.; Richter, W. J. Characterization of a β-Asp33 Isoform of Recombinant Hirudin Sequence Variant 1 by Low-energy Collision-induced Dissociation. 1996, 31 (9), 967-974.

(49) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. J. Am. Chem. Soc. 1998, 120 (13), 3265-3266.

(50) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation of Gaseous Multiply-Charged Proteins Is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity. J. Am. Chem. Soc. 1999, 121 (12), 2857-2862.

(51) Cournoyer, J. J.; Lin, C.; O'Connor, P. B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. Anal. Chem. 2006, 78 (4), 1264-1271.

(52) Brodbelt, J. S. Photodissociation mass spectrometry: new tools for characterization of biological molecules. Chem. Soc. Rev. 2014, 43 (8), 2757-2783.

(53) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. Chem. Rev. 2020, 120 (7), 3328-3380.

(54) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. Ultrafast Ultraviolet Photodissociation at 193 nm and its Applicability to Proteomic Workflows. J. Proteome Res. 2010, 9 (8), 4205-4214.

(55) Riggs, D. L.; Gomez, S. V.; Julian, R. R. Sequence and Solution Effects on the Prevalence of d-Isomers Produced by Deamidation. ACS Chem. Biol. 2017, 12 (11), 2875-2882.

(56) Sun, Q.; Yin, S.; Loo, J. A.; Julian, R. R. Radical Directed Dissociation for Facile Identification of Iodotyrosine Residues Using Electrospray Ionization Mass Spectrometry. Anal. Chem. 2010, 82 (9), 3826-3833.

(57) Ly, T.; Julian, R. R. Residue-Specific Radical-Directed Dissociation of Whole Proteins in the Gas Phase. J. Am. Chem. Soc. 2008, 130 (1), 351-358.

(58) Cournoyer, J. J.; Lin, C.; Bowman, M. J.; O'Connor, P. B. Quantitating the Relative Abundance of Isoaspartyl Residues in Deamidated Proteins by Electron Capture Dissociation. J. Am. Soc. Mass Spectrom. 2007, 18 (1), 48-56.

(59) Ivanov, D. G.; Indeykina, M. I.; Pekov, S. I.; Bugrova, A. E.; Kechko, O. I.; Iusupov, A. E.; Kononikhin, A. S.; Makarov, A. A.; Nikolaev, E. N.; Popov, I. A. Relative Quantitation of Beta-Amyloid Peptide Isomers with Simultaneous Isomerization of Multiple Aspartic Acid Residues by Matrix Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. J. Am. Soc. Mass Spectrom. 2020, 31 (1), 164-168.

(60) Pekov, S.; Indeykina, M.; Popov, I.; Kononikhin, A.; Bocharov, K.; Kozin, S. A.; Makarov, A. A.; Nikolaev, E. Application of MALDI-TOF/TOF-MS for relative quantitation of α- and β-Asp7 isoforms of amyloid-β peptide. 2018, 24 (1), 141-144.

(61) Pekov, S. I.; Ivanov, D. G.; Bugrova, A. E.; Indeykina, M. I.; Zakharova, N. V.; Popov, I. A.; Kononikhin, A. S.; Kozin, S. A.; Makarov, A. A.; Nikolaev, E. N. Evaluation of MALDI-TOF/TOF Mass Spectrometry Approach for Quantitative Determination of Aspartate Residue Isomerization in the Amyloid-β Peptide. J. Am. Soc. Mass Spectrom. 2019.

(62) Robinson, N. E.; Robinson, M. L.; Schulze, S. E. S.; Lai, B. T.; Gray, H. B. Deamidation of α-synuclein. Protein Sci. 2009, 18 (8), 1766-1773.

(63) Robinson, N. E.; Lampi, K. J.; McIver, R. T.; Williams, R. H.; Muster, W. C.; Kruppa, G.; Robinson, A. B. Quantitative measurement of deamidation in lens betaB2-crystallin and peptides by direct electrospray injection and fragmentation in a Fourier transform mass spectrometer. Mol. Vis. 2005, 11, 1211-1219.

(64) Yang, H.; Zubarev, R. A. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis 2010, 31 (11), 1764-1772.

(65) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. Differentiating N-Terminal Aspartic and Isoaspartic Acid Residues in Peptides. Anal. Chem. 2011, 83 (17), 6675-6682.

(66) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. Identification of aspartic and isoaspartic acid residues in amyloid beta peptides, including Abeta1-42, using electron-ion reactions. Anal. Chem. 2009, 81 (23), 9778-9786.

(67) Perez Hurtado, P.; O'Connor, P. B. Deamidation of Collagen. Anal. Chem. 2012, 84 (6), 3017-3025.

(68) Rockwood, A. L.; Erve, J. C. L. Mass Spectral Peak Distortion Due to Fourier Transform Signal Processing. J. Am. Soc. Mass Spectrom. 2014, 25 (12), 2163-2176.

(69) Morrison, L. J.; Brodbelt, J. S. Charge site assignment in native proteins by ultraviolet photodissociation (UVPD) mass spectrometry. Analyst 2016, 141 (1), 166-176.

(70) Guan, Z.; Kelleher, N. L.; O'Connor, P. B.; Aaserud, D. J.; Little, D. P.; McLafferty, F. W. 193 nm photodissociation of larger multiply-charged biomolecules. Int. J. Mass Spectrom. Ion Process. 1996, 157-158, 357-364.

(71) Cannon, J. R.; Martinez-Fonts, K.; Robotham, S. A.; Matouschek, A.; Brodbelt, J. S. Top-Down 193-nm Ultraviolet Photodissociation Mass Spectrometry for Simultaneous Determination of Polyubiquitin Chain Length and Topology. Anal. Chem. 2015, 87 (3), 1812-1820.

(72) Moon, J. H.; Yoon, S. H.; Kim, M. S. Photodissociation of singly protonated peptides at 193 nm investigated with tandem time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 2005, 19 (22), 3248-3252.

(73) Stiving, A. Q.; Harvey, S. R.; Jones, B. J.; Bellina, B.; Brown, J. M.; Barran, P. E.; Wysocki, V. H. Coupling 193 nm Ultraviolet Photodissociation and Ion Mobility for Sequence Characterization of Conformationally-Selected Peptides. J. Am. Soc. Mass Spectrom. 2020, 31 (11), 2313-2320.

(74) Robin, M. B. In Higher Excited States of Polyatomic Molecules, Robin, M. B., Ed.^Eds.; Academic Press, 1974, p^pp 7-68.

(75) R Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. J. Am. Soc. Mass Spectrom. 2017, 28 (9), 1823-1826.

(76) Choi, K. M.; Yoon, S. H.; Sun, M.; Oh, J. Y.; Moon, J. H.; Kim, M. S. Characteristics of Photodissociation at 193 nm of Singly Protonated Peptides Generated by Matrix-Assisted Laser Desorption Ionization (MALDI). J. Am. Soc. Mass Spectrom. 2006, 17 (12), 1643-1653.

(77) Hegyi, G.; Kardos, J.; Kovács, M.; Málnási-Csizmadia, A.; Nyitray, L.; Pál, G.; Radnai, L.; Reményi, A.; Venekei, I. Introduction to practical biochemistry.

(78) Sadakane, Y.; Yamazaki, T.; Nakagomi, K.; Akizawa, T.; Fujii, N.; Tanimura, T.; Kaneda, M.; Hatanaka, Y. Quantification of the isomerization of Asp residue in recombinant human αA-crystallin by reversed-phase HPLC. J. Pharm. Biomed. Anal. 2003, 30 (6), 1825-1833.

(79) Kern, W.; Mende, R.; Denefeld, B.; Sackewitz, M.; Chelius, D. Ion-pair reversed-phase high performance liquid chromatography method for the quantification of isoaspartic acid in a monoclonal antibody. J. Chromatogr. B 2014, 955-956, 26-33.

2.7 **Supplementary Information**

Table S2. 1 Volume percentage of synthetic peptide 2 in mixture for MS/MS experiments.

Calibration Point	**TVME<u>isoD</u>FVAFVDK (%)**	**TVME<u>D</u>FVAFVDK (%)**
1	0	100
2	20	80
3	40	60
4	60	40
5	80	20
6	100	0

Table S2. 2 Scaled percentage of each synthetic peptide for peptide 2 in mixtures with the account of peptide purity.

Calibration Point	**TVME<u>isoD</u>FVAFVDK (%)**	**TVME<u>D</u>FVAFVDK (%)**
1	0	97.5
2	19.2	78
3	38.4	58.5
4	57.6	39
5	76.8	19.5
6	96	0

Table S2. 3 Volume percentage of synthetic peptide 2 in mixture for MS/MS experiments.

Calibration Point	**LGEYGFQ<u>isoD</u>ALIVR (%)**	**LGEYGFQ<u>D</u>ALIVR (%)**
1	0	100
2	20	80
3	40	60
4	60	40
5	80	20
6	0	100

Table S2. 4 Scaled percentage of each synthetic peptide for peptide 3 in mixtures with the account of peptide purity.

Calibration Point	**LGEYGFQ<u>isoD</u>ALIVR (%)**	**LGEYGFQ<u>D</u>ALIVR (%)**
1	0	98.3
2	19.74	78.64
3	39.48	58.98
4	59.22	39.32
5	78.96	19.66
6	98.7	0

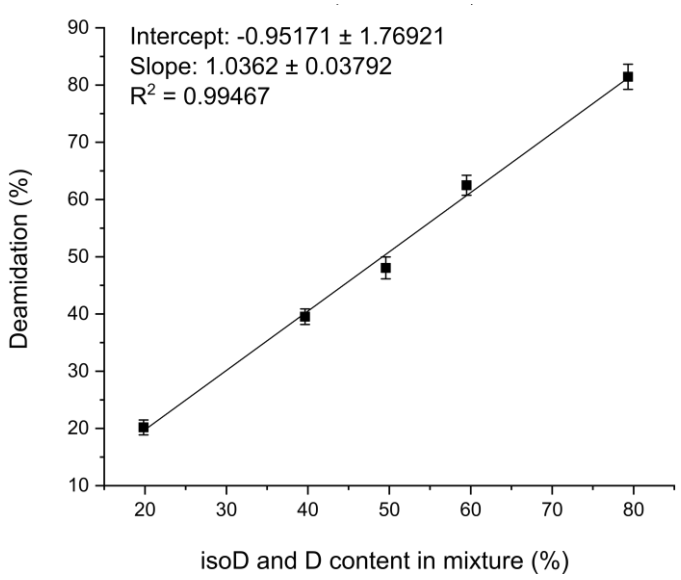


Figure S2. 1 Calibration curve of the calculated deamidation percentage for the synthetic BSA peptide mixtures of [LVNELTEFAK+2H]$^{2+}$, [LVDELTEFAK]$^{2+}$, and [LVisoDELTEFAK]$^{2+}$.

Table S2. 5 Deamidation half-times of the target peptides based on first order deamidation half-times of GlyXxxAsnYyyGly in days at pH 7.4, 37.0 ºC, 0.15 M Tris HCl.

Peptide sequence	**Deamidation half-time ($t_{1/2}$):**
LVNELTEFAK	64.8
TVMENFVAFDK	70.2
LGEYGFQNALIVR	25.8

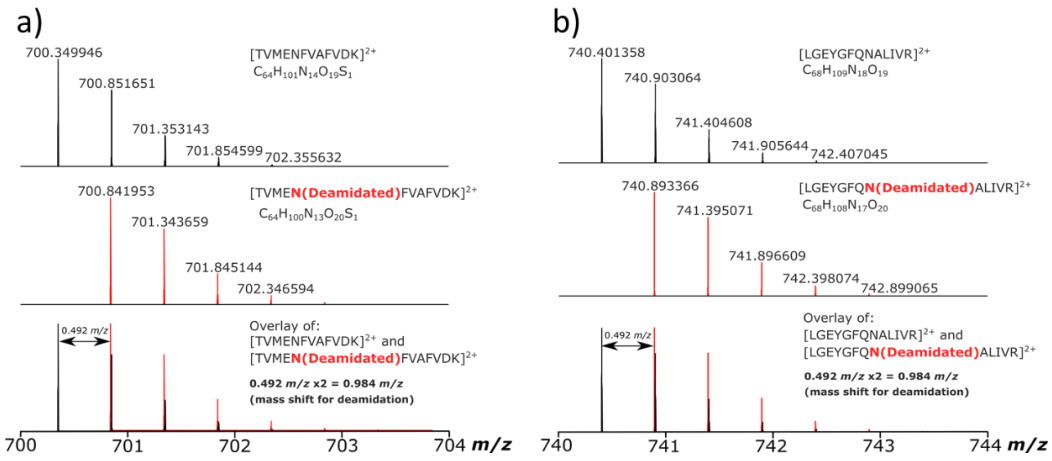


Figure S2. 2 Theoretical isotopic distributions of BSA peptides a) $[TVMENFVAFVDK+2H]^{2+}$ and b) $[LGEYGFQNALIVR+2H]^{2+}$ with 0 % deamidation, 100 % deamidation and an overlay of the non-deamidated and fully deamidated peptide MS spectrum.

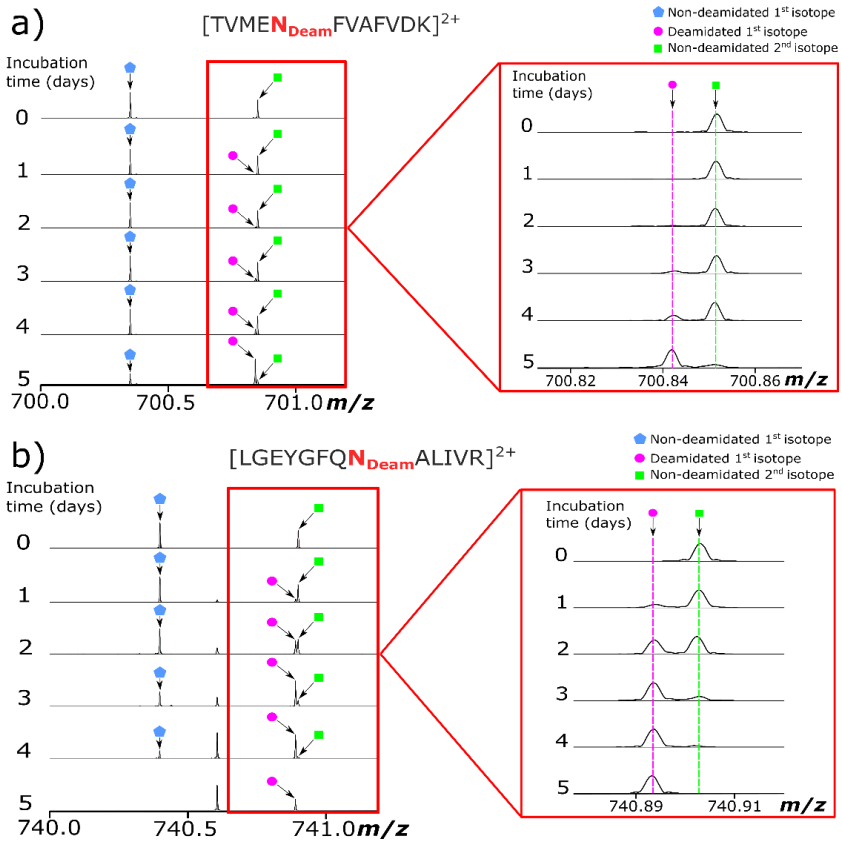


Figure S2. 3 Zoom in of MS spectrum of $[LVN_{Deam}ELTEFAK]^{2+}$ peptide. Theoretical isotopic distributions of BSA peptides a) $[TVMENFVAFVDK+2H]^{2+}$ and b) $[LGEYGFQNALIVR+2H]^{2+}$ with 0 % deamidation, 100 % deamidation and an overlay of the non-deamidated and fully deamidated peptide MS spectrum.

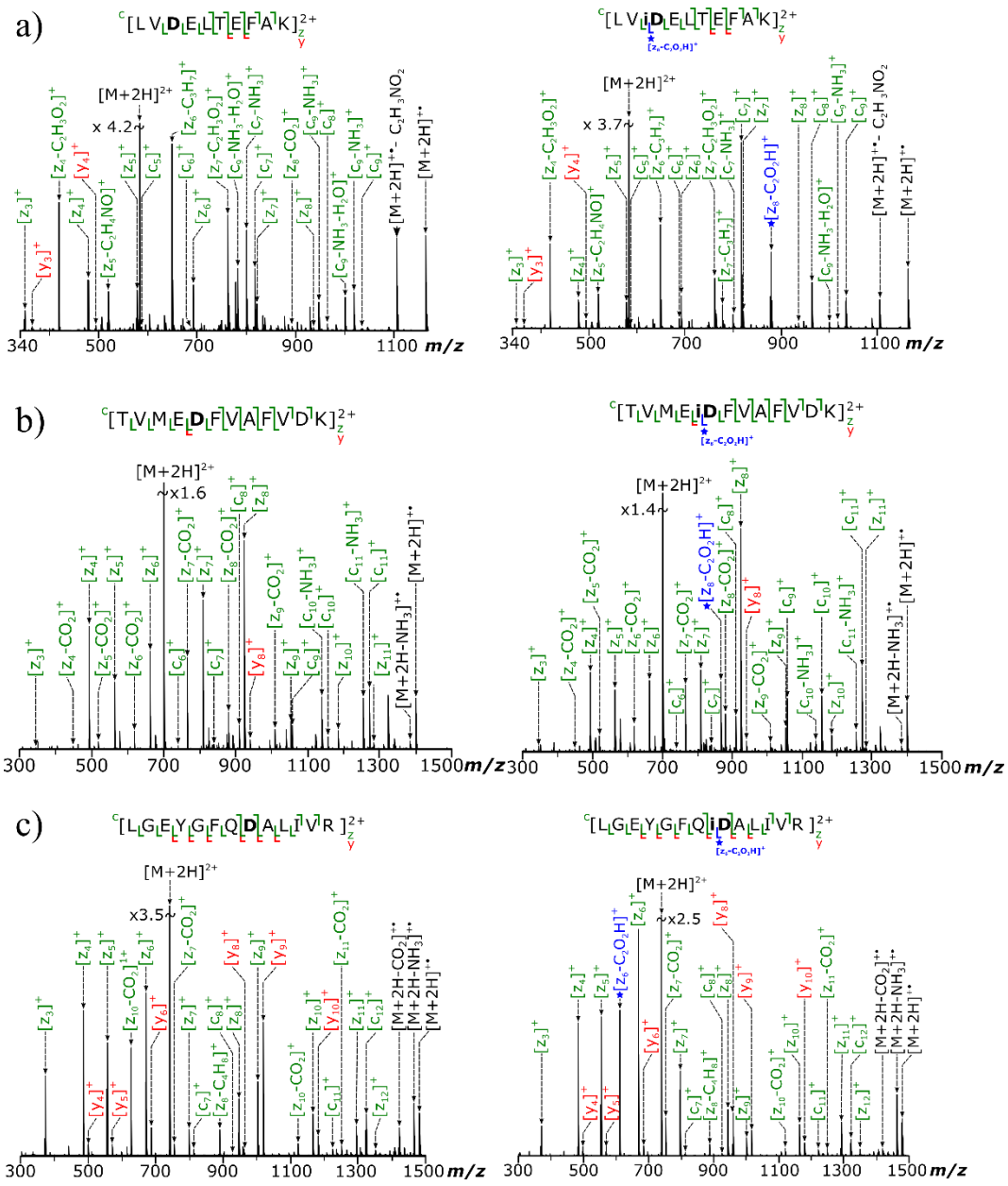


Figure S2. 4 Experimental ECD MS/MS spectra of synthetic isoD and D peptides for a) [LVisoD/DELTEFAK]$^{2+}$ (peptide 1) b) [TVMEisoD/DFVAFVDK+2H]$^{2+}$ (peptide 2) and c) [LGEYGFQisoD/DALIVR+2H]$^{2+}$ (peptide 3) with the specific z_{n-i}-C_2O_2H (z_{n-i}-57) fragment labelled in blue for in the ECD MS/MS spectra of the isoD containing peptides.

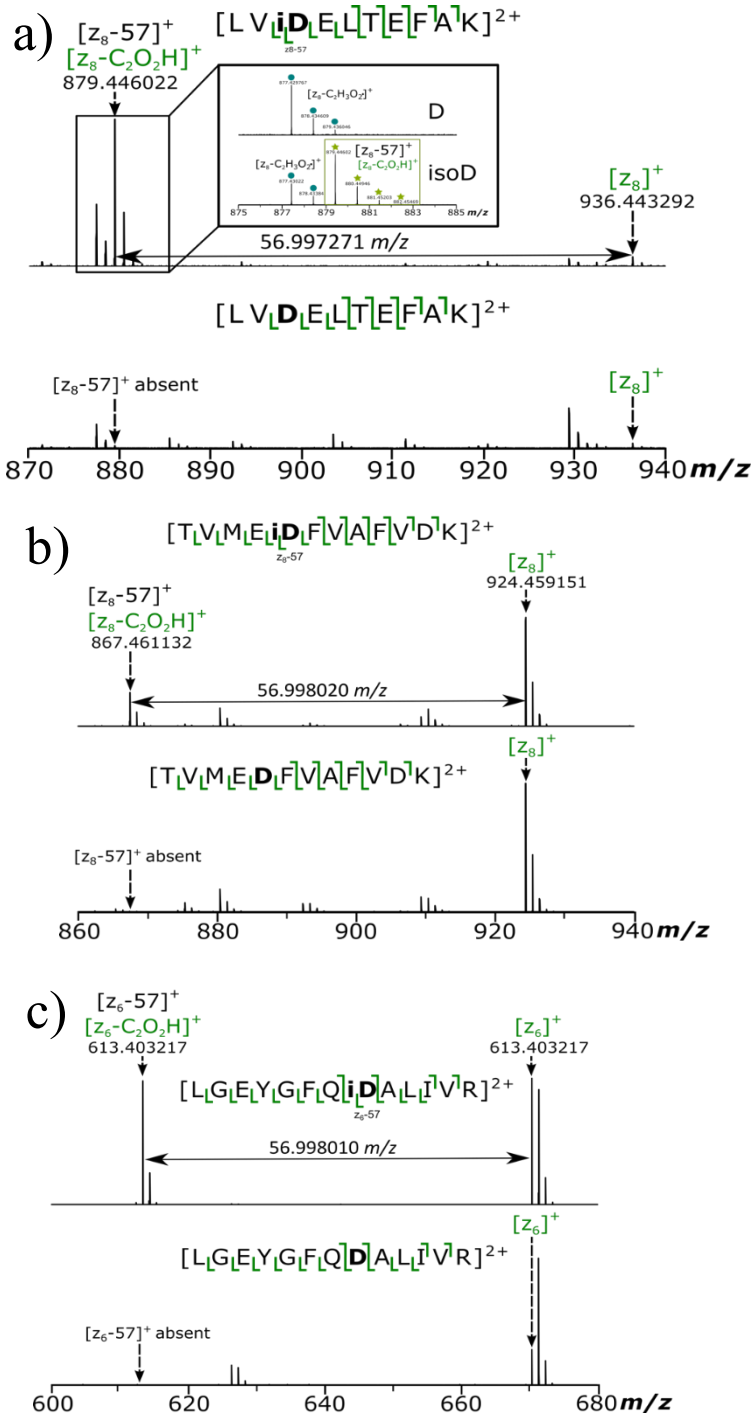


Figure S2. 5 Zoom in of ECD MS/MS spectra showing the absence and presence of specific z_{n-i}-57 fragment from D and isoD peptides respectively for a) [LV<u>isoD/D</u>ELTEFAK]$^{2+}$ (peptide 1) b) [TVME<u>isoD/D</u>FVAFVDK+2H]$^{2+}$ and c) [LGEYGFQ<u>isoD/D</u>ALIVR+2H]$^{2+}$.

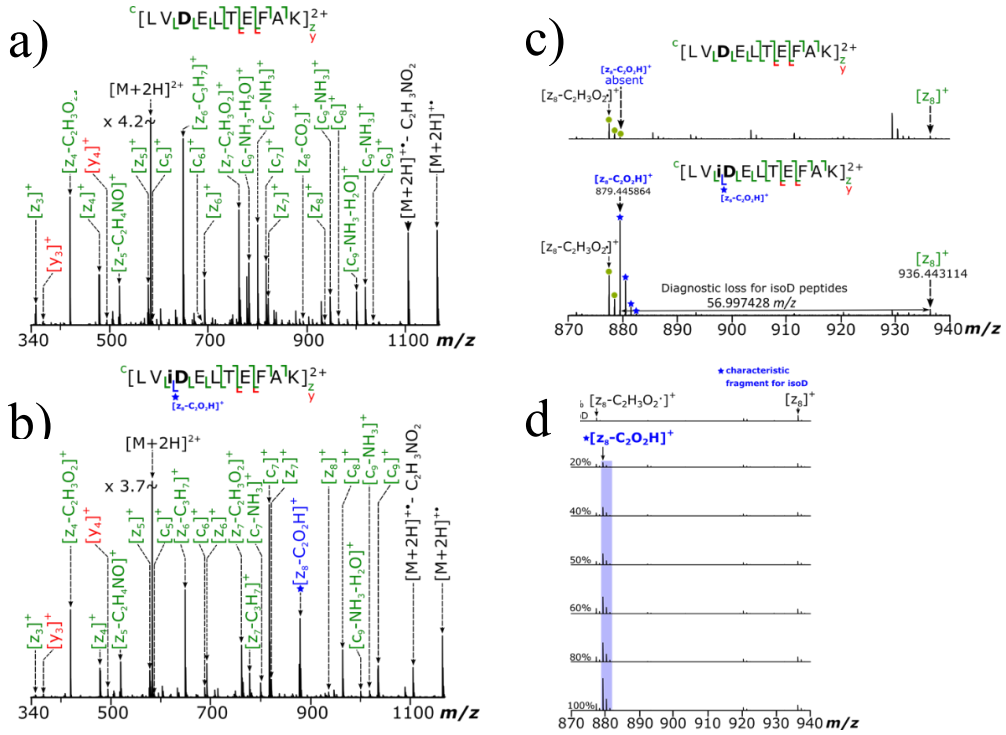


Figure S2. 6 ECD fragmentation spectra of a) [LVDELTEFAK+2H]$^{2+}$ and b) [LVisoDELTEFAK+2H]$^{2+}$ c) The *m/z* 870-940 region of isoD and D synthetic peptide mixtures. d) D and isoD forms of each peptide, in this case [LVDELTEFAK+2H]^{2} and [LVisoDELTEFAK+2H]$^{2+}$ were mixed to obtain mixtures, in which the isoD % varied from 0 to 100% in 20% increments.

Table S2. 6 Peak assignment table for the synthetic peptide [LVDELTEFAK+2H]$^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_3	1+	$C_{18}H_{27}N_3O_4$	349.199608	349.199558	-0.14
y_3-H	1+	$C_{18}H_{28}N_4O_4$	364.210507	364.210458	-0.13
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218332	365.218341	0.02
z_4-$C_2H_3O_2$.	1+	$C_{21}H_{31}N_4O_5$	419.228898	419.228877	-0.05
z_4	1+	$C_{23}H_{34}N_4O_7$	478.242202	478.242211	0.02
z.$_4$	1+	$C_{23}H_{35}N_4O_7$	479.250027	479.249996	-0.06
y_4-H	1+	$C_{23}H_{35}N_5O_7$	493.253101	493.252893	-0.42
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260926	494.260973	0.10
z_5-$C_3H_4O_2$	1+	$C_{24}H_{37}N_5O_7$	507.268751	507.26868	-0.14
z_5-$C_2H_3O_2$.	1+	$C_{25}H_{38}N_5O_7$	520.276027	520.276526	0.96
a_5+H	1+	$C_{25}H_{45}N_5O_8$	543.326266	543.326178	-0.16
z.$_5$-H_2O	1+	$C_{27}H_{40}N_5O_8$	562.287141	562.28735	0.37
z_5-CH_3.	1+	$C_{26}H_{38}N_5O_9$	564.266195	564.266489	0.52
c_5-NH_3	1+	$C_{26}H_{44}N_5O_9$	570.313356	570.313041	-0.55
z_5-H	1+	$C_{27}H_{40}N_5O_9$	578.282605	578.282644	0.07
z_5	1+	$C_{27}H_{41}N_5O_9$	579.289881	579.289753	-0.22
z.$_5$	1+	$C_{27}H_{42}N_5O_9$	580.297706	580.297859	0.26
MH	2+	$C_{53}H_{85}N_{11}O_{18}$	582.810988	582.811043	0.09
c_5	1+	$C_{26}H_{47}N_6O_9$	587.339905	587.340493	1.00
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308605	595.308337	-0.45
z_6-$C_3H_4O_2$	1+	$C_{30}H_{48}N_6O_8$	620.352815	620.353023	0.34
z_6-$C_2H_3O_2$.	1+	$C_{31}H_{49}N_6O_8$	633.360091	633.36049	0.63
z_6-C_3H_7	1+	$C_{25}H_{42}N_7O_{11}S_1$	649.31917	649.319255	0.13
c_6-NH_3	1+	$C_{30}H_{51}N_6O_{11}$	671.361035	671.360808	-0.34
z_6-H_2O	1+	$C_{33}H_{50}N_6O_9$	674.36338	674.363373	-0.01
c_6	1+	$C_{30}H_{54}N_7O_{11}$	688.387584	688.386905	-0.99
z_6	1+	$C_{33}H_{52}N_6O_{10}$	692.373945	692.373911	-0.05
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392669	708.392707	0.05
z_7-$C_3H_4O_2$	1+	$C_{35}H_{55}N_7O_{11}$	749.395409	749.39541	0.00
z_7-$C_2H_3O_2$.	1+	$C_{36}H_{56}N_7O_{11}$	762.402685	762.403211	0.69
c_7-H_2O	1+	$C_{35}H_{59}N_8O_{13}$	799.419613	799.419665	0.07
c_7-NH_3	1+	$C_{35}H_{58}N_7O_{14}$	800.403629	800.403643	0.02
c.$_7$	1+	$C_{35}H_{60}N_8O_{14}$	816.422353	816.422595	0.30
c_7	1+	$C_{35}H_{61}N_8O_{14}$	817.430178	817.429858	-0.39
z_7	1+	$C_{38}H_{59}N_7O_{13}$	821.416539	821.416254	-0.35
z_8-$C_3H_4O_2$	1+	$C_{39}H_{60}N_8O_{14}$	864.422353	864.422439	0.10
z_8-CO_2	1+	$C_{41}H_{64}N_8O_{14}$	43.989864	892.453619	-0.04
z_8	1+	$C_{42}H_{64}N_8O_{16}$	936.443483	936.442771	-0.76
c_8-NH_3	1+	$C_{44}H_{67}N_8O_{15}$	947.472043	947.47187	-0.18
c_8	1+	$C_{44}H_{70}N_9O_{15}$	964.498592	964.49859	0.00
c_9-NH_3	1+	$C_{47}H_{72}N_9O_{16}$	1018.509157	1018.508272	-0.87
c_9	1+	$C_{47}H_{75}N_{10}O_{16}$	1035.535706	1035.536345	0.62
MH	1+	$C_{53}H_{85}N_{11}O_{18}$	1164.614685	1164.613566	-0.96
Average error (ppm):					-0.02
Absolute average error (ppm):					0.32
Standard deviation (ppm):					0.31

Table S2. 7 Peak assignment table for the synthetic peptide [LVisoDELTEFAK+2H]$^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_3	1+	$C_{18}H_{27}N_3O_4$	349.199608	349.199922	0.90
y_3-H	1+	$C_{18}H_{28}N_4O_4$	364.210507	364.210702	0.54
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218332	365.218575	0.67
z_4-$C_2H_3O_2$.	1+	$C_{21}H_{31}N_4O_5$	419.228898	419.229022	0.30
z_4	1+	$C_{23}H_{34}N_4O_7$	478.242202	478.2422	0.00
$z._4$	1+	$C_{23}H_{35}N_4O_7$	479.250027	479.25007	0.09
y_4-H	1+	$C_{23}H_{35}N_5O_7$	493.253101	493.25293	-0.35
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260926	494.260883	-0.09
z_5-$C_3H_4O_2$	1+	$C_{24}H_{37}N_5O_7$	507.268751	507.268831	0.16
z_5-$C_2H_3O_2$.	1+	$C_{25}H_{38}N_5O_7$	520.276027	520.276581	1.06
a_5+H	1+	$C_{25}H_{45}N_5O_8$	543.326266	543.326046	-0.40
$z._5$-H_2O	1+	$C_{27}H_{40}N_5O_8$	562.287141	562.287315	0.31
z_5-CH_3.	1+	$C_{26}H_{38}N_5O_9$	564.266195	564.26653	0.59
c_5-NH_3	1+	$C_{26}H_{44}N_5O_9$	570.313356	570.313665	0.54
z_5-H	1+	$C_{27}H_{40}N_5O_9$	578.282605	578.282653	0.08
z_5	1+	$C_{27}H_{41}N_5O_9$	579.289881	579.289701	-0.31
$z._5$	1+	$C_{27}H_{42}N_5O_9$	580.297706	580.29753	-0.30
MH	2+	$C_{53}H_{85}N_{11}O_{18}$	582.810988	582.810938	-0.09
c_5	1+	$C_{26}H_{47}N_6O_9$	587.339905	587.339962	0.10
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308605	595.308413	-0.32
z_6-$C_3H_4O_2$	1+	$C_{30}H_{48}N_6O_8$	620.352815	620.352435	-0.61
z_6-$C_2H_3O_2$.	1+	$C_{31}H_{49}N_6O_8$	633.360091	633.360528	0.69
z_6-C_3H_7	1+	$C_{25}H_{42}N_7O_{11}S_1$	649.31917	649.319139	-0.05
c_6-NH_3	1+	$C_{30}H_{51}N_6O_{11}$	671.361035	671.360376	-0.98
z_6-H_2O	1+	$C_{33}H_{50}N_6O_9$	674.36338	674.36299	-0.58
c_6	1+	$C_{30}H_{54}N_7O_{11}$	688.387584	688.387536	-0.07
z_6	1+	$C_{33}H_{52}N_6O_{10}$	692.373945	692.373788	-0.23
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392669	708.392854	0.26
z_7-$C_3H_4O_2$	1+	$C_{35}H_{55}N_7O_{11}$	749.395409	749.395688	0.37
z_7-$C_2H_3O_2$.	1+	$C_{36}H_{56}N_7O_{11}$	762.403235	762.403256	0.03
c_7-H_2O	1+	$C_{35}H_{59}N_8O_{13}$	799.419613	799.419776	0.20
c_7-NH_3	1+	$C_{35}H_{58}N_7O_{14}$	800.403629	800.403585	-0.05
z_7-H_2O	1+	$C_{38}H_{57}N_7O_{12}$	803.405974	803.40669	0.89
$c._7$	1+	$C_{35}H_{60}N_8O_{14}$	816.422353	816.421923	-0.53
c_7	1+	$C_{35}H_{61}N_8O_{14}$	817.430178	817.430204	0.03
z_7	1+	$C_{38}H_{59}N_7O_{13}$	821.416539	821.416496	-0.05
z_8-$C_3H_4O_2$	1+	$C_{39}H_{60}N_8O_{14}$	864.422353	864.422629	0.32
z_8-C_2O_2H	1+	$C_{40}H_{63}N_8O_{14}$	879.445828	879.445864	0.04
z_8	1+	$C_{42}H_{64}N_8O_{16}$	936.443483	936.443114	-0.39
c_8-NH_3	1+	$C_{44}H_{67}N_8O_{15}$	947.472043	947.472118	0.08
c_8	1+	$C_{44}H_{70}N_9O_{15}$	964.498592	964.498421	-0.18
c_9-NH_3	1+	$C_{47}H_{72}N_9O_{16}$	1018.509157	1018.509687	0.52
c_9-CH_3.	1+	$C_{46}H_{72}N_{10}O_{16}$	1020.511682	1020.512361	0.67
$c._9$-H	1+	$C_{47}H_{73}N_{10}O_{16}$	1033.520056	1033.520917	0.83

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
c_9	1+	$C_{47}H_{75}N_{10}O_{16}$	1035.535706	1035.5357	-0.01
MH	1+	$C_{53}H_{85}N_{11}O_{18}$	1164.6147	1164.61448	-0.19
Average error (ppm):					0.10
Absolute average error (ppm):					0.36
Standard deviation (ppm):					0.30

Table S2. 8 Peak assignment table for the synthetic peptide [TVMEDFVAFVDK+2H]$^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_4-CO_2	1+	$C_{23}H_{36}N_4O_5$	448.268022	448.268167	0.32
z_4	1+	$C_{24}H_{36}N_4O_7$	492.257852	492.257842	-0.02
$z_{\cdot 4}$	1+	$C_{24}H_{37}N_4O_7$	493.265677	493.265743	0.13
z_5-CO_2	1+	$C_{26}H_{41}N_5O_6$	519.305136	519.305034	-0.20
z_5	1+	$C_{27}H_{41}N_5O_8$	563.294966	563.294964	0.00
$z_{\cdot 5}$	1+	$C_{27}H_{42}N_5O_8$	564.302791	564.302796	0.01
z_6	1+	$C_{32}H_{50}N_6O_9$	662.36338	662.363364	-0.02
$z_{\cdot 6}$	1+	$C_{32}H_{51}N_6O_9$	663.371205	663.371208	0.00
MH	2+	$C_{64}H_{99}N_{13}O_{20}S_1$	700.841938	700.842122	0.26
c_6	1+	$C_{32}H_{50}N_7O_{11}S_1$	740.328355	740.328285	-0.09
z_7-CO_2	1+	$C_{40}H_{59}N_7O_8$	765.441964	765.441859	-0.14
z_7	1+	$C_{41}H_{59}N_7O_{10}$	809.431794	809.431924	0.16
$c_{\cdot 7}$-NH_3	1+	$C_{37}H_{55}N_7O_{12}S_1$	821.362395	821.362828	0.53
y_7	1+	$C_{41}H_{61}N_8O_{10}$	825.450518	825.450045	-0.57
$c_{\cdot 7}$	1+	$C_{37}H_{58}N_8O_{12}S_1$	838.388944	838.388948	0.00
c_7	1+	$C_{37}H_{59}N_8O_{12}S_1$	839.396769	839.396209	-0.67
z_8-CO_2	1+	$C_{44}H_{64}N_8O_{11}$	880.468908	880.46856	-0.40
c_8-NH_3	1+	$C_{40}H_{61}N_8O_{13}S_1$	893.407334	893.407022	-0.35
$c_{\cdot 8}$	1+	$C_{40}H_{63}N_9O_{13}S_1$	909.426058	909.425495	-0.62
c_8	1+	$C_{40}H_{64}N_9O_{13}S_1$	910.433883	910.433743	-0.15
z_8	1+	$C_{45}H_{64}N_8O_{13}$	924.458738	924.459073	0.36
y_8	1+	$C_{45}H_{66}N_9O_{13}$	940.477462	940.476781	-0.72
z_9-CO_2	1+	$C_{49}H_{71}N_9O_{14}$	1009.511502	1009.510707	-0.79
b_9	1+	$C_{49}H_{70}N_9O_{14}S_1$	1040.475748	1040.475148	-0.58
z_9	1+	$C_{50}H_{71}N_9O_{16}$	1053.501332	1053.501322	-0.01
c_9	1+	$C_{49}H_{73}N_{10}O_{14}S_1$	1057.502297	1057.50166	-0.60
c_{10}-NH_3	1+	$C_{57}H_{85}N_{11}O_{17}S_1$	1139.544162	1139.544139	-0.02
c_{10}	1+	$C_{54}H_{82}N_{11}O_{15}S_1$	1156.570711	1156.570092	-0.54
z_{10}	1+	$C_{55}H_{80}N_{10}O_{17}S_1$	1184.541817	1184.54084	-0.82
c_{11}-NH_3	1+	$C_{58}H_{84}N_{11}O_{18}S_1$	1254.571106	1254.570958	-0.12
c_{11}	1+	$C_{58}H_{87}N_{12}O_{18}S_1$	1271.597655	1271.597612	-0.03

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_{11}	1+	$C_{60}H_{89}N_{11}O_{18}S_1$	1283.610231	1283.608863	-1.07
MH	1+	$C_{64}H_{98}N_{13}O_{20}S_1$	1400.67663	1400.676643	0.01
Average error (ppm):					-0.20
Absolute average error (ppm):					0.31
Standard deviation (ppm):					0.30

Table S2. 9 Peak assignment table for synthetic peptide [TVMEisoDFVAFVDK+2H]$^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_3	1+	$C_{15}H_{27}N_3O_6$	345.189438	345.189623	0.54
z_4-CO_2	1+	$C_{23}H_{36}N_4O_5$	448.268022	448.268064	0.09
z_4	1+	$C_{24}H_{36}N_4O_7$	492.257852	492.257846	-0.01
$z_{\cdot 4}$	1+	$C_{24}H_{37}N_4O_7$	493.265677	493.265739	0.13
z_5-CO_2	1+	$C_{26}H_{41}N_5O_6$	519.305136	519.305082	-0.10
z_5	1+	$C_{27}H_{41}N_5O_8$	563.294966	563.294951	-0.03
$z_{\cdot 5}$	1+	$C_{27}H_{42}N_5O_8$	564.302791	564.302725	-0.12
z_6	1+	$C_{32}H_{50}N_6O_9$	662.36338	662.363431	0.08
$z_{\cdot 6}$	1+	$C_{32}H_{51}N_6O_9$	663.371205	663.371277	0.11
MH	2+	$C_{64}H_{99}N_{13}O_{20}S_1$	700.841938	700.842177	0.34
c_6	1+	$C_{32}H_{50}N_7O_{11}S_1$	740.328355	740.327945	-0.55
z_7-CO_2	1+	$C_{40}H_{59}N_7O_8$	765.441964	765.441914	-0.07
z_7-H_2O	1+	$C_{41}H_{57}N_7O_9$	791.421229	791.42076	-0.59
z_7	1+	$C_{41}H_{59}N_7O_{10}$	809.431794	809.431895	0.12
$c_{\cdot 7}$-NH_3	1+	$C_{37}H_{55}N_7O_{12}S_1$	821.362395	821.362469	0.09
c_7-NH_3	1+	$C_{37}H_{56}N_7O_{12}S_1$	822.37022	822.370553	0.40
y_7	1+	$C_{41}H_{61}N_8O_{10}$	825.450518	825.450659	0.17
$c_{\cdot 7}$	1+	$C_{37}H_{58}N_8O_{12}S_1$	838.388944	838.388247	-0.83
c_7	1+	$C_{37}H_{59}N_8O_{12}S_1$	839.396769	839.39628	-0.58
z_8-C_2O_2H	1+	$C_{43}H_{63}N_8O_{11}$	867.4611215	867.461132	0.01
z_8-CO_2	1+	$C_{44}H_{64}N_8O_{11}$	880.468908	880.468777	-0.15
c_8-NH_3	1+	$C_{40}H_{61}N_8O_{13}S_1$	893.407334	893.407047	-0.32
$c_{\cdot 8}$	1+	$C_{40}H_{63}N_9O_{13}S_1$	909.426058	909.425484	-0.63
c_8	1+	$C_{40}H_{64}N_9O_{13}S_1$	910.433883	910.43376	-0.14
z_8	1+	$C_{45}H_{64}N_8O_{13}$	924.458738	924.459151	0.45
y_8	1+	$C_{45}H_{66}N_9O_{13}$	940.477462	940.476969	-0.52
z_9-CO_2	1+	$C_{49}H_{71}N_9O_{14}$	1009.511502	1009.511587	0.08
b_9	1+	$C_{49}H_{70}N_9O_{14}S_1$	1040.475748	1040.474646	-1.06
z_9	1+	$C_{50}H_{71}N_9O_{16}$	1053.501332	1053.50093	-0.38
c_9	1+	$C_{49}H_{73}N_{10}O_{14}S_1$	1057.502297	1057.501952	-0.33
c_{10}-NH_3	1+	$C_{57}H_{85}N_{11}O_{17}S_1$	1139.544162	1139.543194	-0.85
c_{10}	1+	$C_{54}H_{82}N_{11}O_{15}S_1$	1156.570711	1156.57034	-0.32
z_{10}	1+	$C_{55}H_{80}N_{10}O_{17}S_1$	1184.541817	1184.541128	-0.58
a_{11}+H	1+	$C_{57}H_{85}N_{11}O_{17}S_1$	1227.584016	1227.583997	-0.02

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
c_{11}-NH_3	1+	$C_{58}H_{84}N_{11}O_{18}S_1$	1254.571106	1254.570939	-0.13
c_{11}	1+	$C_{58}H_{87}N_{12}O_{18}S_1$	1271.597655	1271.597643	-0.01
z_{11}	1+	$C_{60}H_{89}N_{11}O_{18}S_1$	1283.610231	1283.609295	-0.73
MH	1+	$C_{64}H_{98}N_{13}O_{20}S_1$	1400.67663	1400.67662	-0.01
Average error (ppm):					-0.17
Absolute average error (ppm):					0.31
Standard deviation (ppm):					0.28

Table S2. 10 Peak assignment table for the synthetic peptide $[LGEYGFQ\underline{D}ALIVR+2H]^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_5	2+	$C_{26}H_{50}N_7O_6$	278.19058	278.190368	-0.76
z_6-H	2+	$C_{30}H_{54}N_8O_9$	335.20014	335.200054	-0.26
z_6	2+	$C_{30}H_{55}N_8O_9$	335.704052	335.703996	-0.17
z_4	2+	$C_{23}H_{44}N_6O_5$	484.33677	484.336446	-0.67
$z._4$	1+	$C_{23}H_{45}N_6O_5$	485.344595	485.344353	-0.50
y_4	1+	$C_{23}H_{46}N_7O_5$	500.355494	500.355296	-0.40
z_5	1+	$C_{26}H_{49}N_7O_6$	555.373884	555.373879	-0.01
$z._5$	1+	$C_{26}H_{50}N_7O_6$	556.381709	556.381672	-0.07
y_5	1+	$C_{26}H_{51}N_8O_6$	571.392608	571.392362	-0.43
z_6-CO_2	1+	$C_{29}H_{54}N_8O_7$	626.410998	626.411271	0.44
z_6	1+	$C_{30}H_{54}N_8O_9$	670.400828	670.40126	0.64
$z._6$	1+	$C_{30}H_{55}N_8O_9$	671.408653	671.409255	0.90
y_6	1+	$C_{30}H_{56}N_9O_9$	686.419552	686.419613	0.09
MH	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893368	740.894034	0.90
z_7-CO_2	1+	$C_{34}H_{62}N_{10}O_9$	754.469576	754.469308	-0.36
z_7	1+	$C_{35}H_{62}N_{10}O_{11}$	798.459406	798.459387	-0.02
c_7	1+	$C_{38}H_{54}N_9O_{11}$	812.393732	812.393678	-0.07
z_8 -C_4H_8	1+	$C_{40}H_{63}N_{11}O_{12}$	889.46522	889.464993	-0.26
c_8	1+	$C_{42}H_{59}N_{10}O_{14}$	927.420676	927.420471	-0.22
z_8	1+	$C_{44}H_{71}N_{11}O_{12}$	945.52782	945.528234	0.44
y_8	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546544	961.546332	-0.22
z_9	1+	$C_{46}H_{74}N_{12}O_{13}$	1002.549284	1002.549729	0.44
y_9	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.568008	1018.568746	0.72
z_{10}-CO_2	1+	$C_{54}H_{83}N_{13}O_{13}$	1121.622783	1121.623414	0.56
z_{10}	1+	$C_{55}H_{83}N_{13}O_{15}$	1165.612613	1165.613072	0.39
y_{10}	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631337	1181.631773	0.37
c_{11}	1+	$C_{57}H_{86}N_{13}O_{17}$	1224.625918	1224.626771	0.70
z_{11}-CO_2	1+	$C_{59}H_{90}N_{14}O_{16}$	1250.665377	1250.665811	0.35

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_{11}	1+	$C_{60}H_{90}N_{14}O_{18}$	1294.655207	1294.655618	0.32
z_{12}	1+	$C_{62}H_{93}N_{15}O_{19}$	1351.676671	1351.676272	-0.30
MH	1+	$C_{68}H_{106}N_{17}O_{20}$	1480.7795	1480.77951	0.01
Average error (ppm):					0.08
Absolute average error (ppm):					0.39
Standard deviation (ppm):					0.25

Table S2. 11 Peak assignment table for synthetic peptide [LGEYGFQisoDALIVR+2H]$^{2+}$ ECD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_4	2+	$C_{23}H_{45}N_6O_5$	242.672023	242.671813	-0.87
z_5	2+	$C_{26}H_{50}N_7O_6$	278.19058	278.190449	-0.47
z_6-H	2+	$C_{30}H_{54}N_8O_9$	335.20014	335.200147	0.02
z_6	2+	$C_{30}H_{55}N_8O_9$	335.704052	335.703961	-0.27
z_7-H	2+	$C_{35}H_{62}N_{10}O_{11}$	399.229429	399.22953	0.25
z_7	2+	$C_{35}H_{63}N_{10}O_{11}$	399.733341	399.733434	0.23
z_8-H	2+	$C_{44}H_{71}N_{11}O_{12}$	472.763636	472.76395	0.66
z_4	2+	$C_{23}H_{44}N_6O_5$	484.33677	484.336334	-0.90
$z\cdot_4$	1+	$C_{23}H_{45}N_6O_5$	485.344595	485.344265	-0.68
y_4	1+	$C_{23}H_{46}N_7O_5$	500.355494	500.355051	-0.89
z_5	1+	$C_{26}H_{49}N_7O_6$	555.373884	555.373804	-0.14
$z\cdot_5$	1+	$C_{26}H_{50}N_7O_6$	556.381709	556.381618	-0.16
y_5	1+	$C_{26}H_{51}N_8O_6$	571.392608	571.392172	-0.76
z_6-C_2O_2H	1+	$C_{28}H_{53}N_8O_7$	613.4032155	613.403217	0.00
z_6-CO_2	1+	$C_{29}H_{54}N_8O_7$	626.410998	626.410409	-0.94
z_6	1+	$C_{30}H_{54}N_8O_9$	670.400828	670.401226	0.59
$z\cdot_6$	1+	$C_{30}H_{55}N_8O_9$	671.408653	671.409126	0.70
y_6-H	1+	$C_{30}H_{55}N_9O_9$	685.411727	685.411117	-0.89
y_6	1+	$C_{30}H_{56}N_9O_9$	686.419552	686.419368	-0.27
MH	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893368	740.893966	0.81
z_7-CO_2	1+	$C_{34}H_{62}N_{10}O_9$	754.469576	754.469688	0.15
z_7	1+	$C_{35}H_{62}N_{10}O_{11}$	798.459406	798.459638	0.29
c_7	1+	$C_{38}H_{54}N_9O_{11}$	812.393732	812.393057	-0.83
z_8 -C_4H_8	1+	$C_{40}H_{63}N_{11}O_{12}$	889.46522	889.465466	0.28
c_8	1+	$C_{42}H_{59}N_{10}O_{14}$	927.420676	927.420755	0.09
z_8	1+	$C_{44}H_{71}N_{11}O_{12}$	945.52782	945.52797	0.16
y_8	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546544	961.546678	0.14
z_9	1+	$C_{46}H_{74}N_{12}O_{13}$	1002.549284	1002.549394	0.11
y_9	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.568008	1018.568132	0.12
z_{10}-CO_2	1+	$C_{54}H_{83}N_{13}O_{13}$	1121.622783	1121.622585	-0.18
z_{10}	1+	$C_{55}H_{83}N_{13}O_{15}$	1165.612613	1165.612513	-0.09
y_{10}	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631337	1181.631118	-0.19
c_{11}	1+	$C_{57}H_{86}N_{13}O_{17}$	1224.625918	1224.62658	0.54

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_{11}-CO_2	1+	$C_{59}H_{90}N_{14}O_{16}$	1250.665377	1250.66603	0.52
z_{11}	1+	$C_{60}H_{90}N_{14}O_{18}$	1294.655207	1294.655153	-0.04
z_{12}	1+	$C_{62}H_{93}N_{15}O_{19}$	1351.676671	1351.676688	0.01
MH	1+	$C_{68}H_{106}N_{17}O_{20}$	1480.7795	1480.779542	0.03
Average error (ppm):					-0.08
Absolute average error (ppm):					0.37
Standard deviation (ppm):					0.31

Table S2. 12 Peak assignment table for the ECD MS/MS spectrum of the peptide $[LVN_{Deam}ELTEFAK+2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_3	1+	$C_{18}H_{27}N_3O_4$	349.199608	349.199608	0.00
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218332	365.218350	0.05
z_4-$C_2H_3O_2$.	1+	$C_{21}H_{31}N_4O_5$	419.228898	419.228917	0.05
z_4	1+	$C_{23}H_{35}N_4O_7$	479.250027	479.249982	-0.09
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260926	494.261233	0.62
z_5	1+	$C_{27}H_{41}N_5O_9$	579.289881	579.289880	0.00
MH	2+	$C_{53}H_{88}N_{12}O_{17}$	582.318972	582.318919	-0.09
MH+Deam	2+	$C_{53}H_{87}N_{11}O_{18}$	582.810981	582.811032	0.09
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308605	595.308693	0.15
z_6-C_3H_7	1+	$C_{30}H_{45}N_6O_{10}$	649.319170	649.319208	0.06
c_6	1+	$C_{30}H_{55}N_8O_{10}$	687.403568	687.404202	0.92
z_6	1+	$C_{33}H_{52}N_6O_{10}$	692.373945	692.373947	0.00
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392669	708.392726	0.08
z_7-$C_2H_3O_2$	1+	$C_{38}H_{59}N_7O_{13}$	762.403235	762.403477	0.32
z_7-C_4H_8	1+	$C_{34}H_{51}N_7O_{13}$	765.353939	765.354003	0.08
z_7-C_3H_7.	1+	$C_{35}H_{52}N_7O_{13}$	778.361215	778.361858	0.83
c_7+Deam-H_2O	1+	$C_{35}H_{59}N_8O_{13}$	799.419610	799.419789	0.22
c_7+Deam	1+	$C_{35}H_{61}N_8O_{14}$	817.430178	817.430300	0.15
z_7	1+	$C_{38}H_{59}N_7O_{13}$	821.416539	821.416646	0.13
y_7	1+	$C_{38}H_{61}N_8O_{13}$	837.435263	837.435447	0.22
VNELTEFA-CO	1+	$C_{40}H_{62}N_9O_{13}$	876.446160	876.446607	0.51
z_8-C_2O_2H+Deam	1+	$C_{40}H_{63}N_8O_{14}$	879.445828	879.445860	0.04
z_8	1+	$C_{42}H_{65}N_9O_{15}$	935.459467	935.459112	-0.38
z_8+Deam	1+	$C_{42}H_{64}N_8O_{16}$	936.443479	936.443572	0.10

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_8	1+	$C_{44}H_{68}N_9O_{14}$	946.488027	946.488175	0.16
c_8	1+	$C_{44}H_{71}N_{10}O_{14}$	963.514573	963.514547	-0.03
c_8+Deam	1+	$C_{44}H_{70}N_9O_{15}$	964.498592	964.498894	0.31
b_9-H_2O+Deam	1+	$C_{47}H_{70}N_9O_{15}$	1000.498589	1000.497963	-0.63
b_9	1+	$C_{47}H_{73}N_{10}O_{15}$	1017.525141	1017.525259	0.12
b_9+Deam	1+	$C_{47}H_{72}N_9O_{16}$	1018.509154	1018.509938	0.77
c_9	1+	$C_{47}H_{76}N_{11}O_{15}$	1034.551690	1034.551498	-0.19
c_9+Deam	1+	$C_{47}H_{75}N_{10}O_{16}$	1035.535706	1035.535984	0.27
MH-NH_3-CO_2	1+	$C_{50}H_{78}N_{11}O_{17}$	1104.557717	1104.557438	-0.25
MH	1+	$C_{53}H_{87}N_{12}O_{17}$	1163.630669	1163.630691	0.02
MH+Deam	1+	$C_{53}H_{86}N_{11}O_{18}$	1164.614685	1164.615067	0.33
Average error (ppm):					0.14
Absolute average error (ppm):					0.24
Standard deviation (ppm):					0.25

Table S2. 13 Peak assignment table for the ECD MS/MS spectrum of the peptide $[TVMEN_{Deam}FVAFVDK +2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_4+H	1+	$C_{24}H_{37}N_4O_7$	493.265676	493.265702	0.05
z_5	1+	$C_{27}H_{41}N_5O_8$	563.294965	563.295562	1.06
z_5 + H	1+	$C_{27}H_{42}N_5O_8$	564.302791	564.302795	0.01
z_6	1+	$C_{32}H_{50}N_6O_9$	662.363379	662.363333	-0.07
y_6-H	1+	$C_{32}H_{51}N_7O_9$	677.374278	677.37384	-0.65
y_6	1+	$C_{32}H_{52}N_7O_9$	678.382103	678.382077	-0.04
MH	2+	$C_{64}H_{99}N_{14}O_{19}S_1$	700.349947	700.349967	0.03
MH+Deam	2+	$C_{64}H_{98}N_{13}O_{20}S_1$	700.841938	700.842516	0.82
z_7-CO_2	1+	$C_{40}H_{59}N_7O_8$	765.441963	765.442035	0.09
z_7	1+	$C_{41}H_{59}N_7O_{10}$	809.431793	809.431808	0.02
y_7-H	1+	$C_{41}H_{60}N_8O_{10}$	824.442692	824.441861	-1.01
z_8-C_2O_2H+Deam	1+	$C_{43}H_{63}N_8O_{11}$	867.4611215	867.460189	-1.07
c_8-H	1+	$C_{40}H_{64}N_{10}O_{12}S_1$	908.442040	908.441838	-0.22
z_8	1+	$C_{45}H_{65}N_9O_{12}$	923.474720	923.475067	0.38
z_8+Deam	1+	$C_{45}H_{64}N_8O_{13}$	924.458738	924.457988	-0.81
z_9	1+	$C_{50}H_{72}N_{10}O_{15}$	1052.517313	1052.515998	-1.25

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
c_9	1+	$C_{49}H_{74}N_{11}O_{13}S_1$	1056.518279	1056.517189	-1.03
c_{10}	1+	$C_{54}H_{83}N_{12}O_{14}S_1$	1155.586693	1155.585002	-1.46
z_{10}	1+	$C_{55}H_{81}N_{11}O_{16}S_1$	1183.557798	1183.555389	-2.04
z_{10}+H	1+	$C_{55}H_{80}N_{10}O_{17}S_1$	1184.565623	1184.566105	0.41
c_{11}	1+	$C_{58}H_{88}N_{13}O_{17}S_1$	1270.613636	1270.613594	-0.03
MH	1+	$C_{64}H_{99}N_{14}O_{19}S_1$	1399.692915	1399.690568	-1.68
Average error (ppm):					-0.39
Absolute average error (ppm):					0.65
Standard deviation (ppm):					0.62

Table S2. 14 Peak assignment table for the ECD MS/MS spectrum of the peptide $[LGEYGFQN_{Deam}ALIVR+2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_4+H	1+	$C_{23}H_{45}N_6O_5$	485.344595	485.344612	0.04
z_5	1+	$C_{26}H_{49}N_7O_6$	555.373884	555.373760	-0.22
z_5+H	1+	$C_{26}H_{50}N_7O_6$	556.381709	556.381700	-0.02
z_6-C_2O_2H+Deam	1+	$C_{28}H_{53}N_8O_7$	613.403216	613.403163	-0.09
z_6	1+	$C_{30}H_{55}N_9O_8$	669.416811	669.416767	-0.07
z_6+Deam	1+	$C_{30}H_{54}N_8O_9$	670.400828	670.400718	-0.16
y_6+Deam	1+	$C_{30}H_{56}N_9O_9$	686.419552	686.419148	-0.59
MH	2+	$C_{68}H_{107}N_{18}O_{19}$	740.401358	740.401491	0.18
MH+Deam	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893368	740.893400	0.04
z_7-CO_2+Deam	1+	$C_{34}H_{62}N_{10}O_9$	754.469576	754.469832	0.34
z_7+Deam	1+	$C_{35}H_{62}N_{10}O_{11}$	798.459406	798.459501	0.12
z_8+Deam	1+	$C_{44}H_{71}N_{11}O_{12}$	945.527820	945.527380	-0.47
y_8-H+Deam	1+	$C_{44}H_{72}N_{12}O_{12}$	960.538717	960.538822	0.11
y_8+Deam	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546544	961.546216	-0.34
z_9	1+	$C_{46}H_{75}N_{13}O_{12}$	1001.565266	1001.565102	-0.16
z_9+Deam	1+	$C_{46}H_{74}N_{12}O_{13}$	1002.549284	1002.549038	-0.25
y_9+Deam	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.568008	1018.568635	0.62
z_{10}+Deam	1+	$C_{55}H_{83}N_{13}O_{15}$	1165.612613	1165.612861	0.21
y_{10}+Deam	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631337	1181.630914	-0.36
z_{11}+Deam	1+	$C_{60}H_{90}N_{14}O_{18}$	1294.655207	1294.655200	-0.01
c_{12}+Deam	1+	$C_{62}H_{95}N_{14}O_{18}$	1323.694329	1323.694408	0.06

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
MH	1+	$C_{68}H_{106}N_{17}O_{20}$	1480.779500	1480.779500	0.00
Average error (ppm):					-0.05
Absolute average error (ppm):					0.20
Standard deviation (ppm):					0.18

Table S2. 15 Peak assignment table for the synthetic peptide [LV<u>D</u>ELTEFAK+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_{10}H_{21}N_2O_1$	185.164836	185.164836	0.00
z_2	1+	$C_9H_{18}N_2O_3$	202.131194	202.131269	0.37
b_2	1+	$C_{11}H_{21}N_2O_2$	213.159750	213.159728	-0.11
y_2	1+	$C_9H_{20}N_3O_3$	218.149918	218.149983	0.30
DE	1+	$C_9H_{13}N_2O_6$	245.076812	245.076881	0.28
v_3+H	1+	$C_{11}H_{21}N_4O_4$	273.155732	273.155682	-0.18
ELT-H_2O	1+	$C_{15}H_{24}N_3O_5$	326.171047	326.171208	0.49
b_3	1+	$C_{15}H_{26}N_3O_5$	328.186690	328.186781	0.28
z_3	1+	$C_{18}H_{27}N_3O_4$	349.199604	349.199628	0.07
DEL	1+	$C_{15}H_{24}N_3O_7$	358.160876	358.160986	0.31
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218328	365.218387	0.16
TEF	1+	$C_{18}H_{24}N_3O_6$	378.165961	378.166004	0.11
x_3+H	1+	$C_{19}H_{28}N_4O_3$	392.205418	392.205440	0.06
a_4+H	1+	$C_{19}H_{34}N_4O_7$	430.242202	430.242326	0.29
TEFA-H_2O	1+	$C_{21}H_{27}N_4O_6$	431.192511	431.192793	0.65
b_4	1+	$C_{20}H_{33}N_4O_8$	457.229280	457.229418	0.30
y_4-H_2O	1+	$C_{23}H_{34}N_5O_6$	476.250354	476.250375	0.05
z_4	1+	$C_{23}H_{34}N_4O_7$	478.242194	478.242217	0.05
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260918	494.260920	0.00
MH-C_7H_7	2+	$C_{46}H_{80}N_{11}O_{18}$	537.283591	537.283843	0.47
a_5+H	1+	$C_{25}H_{45}N_5O_8$	543.326266	543.326380	0.21
b_5-H_2O	1+	$C_{26}H_{42}N_5O_8$	552.302776	552.303032	0.46
b_5	1+	$C_{26}H_{44}N_5O_9$	570.313340	570.313355	0.03
MH-H_2O	2+	$C_{53}H_{85}N_{11}O_{17}$	573.805697	573.805808	0.19
MH	2+	$C_{53}H_{87}N_{11}O_{18}$	582.810979	582.810745	-0.40
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308598	595.308590	-0.01
b_6-H_2O	1+	$C_{30}H_{49}N_6O_{10}$	653.350456	653.350703	0.38
b_6	1+	$C_{30}H_{51}N_6O_{11}$	671.361020	671.361284	0.39
y_6-H_2O	1+	$C_{33}H_{52}N_7O_9$	690.382094	690.382067	-0.04
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392658	708.392886	0.32

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_7+H	1+	$C_{34}H_{59}N_7O_{13}$	773.416521	773.416704	0.24
b_7-H_2O	1+	$C_{35}H_{56}N_7O_{13}$	782.393046	782.393456	0.52
b_7	1+	$C_{35}H_{58}N_7O_{14}$	800.403610	800.403773	0.20
y_7-H_2O	1+	$C_{38}H_{59}N_8O_{12}$	819.424684	819.425187	0.61
y_7	1+	$C_{38}H_{61}N_8O_{13}$	837.435248	837.435272	0.03
b_8-CO_2	1+	$C_{43}H_{67}N_8O_{13}$	903.482213	903.481520	-0.77
a_8+H	1+	$C_{43}H_{68}N_8O_{14}$	920.484953	920.484490	-0.50
b_8-H_2O	1+	$C_{44}H_{65}N_8O_{14}$	929.461456	929.462287	0.89
b_8	1+	$C_{44}H_{67}N_8O_{15}$	947.472020	947.472612	0.62
y_8	1+	$C_{42}H_{66}N_9O_{16}$	952.462188	952.462164	-0.03
b_9-H_2O	1+	$C_{47}H_{70}N_9O_{15}$	1000.498570	1000.498810	0.24
b_9	1+	$C_{47}H_{72}N_9O_{16}$	1018.509134	1018.509115	-0.02
Average error (ppm):					0.18
Absolute average error (ppm):					0.28
Standard deviation (ppm):					0.23

Table S2. 16 Peak assignment table for the synthetic peptide [LV<u>isoD</u>ELTEFAK+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_{10}H_{21}N_2O_1$	185.164836	185.164837	0.01
z_2	1+	$C_9H_{18}N_2O_3$	202.131194	202.131168	-0.13
b_2	1+	$C_{11}H_{21}N_2O_2$	213.159750	213.159735	-0.07
y_2	1+	$C_9H_{20}N_3O_3$	218.149918	218.149915	-0.01
DE	1+	$C_9H_{13}N_2O_6$	245.076812	245.076782	-0.12
v_3+H	1+	$C_{11}H_{21}N_4O_4$	273.155732	273.155670	-0.23
ELT-H_2O	1+	$C_{15}H_{24}N_3O_5$	326.171047	326.171079	0.10
b_3	1+	$C_{15}H_{26}N_3O_5$	328.186690	328.186690	0.00
z_3	1+	$C_{18}H_{27}N_3O_4$	349.199604	349.199587	-0.05
DEL	1+	$C_{15}H_{24}N_3O_7$	358.160876	358.160929	0.15
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218328	365.218373	0.12
TEF	1+	$C_{18}H_{24}N_3O_6$	378.165961	378.165956	-0.01
x_3+H	1+	$C_{19}H_{28}N_4O_3$	392.205418	392.205377	-0.10
a_4+H	1+	$C_{19}H_{34}N_4O_7$	430.242202	430.242322	0.28
b_4	1+	$C_{20}H_{33}N_4O_8$	457.229280	457.229375	0.21
z_4	1+	$C_{23}H_{34}N_4O_7$	478.242194	478.242347	0.32
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260918	494.260977	0.12
MH-C_7H_7	2+	$C_{46}H_{80}N_{11}O_{18}$	537.283591	537.283783	0.36
a_5+H	1+	$C_{25}H_{45}N_5O_8$	543.326266	543.326625	0.66
b_5	1+	$C_{26}H_{44}N_5O_9$	570.313340	570.313364	0.04

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
MH-H_2O	1+	$C_{53}H_{85}N_{11}O_{17}$	573.805697	573.805865	0.29
MH	2+	$C_{53}H_{87}N_{11}O_{18}$	582.810979	582.810308	-1.15
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308598	595.308363	-0.40
b_6-H_2O	1+	$C_{30}H_{49}N_6O_{10}$	653.350456	653.351008	0.85
b_6	1+	$C_{30}H_{51}N_6O_{11}$	671.361020	671.361139	0.18
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392658	708.392974	0.45
a_7+H	1+	$C_{34}H_{59}N_7O_{13}$	773.416521	773.416113	-0.53
b_7	1+	$C_{35}H_{58}N_7O_{14}$	800.403610	800.404230	0.77
y_7-H_2O	1+	$C_{38}H_{59}N_8O_{12}$	819.424684	819.425182	0.61
y_7	1+	$C_{38}H_{61}N_8O_{13}$	837.435248	837.435513	0.32
a_8+H	1+	$C_{43}H_{68}N_8O_{14}$	920.484953	920.485570	0.67
y_8	1+	$C_{42}H_{66}N_9O_{16}$	952.462188	952.462906	0.75
b_9	1+	$C_{47}H_{72}N_9O_{16}$	1018.509134	1018.509086	-0.05
Average error (ppm):					0.13
Absolute average error (ppm):					0.31
Standard deviation (ppm):					0.29

Table S2. 17 Peak assignment table for the synthetic peptide [TVMEDFVAFVDK+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_8H_{17}N_2O_2$	173.128456	173.128495	0.23
b_2	1+	$C_9H_{17}N_2O_3$	201.123370	201.123371	0.00
y_2-H_2O	1+	$C_{10}H_{18}N_3O_4$	244.129180	244.129186	0.03
ME	1+	$C_{10}H_{16}N_2O_4S_1$	261.090354	261.090348	-0.02
y_2	1+	$C_{10}H_{20}N_3O_5$	262.139744	262.139735	-0.04
VAF/VFA/AFV	1+	$C_{17}H_{23}N_3O_3$	318.181218	318.181200	-0.06
b_3	1+	$C_{14}H_{26}N_3O_4S_1$	332.163850	332.163866	0.05
z_3	1+	$C_{15}H_{27}N_3O_6$	345.189434	345.189478	0.13
y_3	1+	$C_{15}H_{29}N_4O_6$	361.208158	361.208155	-0.01
DFV	1+	$C_{18}H_{23}N_3O_5$	362.171047	362.171003	-0.12
MED	1+	$C_{14}H_{21}N_3O_7S_1$	376.117297	376.117276	-0.06
DFVA	1+	$C_{21}H_{28}N_4O_6$	433.208161	433.208130	-0.07
b_4	1+	$C_{19}H_{33}N_4O_7S_1$	461.206440	461.206400	-0.09
EDFV-H_2O	1+	$C_{23}H_{28}N_4O_7$	473.203075	473.203058	-0.04
y_4	1+	$C_{24}H_{38}N_5O_7$	508.276568	508.276576	0.02
MEDF	1+	$C_{23}H_{30}N_4O_8S_1$	523.185711	523.185685	-0.05
VAFVD	1+	$C_{26}H_{37}N_5O_7$	532.276575	532.276404	-0.32
b_5	1+	$C_{23}H_{38}N_5O_{10}S_1$	576.233380	576.233429	0.08
y_5	1+	$C_{27}H_{43}N_6O_8$	579.313682	579.313710	0.05

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_{10}	2+	$C_{55}H_{83}N_{11}O_{17}S_1$	600.783880	600.784375	0.82
MH-C_7H_7	1+	$C_{57}H_{92}N_{13}O_{20}S_1$	655.314550	655.314291	-0.40
y_6	1+	$C_{32}H_{52}N_7O_9$	678.382096	678.382145	0.07
MH - CH_3	1+	$C_{63}H_{96}N_{13}O_{20}S_1$	693.330200	693.330152	-0.07
MH	2+	$C_{64}H_{99}N_{13}O_{20}S_1$	700.841938	700.841377	-0.80
b_6	1+	$C_{32}H_{47}N_6O_{11}S_1$	723.30179	723.301802	0.02
b_7	1+	$C_{37}H_{56}N_7O_{12}S_1$	822.370204	822.370395	0.23
y_7	1+	$C_{41}H_{61}N_8O_{10}$	825.450506	825.450192	-0.38
b_8	1+	$C_{40}H_{61}N_8O_{13}S_1$	893.407318	893.407411	0.10
y_8	1+	$C_{45}H_{66}N_9O_{13}$	940.477446	940.477064	-0.41
b_9	1+	$C_{49}H_{70}N_9O_{14}S_1$	1040.475748	1040.475455	-0.28
y_9-H_2O	1+	$C_{50}H_{71}N_{10}O_{15}$	1051.509472	1051.508506	-0.92
b_{10}	1+	$C_{54}H_{79}N_{10}O_{15}S_1$	1139.544142	1139.544038	-0.09
y_{10}	1+	$C_{55}H_{82}N_{11}O_{17}S_1$	1200.560516	1200.560250	-0.22
b_{11}	1+	$C_{58}H_{84}N_{11}O_{18}S_1$	1254.571082	1254.571067	-0.01
Average error (ppm):					-0.08
Absolute average error (ppm):					0.18
Standard deviation (ppm):					0.24

Table S2. 18 Peak assignment table for synthetic peptide [TVME<u>isoD</u>FVAFVDK+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_8H_{17}N_2O_2$	173.128456	173.128538	0.47
b_2	1+	$C_9H_{17}N_2O_3$	201.123370	201.123370	0.00
y_2-H_2O	1+	$C_{10}H_{18}N_3O_4$	244.129180	244.129263	0.34
ME	1+	$C_{10}H_{16}N_2O_4S_1$	261.090354	261.090333	-0.08
y_2	1+	$C_{10}H_{20}N_3O_5$	262.139744	262.139752	0.03
VAF/VFA/AFV	1+	$C_{17}H_{23}N_3O_3$	318.181218	318.181142	-0.24
b_3	1+	$C_{14}H_{26}N_3O_4S_1$	332.163850	332.163871	0.06
z_3	1+	$C_{15}H_{27}N_3O_6$	345.189434	345.189541	0.31
y_3	1+	$C_{15}H_{29}N_4O_6$	361.208158	361.208160	0.01
DFV	1+	$C_{18}H_{23}N_3O_5$	362.171047	362.170920	-0.35
MED	1+	$C_{14}H_{21}N_3O_7S_1$	376.117297	376.117342	0.12
DFVA	1+	$C_{21}H_{28}N_4O_6$	433.208161	433.207998	-0.38
b_4	1+	$C_{19}H_{33}N_4O_7S_1$	461.206440	461.206348	-0.20
EDFV-H_2O	1+	$C_{23}H_{28}N_4O_7$	473.203075	473.202897	-0.38
y_4	1+	$C_{24}H_{38}N_5O_7$	508.276568	508.276411	-0.31

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
MEDF	1+	$C_{23}H_{30}N_4O_8S_1$	523.185711	523.185446	-0.51
VAFVD	1+	$C_{26}H_{37}N_5O_7$	532.276575	532.276142	-0.81
b_5	1+	$C_{23}H_{38}N_5O_{10}S_1$	576.233380	576.233295	-0.15
y_5	1+	$C_{27}H_{43}N_6O_8$	579.313682	579.313677	-0.01
y_{10}	2+	$C_{55}H_{83}N_{11}O_{17}S_1$	600.783880	600.783860	-0.03
MH-C_7H_7	2+	$C_{57}H_{92}N_{13}O_{20}S_1$	655.314550	655.314410	-0.21
y_6	1+	$C_{32}H_{52}N_7O_9$	678.382096	678.382027	-0.10
MH-CH_3	2+	$C_{63}H_{96}N_{13}O_{20}S_1$	693.330200	693.330118	-0.12
a_6+H	1+	$C_{31}H_{48}N_6O_{10}S_1$	696.314716	696.315299	0.84
MH	2+	$C_{64}H_{99}N_{13}O_{20}S_1$	700.841938	700.841303	-0.91
b_6	1+	$C_{32}H_{47}N_6O_{11}S_1$	723.301790	723.301020	-1.07
b_7	1+	$C_{37}H_{56}N_7O_{12}S_1$	822.370204	822.369907	-0.36
y_7	1+	$C_{41}H_{61}N_8O_{10}$	825.450506	825.450584	0.09
b_8	1+	$C_{40}H_{61}N_8O_{13}S_1$	893.407318	893.407774	0.51
y_8	1+	$C_{45}H_{66}N_9O_{13}$	940.477446	940.477256	-0.20
b_9	1+	$C_{49}H_{70}N_9O_{14}S_1$	1040.475748	1040.475783	0.03
y_9-H_2O	1+	$C_{50}H_{71}N_{10}O_{15}$	1051.509472	1051.509541	0.07
b_{10}	1+	$C_{54}H_{79}N_{10}O_{15}S_1$	1139.544142	1139.543191	-0.83
y_{10}	1+	$C_{55}H_{82}N_{11}O_{17}S_1$	1200.560516	1200.560773	0.21
b_{11}	1+	$C_{58}H_{84}N_{11}O_{18}S_1$	1254.571082	1254.571092	0.01
Average error (ppm):					-0.12
Absolute average error (ppm):					0.31
Standard deviation (ppm):					0.30

Table S2. 19 Peak assignment table for synthetic peptide [LGEYGFQDALIVR+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_2	1+	$C_8H_{15}N_2O_2$	171.112800	171.112801	0.00
y_1	1+	$C_6H_{15}N_4O_2$	175.118952	175.118944	-0.05
z_2	1+	$C_{11}H_{22}N_4O_3$	258.168642	258.168595	-0.18
a_3	1+	$C_{12}H_{22}N_3O_4$	272.160483	272.160347	-0.50
y_2	1+	$C_{11}H_{24}N_5O_3$	274.187366	274.187274	-0.34
EY	1+	$C_{14}H_{17}N_2O_5$	293.113198	293.113281	0.28
b_3	1+	$C_{13}H_{22}N_3O_5$	300.155390	300.155429	0.13
c_3	1+	$C_{13}H_{25}N_4O_5$	317.181940	317.181876	-0.20
EYG/EGY	1+	$C_{16}H_{19}N_3O_6$	350.134661	350.134565	-0.28
y_3	1+	$C_{17}H_{35}N_6O_4$	387.271426	387.271389	-0.10
a_4	1+	$C_{21}H_{31}N_4O_6$	435.223806	435.223692	-0.26
b_4-H_2O	1+	$C_{22}H_{29}N_4O_6$	445.208156	445.208011	-0.33

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_4	1+	$C_{22}H_{31}N_4O_7$	463.218720	463.218451	-0.58
EYGF-H_2O	1+	$C_{25}H_{26}N_4O_6$	479.192511	479.192065	-0.93
y_4	1+	$C_{23}H_{46}N_7O_5$	500.355486	500.355487	0.00
z_5	1+	$C_{26}H_{49}N_7O_6$	555.373876	555.373742	-0.24
y_5	1+	$C_{26}H_{51}N_8O_6$	571.392600	571.392596	-0.01
a_6	1+	$C_{32}H_{43}N_6O_8$	639.313680	639.313256	-0.66
b_6-H_2O	1+	$C_{33}H_{41}N_6O_8$	649.298030	649.298178	0.23
b_6	1+	$C_{33}H_{43}N_6O_9$	667.308594	667.309070	0.71
y_6	1+	$C_{30}H_{56}N_9O_9$	686.419540	686.419283	-0.37
MH- C_7H_7O	2+	$C_{61}H_{100}N_{17}O_{19}$	687.368543	687.368865	0.47
MH- C_3H_7	2+	$C_{65}H_{100}N_{17}O_{20}$	712.358175	712.357643	-0.75
MH-H_2O	2+	$C_{68}H_{105}N_{17}O_{19}$	731.888084	731.887637	-0.61
MH	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893388	740.892573	-1.10
b_7	1+	$C_{38}H_{51}N_8O_{11}$	795.367174	795.366234	-1.18
y_7-H_2O	1+	$C_{35}H_{62}N_{11}O_{10}$	796.467565	796.467192	-0.47
y_7-NH_3	1+	$C_{35}H_{61}N_{10}O_{11}$	797.451571	797.451399	-0.22
y_7	1+	$C_{35}H_{64}N_{11}O_{11}$	814.478120	814.477471	-0.80
b_8-H_2O	1+	$C_{42}H_{54}N_9O_{13}$	892.383550	892.382686	-0.97
b_8	1+	$C_{42}H_{56}N_9O_{14}$	910.394114	910.394374	0.29
y_8	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546530	961.546557	0.03
b_9-H_2O	1+	$C_{45}H_{59}N_{10}O_{14}$	963.420664	963.420052	-0.64
y_9	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.567994	1018.568487	0.48
b_{10}	1+	$C_{51}H_{72}N_{11}O_{16}$	1094.515305	1094.515067	-0.22
b_{10}-H_2O	1+	$C_{51}H_{70}N_{11}O_{15}$	1076.504724	1076.505189	0.43
y_{10}	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631324	1181.631295	-0.02
b_{11}-H_2O	1+	$C_{57}H_{81}N_{12}O_{16}$	1189.588784	1189.587602	-0.99
x_{12}+H	1+	$C_{63}H_{94}N_{16}O_{20}$	1394.682481	1394.680937	-1.11
Average error (ppm):					-0.28
Absolute average error (ppm):					0.44
Standard deviation (ppm):					0.34

Table S2. 20 Peak assignment table for synthetic peptide [LGEYGFQ<u>isoD</u>ALIVR+2H]$^{2+}$ UVPD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_2	1+	$C_8H_{15}N_2O_2$	171.112800	171.112801	0.00
y_1	1+	$C_6H_{15}N_4O_2$	175.118952	175.118933	-0.11
z_2	1+	$C_{11}H_{22}N_4O_3$	258.168642	258.168486	-0.60
a_3	1+	$C_{12}H_{22}N_3O_4$	272.160483	272.160253	-0.85
y_2	1+	$C_{11}H_{24}N_5O_3$	274.187366	274.187187	-0.65

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
EY	1+	$C_{14}H_{17}N_2O_5$	293.113198	293.113216	0.06
b_3	1+	$C_{13}H_{22}N_3O_5$	300.155390	300.155390	0.00
c_3	1+	$C_{13}H_{25}N_4O_5$	317.181940	317.181774	-0.52
EYG/EGY	1+	$C_{16}H_{19}N_3O_6$	350.134661	350.134607	-0.16
y_3	1+	$C_{17}H_{35}N_6O_4$	387.271426	387.271319	-0.28
a_4	1+	$C_{21}H_{31}N_4O_6$	435.223806	435.223976	0.39
b_4-H_2O	1+	$C_{22}H_{29}N_4O_6$	445.208156	445.207667	-1.10
b_4	1+	$C_{22}H_{31}N_4O_7$	463.218720	463.218383	-0.73
EYGF-H_2O	1+	$C_{25}H_{26}N_4O_6$	479.192511	479.192084	-0.89
y_4	1+	$C_{23}H_{46}N_7O_5$	500.355486	500.355489	0.01
z_5	1+	$C_{26}H_{49}N_7O_6$	555.373876	555.373274	-1.08
y_5	1+	$C_{26}H_{51}N_8O_6$	571.392600	571.392430	-0.30
a_6	1+	$C_{32}H_{43}N_6O_8$	639.313680	639.313067	-0.96
b_6-H_2O	1+	$C_{33}H_{41}N_6O_8$	649.298030	649.297552	-0.74
b_6	1+	$C_{33}H_{43}N_6O_9$	667.308594	667.308404	-0.29
y_6-H_2O	1+	$C_{30}H_{54}N_9O_8$	668.408976	668.408266	-1.06
y_6-NH_3	1+	$C_{30}H_{53}N_8O_9$	669.392991	669.392415	-0.86
y_6	1+	$C_{30}H_{56}N_9O_9$	686.419540	686.419621	0.12
MH-C_7H_7O	2+	$C_{61}H_{100}N_{17}O_{19}$	687.368543	687.368757	0.31
MH - C_7H_7	2+	$C_{61}H_{100}N_{17}O_{20}$	695.365978	695.366128	0.22
MH- C_3H_7	2+	$C_{65}H_{100}N_{17}O_{20}$	712.358175	712.358296	0.17
MH-H_2O	2+	$C_{68}H_{105}N_{17}O_{19}$	731.888084	731.888037	-0.06
y_{13}-NH_3	2+	$C_{68}H_{104}N_{16}O_{20}$	732.380093	732.379634	-0.63
MH	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893388	740.892587	-1.08
b_7-H_2O	1+	$C_{38}H_{49}N_8O_{10}$	777.356618	777.356054	-0.73
b_7	1+	$C_{38}H_{51}N_8O_{11}$	795.367174	795.366381	-1.00
y_7-H_2O	1+	$C_{35}H_{62}N_{11}O_{10}$	796.467565	796.466796	-0.97
y_7-NH_3	1+	$C_{35}H_{61}N_{10}O_{11}$	797.451571	797.451383	-0.24
y_7	1+	$C_{35}H_{64}N_{11}O_{11}$	814.478120	814.477494	-0.77
b_8-H_2O	1+	$C_{42}H_{54}N_9O_{13}$	892.383550	892.382597	-1.07
b_8	1+	$C_{42}H_{56}N_9O_{14}$	910.394114	910.393433	-0.75
y_8	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546530	961.545697	-0.87
b_9-H_2O	1+	$C_{45}H_{59}N_{10}O_{14}$	963.420664	963.419968	-0.72
y_9	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.567994	1018.567840	-0.15
b_{10}-H_2O	1+	$C_{51}H_{70}N_{11}O_{15}$	1076.504724	1076.504450	-0.25
y_{10}	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631324	1181.631318	-0.01
b_{11}-H_2O	1+	$C_{57}H_{81}N_{12}O_{16}$	1189.588784	1189.589741	0.80
x_{12}+H	1+	$C_{63}H_{94}N_{16}O_{20}$	1394.682481	1394.680873	-1.15
Average error (ppm):					-0.45
Absolute average error (ppm):					0.55
Standard deviation (ppm):					0.38

Table S2. 21 Peak assignment table for the UVPD MS/MS spectrum of the peptide $[LVN_{Deam}ELTEFAK+2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_{10}H_{21}N_2O_1$	185.1648358	185.164836	0
b_2	1+	$C_{11}H_{21}N_2O_2$	213.1597505	213.159706	-0.21
y_2	1+	$C_9H_{20}N_3O_3$	218.1499182	218.15002	0.47
NE	1+	$C_9\ H_{14}\ N_3\ O_5$	244.092797	244.092972	0.72
b_3	1+	$C_{15}H_{27}N_4O_4$	327.202682	327.20268	-0.01
NEL	1+	$C_{15}H_{25}N_4O_6$	357.176861	357.176935	0.21
y_3	1+	$C_{18}H_{29}N_4O_4$	365.218332	365.218272	-0.16
x_3+H	1+	$C_{19}H_{28}N_4O_3$	392.2054178	392.205047	-0.95
b_4-NH_3	1+	$C_{20}H_{31}N_4O_7$	439.2187158	439.218363	-0.8
b_4	1+	$C_{20}H_{34}N_5O_7$	456.245276	456.245455	0.39
y_4	1+	$C_{23}H_{36}N_5O_7$	494.260926	494.260846	-0.16
MH-C_7H_7	2+	$C_{46}H_{80}N_{11}O_{18}$	537.2835914	537.283709	0.22
LTEFA-H_2O	1+	$C_{27}H_{38}N_5O_7$	544.276575	544.276481	-0.17
b_5-H_2O	1+	$C_{26}H_{42}N_5O_8$	552.3027758	552.303011	0.43
b_5	1+	$C_{26}H_{45}N_6O_8$	569.32934	569.329353	0.02
MH-H_2O	2+	$C_{53}H_{84}N_{12}O_{16}$	573.8056967	573.80615	0.79
MH	2+	$C_{53}H_{86}N_{12}O_{17}$	582.318971	582.318544	-0.73
MH+Deam	2+	$C_{53}H_{87}N_{11}O_{18}$	582.810979	582.811547	0.97
y_5	1+	$C_{27}H_{43}N_6O_9$	595.308605	595.308613	0.01
b_6-NH_3	1+	$C_{30}H_{49}N_6O_{10}$	653.35047	653.350785	0.48
b_6	1+	$C_{30}H_{52}N_7O_{10}$	670.377019	670.37724	0.33
y_6	1+	$C_{33}H_{54}N_7O_{10}$	708.392669	708.392969	0.42
a_7+H	1+	$C_{34}H_{60}N_8O_{12}$	772.432523	772.432802	0.36
b_7-NH_3	1+	$C_{35}H_{56}N_7O_{13}$	782.393064	782.393423	0.46
b_7	1+	$C_{35}H_{59}N_8O_{13}$	799.419613	799.419816	0.25
y_7-H_2O	1+	$C_{38}H_{59}N_8O_{12}$	819.424698	819.425386	0.84
y_7	1+	$C_{38}H_{61}N_8O_{13}$	837.435263	837.43561	0.41
y_8-NH_3	1+	$C_{42}H_{64}N_9O_{15}$	934.451639	934.452691	1.13
y_8	1+	$C_{42}H_{67}N_{10}O_{15}$	951.478736	951.478966	0.24
y_8+Deam	1+	$C_{42}H_{66}N_9O_{16}$	952.4621882	952.463129	0.99
b_9-NH_3	1+	$C_{47}H_{70}N_9O_{15}$	1000.498592	1000.499375	0.78
b_9	1+	$C_{47}H_{73}N_{10}O_{15}$	1017.525141	1017.525662	0.51
Average error (ppm):					0.26
Absolute average error (ppm):					0.46
Standard deviation (ppm):					0.32

Table S2. 22 Peak assignment table for the UVPD MS/MS spectrum of the peptide $[TVMEN_{Deam}FVAFVDK +2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
a_2	1+	$C_8H_{17}N_2O_2$	173.1284558	173.128454	-0.01
b_2	1+	$C_9H_{17}N_2O_3$	201.1233705	201.123355	-0.08
N(Deam)F-CO	1+	$C_{12}H_{15}N_2O_3$	235.107719	235.107699	-0.09
y_2-H_2O	1+	$C_{10}H_{18}N_3O_4$	244.1291795	244.129164	-0.06
ME	1+	$C_{10}H_{16}N_2O_4S_1$	261.090354	261.09034	-0.05
y_2	1+	$C_{10}H_{20}N_3O_5$	262.1397442	262.139731	-0.05
VAF/VFA/AFV	1+	$C_{17}H_{23}N_3O_3$	318.1812176	318.181199	-0.06
b_3	1+	$C_{14}H_{26}N_3O_4S_1$	332.1638505	332.163829	-0.06
y_3-H_2O	1+	$C_{15}H_{27}N_4O_5$	343.197596	343.19757	-0.08
z_3	1+	$C_{15}H_{27}N_3O_6$	345.1894335	345.189387	-0.13
y_3	1+	$C_{15}H_{29}N_4O_6$	361.2081582	361.208148	-0.03
MEN(Deam)	1+	$C_{14}H_{22}N_3O_7S_1$	376.117298	376.11729	-0.02
AFVD-CO-H_2O	1+	$C_{20}H_{27}N_4O_4$	387.202682	387.202663	-0.05
VAFV-CO	1+	$C_{21}H_{33}N_4O_3$	389.254717	389.254676	-0.11
AFVD-H_2O	1+	$C_{21}H_{27}N_4O_5$	415.197596	415.19758	-0.04
DFVA	1+	$C_{21}H_{28}N_4O_6$	433.2081606	433.208149	-0.03
b_4-H_2O	1+	$C_{19}H_{31}N_4O_6S_1$	443.195882	443.195827	-0.12
b_4	1+	$C_{19}H_{33}N_4O_7S_1$	461.2064405	461.206445	0.01
y_4-H_2O	1+	$C_{24}H_{36}N_5O_6$	490.26601	490.265912	-0.20
y_4	1+	$C_{24}H_{38}N_5O_7$	508.2765682	508.276593	0.05
b_5	1+	$C_{23}H_{39}N_6O_9S_1$	575.249374	575.249321	-0.09
b_5+Deam	1+	$C_{23}H_{38}N_5O_{10}S_1$	576.23339	576.233325	-0.11
y_5	1+	$C_{27}H_{43}N_6O_8$	579.3136822	579.313718	0.06
FVAFVD-H_2O	1+	$C_{35}H_{45}N_6O_7$	661.334424	661.334292	-0.20
y_6	1+	$C_{32}H_{52}N_7O_9$	678.3820962	678.382111	0.02
MH	2+	$C_{64}H_{99}N_{14}O_{19}S_1$	700.349946	700.349101	-1.21
MH+Deam	2+	$C_{64}H_{98}N_{13}O_{20}S_1$	700.841953	700.841987	0.05
b_6-H_2O	1+	$C_{32}H_{45}N_6O_{10}S_1$	705.291239	705.291173	-0.09
b_6	1+	$C_{32}H_{48}N_7O_{10}S_1$	722.317788	722.317627	-0.22
N(Deam)FVAFVD-H_2O	1+	$C_{39}H_{50}N_7O_{10}$	776.361367	776.361015	-0.45
y_7	1+	$C_{41}H_{61}N_8O_{10}$	825.4505062	825.450366	-0.17
b_8-H_2O	1+	$C_{40}H_{60}N_9O_{11}S_1$	874.412751	874.412413	-0.39

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_8-H_2O+Deam	1+	$C_{40}H_{59}N_8O_{12}S_1$	875.396767	875.396533	-0.27
b_8	1+	$C_{40}H_{62}N_9O_{12}S_1$	892.423316	892.42347	0.17
b_8+Deam	1+	$C_{40}H_{61}N_8O_{13}S_1$	893.4073185	893.407304	-0.02
y_8	1+	$C_{45}H_{67}N_{10}O_{12}$	939.493444	939.493255	-0.20
y_8+Deam	1+	$C_{45}H_{66}N_9O_{13}$	940.4774462	940.477124	-0.34
b_9	1+	$C_{49}H_{71}N_{10}O_{13}S_1$	1039.49173	1039.491679	-0.05
b_9+Deam	1+	$C_{49}H_{70}N_9O_{14}S_1$	1040.475746	1040.475684	-0.06
y_9-H_2O	1+	$C_{50}H_{72}N_{11}O_{14}$	1050.525472	1050.525073	-0.38
y_9-H_2O+Deam	1+	$C_{50}H_{71}N_{10}O_{15}$	1051.509488	1051.508793	-0.66
b_{10}	1+	$C_{54}H_{80}N_{11}O_{14}S_1$	1138.560144	1138.560282	0.12
b_{10}+Deam	1+	$C_{54}H_{79}N_{10}O_{15}S_1$	1139.544159	1139.543317	-0.74
y_{10}	1+	$C_{55}H_{83}N_{12}O_{16}S_1$	1199.576522	1199.576511	-0.01
y_{10}+Deam	1+	$C_{55}H_{82}N_{11}O_{17}S_1$	1200.560538	1200.55935	-0.99
b_{11}	1+	$C_{58}H_{85}N_{12}O_{17}S_1$	1253.587087	1253.58691	-0.14
Average error (ppm):					-0.16
Absolute average error (ppm):					0.18
Standard deviation (ppm):					0.25

Table S2. 23 Peak assignment table for the UVPD MS/MS spectrum of the peptide $[LGEYGFQN_{Deam}ALIVR+2H]^{2+}$ from the BSA digest sample (incubation day 5).

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_2	1+	$C_8H_{15}N_2O_2$	171.112800	171.112801	0.00
y_1	1+	$C_6H_{15}N_4O_2$	175.118952	175.118922	-0.17
z_2	1+	$C_{11}H_{22}N_4O_3$	258.168642	258.168656	0.06
a_3	1+	$C_{12}H_{22}N_3O_4$	272.160483	272.160547	0.24
y_2	1+	$C_{11}H_{24}N_5O_3$	274.187366	274.187213	-0.56
EY	1+	$C_{14}H_{17}N_2O_5$	293.113198	293.113278	0.27
b_3	1+	$C_{13}H_{22}N_3O_5$	300.155390	300.155425	0.12
c_3	1+	$C_{13}H_{25}N_4O_5$	317.181940	317.181679	-0.82
EYG/EGY	1+	$C_{16}H_{19}N_3O_6$	350.134661	350.13455	-0.32
y_3	1+	$C_{17}H_{35}N_6O_4$	387.271426	387.271358	-0.18
a_4	1+	$C_{21}H_{31}N_4O_6$	435.223806	435.2239	0.22
b_4-H_2O	1+	$C_{22}H_{29}N_4O_6$	445.208156	445.20812	-0.08

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_4	1+	$C_{22}H_{31}N_4O_7$	463.218720	463.21875	0.06
EYGF-H_2O	1+	$C_{25}H_{26}N_4O_6$	479.192511	479.192326	-0.39
y_4	1+	$C_{23}H_{46}N_7O_5$	500.355486	500.355485	0.00
y_5	1+	$C_{26}H_{51}N_8O_6$	571.392600	571.392601	0.00
b_6-H_2O	1+	$C_{33}H_{41}N_6O_8$	649.298030	649.298198	0.26
y_6+Deam	1+	$C_{30}H_{56}N_9O_9$	686.419551	686.419726	0.25
MH-C_7H_7O	2+	$C_{61}H_{100}N_{17}O_{19}$	687.368543	687.368977	0.63
MH-C_7H_7	2+	$C_{61}H_{100}N_{17}O_{20}$	695.365978	695.366215	0.34
MH-H_2O	2+	$C_{68}H_{105}N_{17}O_{19}$	731.888084	731.88832	0.32
MH	2+	$C_{68}H_{108}N_{18}O_{19}$	740.401358	740.400332	-1.39
MH+Deam	2+	$C_{68}H_{107}N_{17}O_{20}$	740.893388	740.893737	0.47
b_7-H_2O	1+	$C_{38}H_{49}N_8O_{10}$	777.356618	777.356445	-0.22
b_7	1+	$C_{38}H_{51}N_8O_{11}$	795.367181	795.367157	-0.03
y_7-NH_3+Deam	1+	$C_{35}H_{61}N_{10}O_{11}$	797.451579	797.451684	0.13
y_7+Deam	1+	$C_{35}H_{64}N_{11}O_{11}$	814.478120	814.478469	0.43
b_8-H_2O+Deam	1+	$C_{42}H_{54}N_9O_{13}$	892.383550	892.384122	0.64
b_8+Deam	1+	$C_{42}H_{56}N_9O_{14}$	910.394114	910.394674	0.61
y_8+Deam	1+	$C_{44}H_{73}N_{12}O_{12}$	961.546530	961.546959	0.45
y_9	1+	$C_{46}H_{77}N_{14}O_{12}$	1017.583990	1017.583525	-0.46
y_9+Deam	1+	$C_{46}H_{76}N_{13}O_{13}$	1018.567994	1018.56909	1.08
y_{10}+Deam	1+	$C_{55}H_{85}N_{14}O_{15}$	1181.631324	1181.631323	0.00
Average error (ppm):					0.06
Absolute average error (ppm):					0.34
Standard deviation (ppm):					0.32

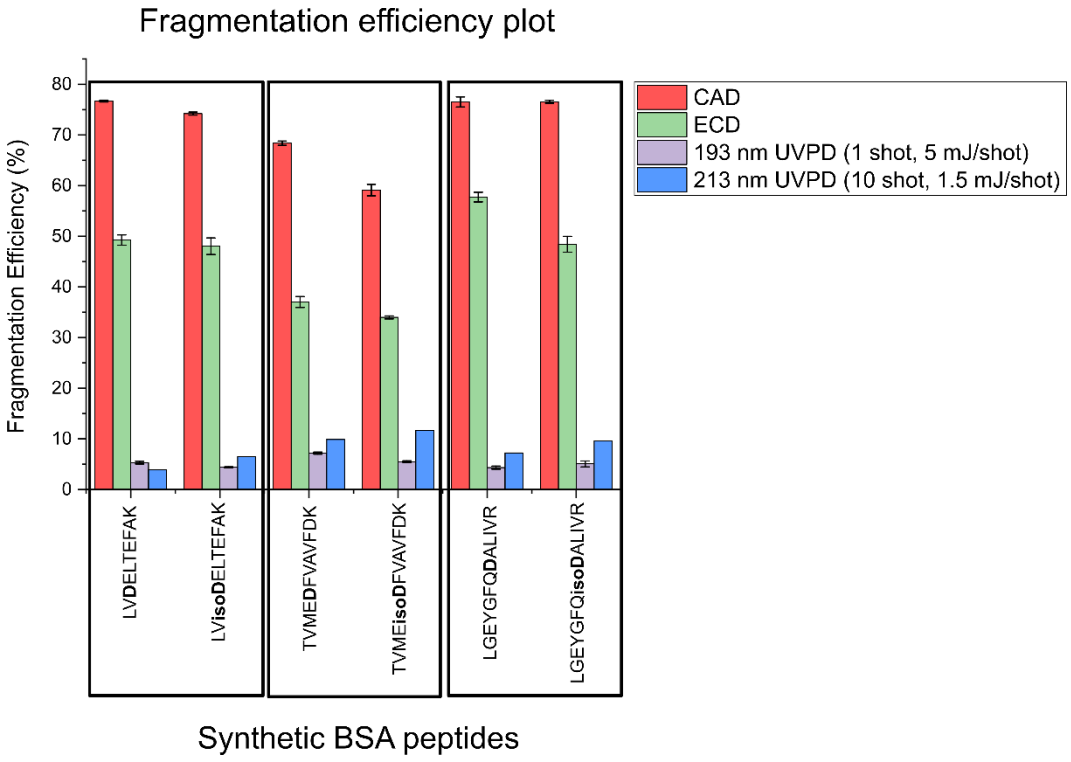


Figure S2. 7 Fragmentation efficiency plot using different MS/MS methods on synthetic isoD and D BSA peptides.

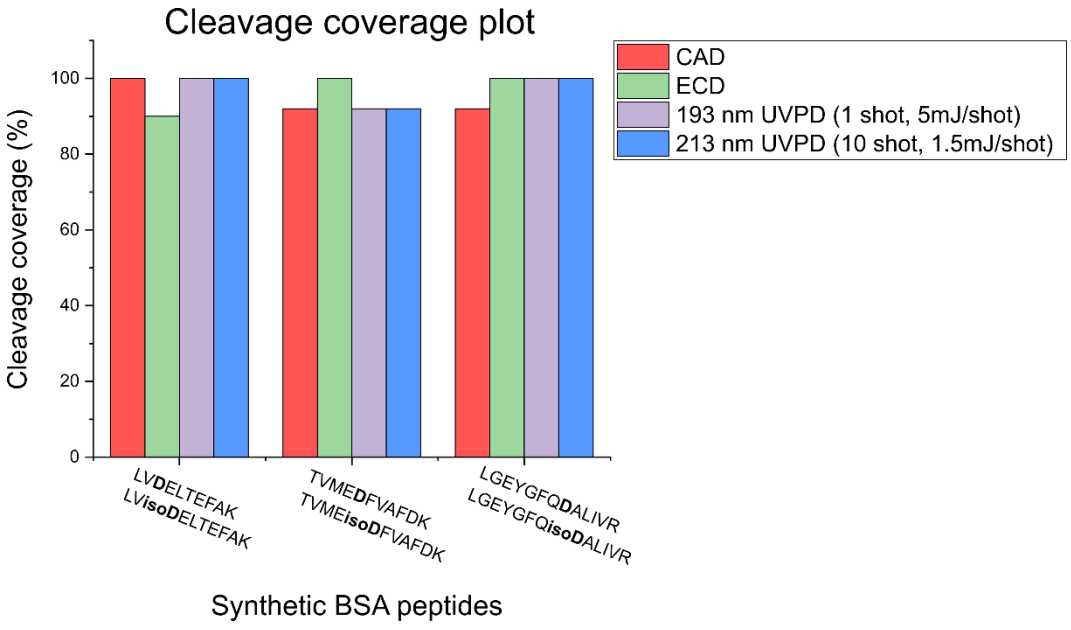


Figure S2. 8 Sequence coverage cleavage plot using different MS/MS methods on synthetic isoD and D BSA peptides.

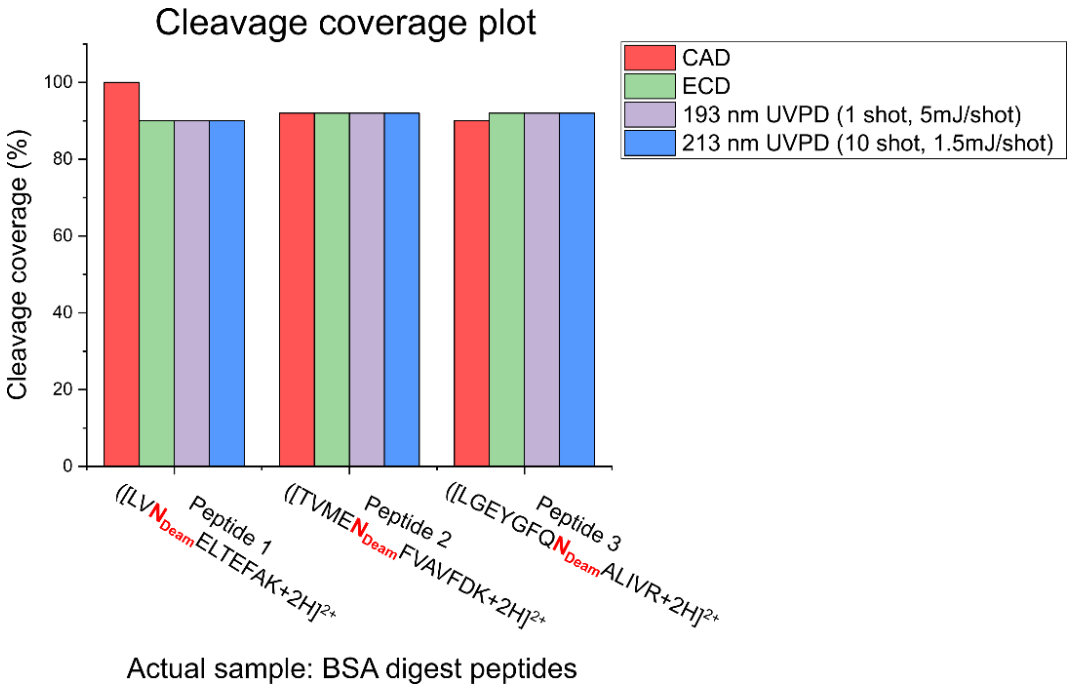


Figure S2. 9 Sequence cleavage coverage plot using different MS/MS methods on deamidated BSA digest peptides (incubation day 5).

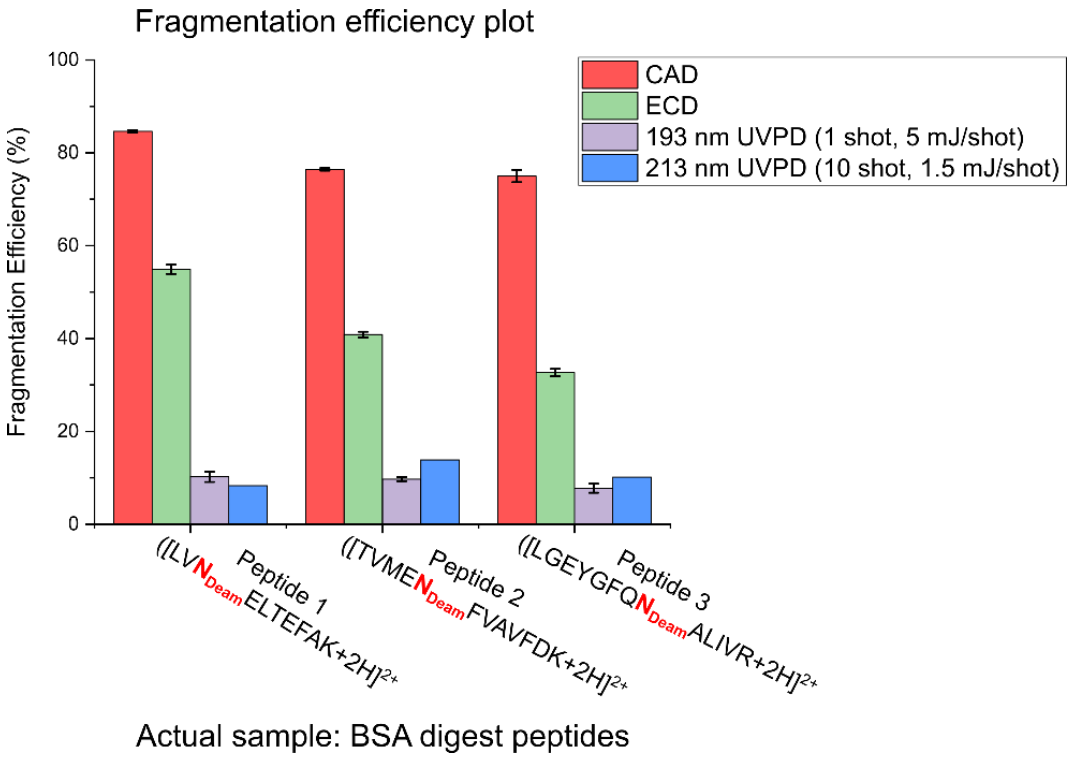


Figure S2. 10 Fragmentation efficiency plot using different MS/MS methods on deamidated BSA digest peptides (incubation day 5).

Table S2. 24 Comparison of MS/MS methods on y fragment intensity fold change between isoD and D synthetic BSA peptides.

Peptide pair		193 nm UVPD	213 nm UVPD	CAD	IRMPD
		y fragment at modification site peak intensity change			
1	D	1	1	1	1
	isoD	7.4	3.2	1.79	50
2	D	1	1	1	1.8
	isoD	6.1	2.5	2.51	1
3	D	1	1	1	1
	isoD	26.5	4.8	1.94	5.4

3. Distinguishing between methylated histidine isomers generated as a post-translational modification of actin

This chapter focuses on the application of fragmentation methods available on the 12 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) for the differentiation and relative quantification of isomeric N-methylated histidine peptides from the cytoskeletal protein, actin.

This was a collaborative project, where sample preparation, data acquisition and analysis results presented in this chapter were carried out by the thesis author and Dr. Yuko P. Y. Lam. Samples were provided by Dr. Hamdi Hussain in Professor Mohan Balasubramanian's research group at the Warwick Medical School - Biomedical Sciences department, University of Warwick.

One manuscript entitled "Distinguishing between methylated histidine isomers generated as a post-translational modification of actin" by Anisha Haris, Yuko P. Y. Lam, Christopher A. Wootton, Hamdi Hussain, Mohan Balasubramanian, and Peter B. O'Connor, is being prepared for submission to the journal, Analytical Chemistry based on the results presented in this chapter.

3.1. Abstract

Methylation is a post-translational modification (PTM) involved in key cellular processes such as the regulation of gene transcription and expression. Actin methylation, specifically N-methylation at histidine-73 residue, has been identified as a regulatory mechanism, which contributes to the function of the cytoskeletal protein. Histidine N-methylation can result in the formation of tele- or pros-methylhistidine, which is challenging to differentiate using various analytical techniques due to subtle changes in their structures and the zero-mass difference between the residues. Herein, we demonstrate the application of collisionally activated dissociation tandem mass spectrometry (CAD MS/MS) to differentiate and quantify the N-methylated histidine isomers, including amino acid residues and actin peptides, directly via the generation of characteristic fragments. We further demonstrate the applicability of various common fragmentation methods, including electron-mediated dissociation (ExD) and photodissociation methods, to differentiate the methylated isomeric products. We also applied CAD MS/MS to quantify the percentage of in actin from a complex proteomic sample. Using the characteristic fragments obtained in the CAD MS/MS, a good linearity relationship ($R^2 > 0.99$) is easily obtained in the calibration curve of tele- and pros-N-methylated H73 peptides. These results demonstrate the applicability of a direct infusion tandem mass spectrometry approach without prior separation, which is a fast, reliable, and robust method to differentiate and relatively quantify the isomeric N-methylated histidine actin peptides.

3.2. Introduction

Methylation is a frequently observed post-translational modification (PTM) in proteins, which plays an essential role in various cellular and biological processes including the regulation of gene transcription,[1] cellular signalling,[2] processing of RNA,[3] protein ageing/repair,[4,5] and the regulation of protein-protein interactions.[6] Muscle contractile proteins, as well as histone proteins and DNA are subject to methylation, which can be either necessary or detrimental to their function.[7-9] Due to the significant effects of methylation on the activities of certain proteins and DNA, this prevalent modification has been implicated in diseases of major concern such as cancer,[10] neurodegenerative disorders,[11] genomic imprinting disorders,[12] and cardiovascular diseases.[13]

In protein methylation, a methyl group (CH_3) is transferred to a nitrogen or oxygen atom (N-methylation or O-methylation respectively) on the amino acid side chains or at the protein N- and C-termini.[14] The transfer of the methyl group is catalysed by specific enzymes called methyltransferases from a primary methyl group donor, known as S-Adenosyl methionine (SAM/AdoMet) to the target substrate (Figure 3.1).[15] Lysine and arginine residues predominantly undergo methylation and the role of lysine and arginine-specific methylation on histone proteins, has been extensively studied.[2,10] However, methylations of other amino acids such as histidine,[16] cysteine,[17] and carboxyl residues[18] have also been reported.

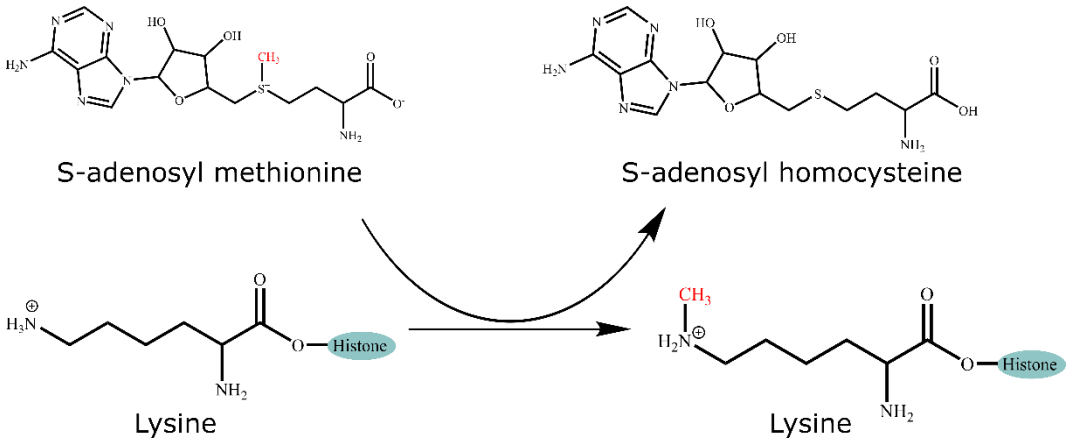


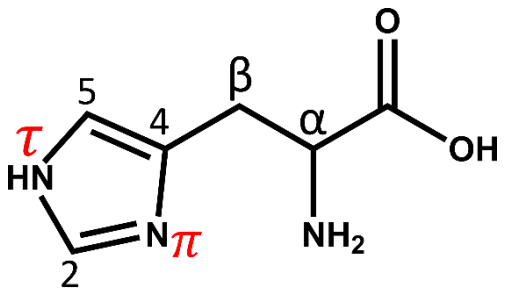
Figure 3. 1 General reaction scheme for histone protein N-methylation on a lysine residue.

Over 50 years ago, histidine methylation was first reported in the muscle proteins, actin[16] and myosin.[19,20] Skeletal muscle tissues are composed of repeating units of both

proteins, which are crucial for muscle contraction. The thin and thick filaments in muscle cells are formed by actin and myosin respectively,[21,22] where both are primarily involved in generating the force required for the contractile response of muscles. Interactions of actin and myosin have also shown to provide support in the cellular functions of non-muscle cells, including cell movement[23,24] and cell division.[25,26] Furthermore, the actin cytoskeleton, which consists of a network of the polymers of actin and actin-binding proteins, is responsible for the movement across cell surfaces, driven by the interactions of actin and myosin interactions as well as the polymerisation of actin filaments.[27]

The detection of 3-methylhistidine (3-MeH) as a free amino acid in urine has been described previously.[28,29] The modified amino acid is often excreted via urine and has often been used as a biomarker for measuring the rate of skeletal muscle protein breakdown.[30,31] Histidine methylation, however, can take place at two different positions on the imidazole ring, and they are therefore noted as 1-methylhistidine (1-MeH) and 3-methylhistidine (3-MeH). In contrast, 1-MeH has been used as a biomarker for meat consumption, as it is obtained from dietary sources.[32,33] In 1967, Johnson *et. al.* isolated and detected 3-MeH in rabbit skeletal muscle actin via paper electrophoresis.[16] 3-MeH was obtained from an isolated tryptic peptide, where it appears to be restricted to the H73 residue in the rabbit actin sample and the ratio of 3-MeHis:His was determined to be 1: 7:6. [16] In previous studies, 3-MeH was also detected in myosin, however, the amount detected was minimal compared to actin and the content was also observed to decrease on purification.[16,34,35]

As mentioned earlier, methylation can take place at two different positions on the nitrogen atom on the imidazole ring of histidine (Figure 3.2). Different systems for numbering the atoms on the imidazole ring have been used previously by biochemists and organic chemists, where the nitrogen atom adjacent to the side chain was numbered as 1 by biochemists and the same nitrogen atom was numbered as 3 by organic chemists. Therefore, to prevent confusion, numbers will not be used herein to designate the position of the methyl group. In accordance with IUPAC guidelines,[36] the position of the methyl group will be referred to as tele/τ-MeH and pros/π-MeH. The nitrogen atom on the imidazole ring of histidine is denoted by tele ('far', abbreviated as τ) as it is further away from the histidine sidechain compared to the nitrogen atom denoted as pros ('near', abbreviated as π), which is closer to the side chain as shown in Figure 3.2 below.



Histidine

Figure 3. 2 Structure of histidine with τ-Me and π-Me position and numbering according to IUPAC.

The methylation of H73 (MeH73) of actin affects the polymerisation of actin filaments, which is necessary for cell motility.[37,38] His73 methylation is also known to regulate the interdomain flexibility and stability of actin. As mentioned earlier, SAM is generally known as the methyl group donor involved in methylation of mainly arginine and lysins residues, however for methylation of histidine, the SET domain containing 3 (SETD3) was discovered as the first human methyltransferase enzyme, which targeted histidine and was responsible for the methylation of H73 in actin.[39-43]

Methylation of actin at H73 position in the protein sequence can result in either tele-methylhistidine (τ-MeH) or pros-methylhistidine (π-MeH) (Figure 3.3).

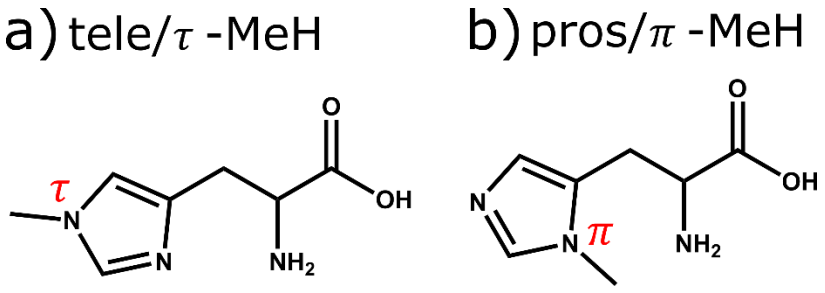


Figure 3. 3 The chemical structures for the modified amino acids a) tele-methylhistidine (τ-MeH) and b) pros-methylhistidine (π-MeH).

Most PTMs can be easily detected using MS instrument due to the mass shift resulted from an additional of the modification. Methylation results in an additional mass shift of 14.015650 Da which can be easily detected by various types of MS instrument. However, τ-MeH and π-MeH are isomeric products with same molecular composition but different methylation position at the side chain, which results in no mass difference

observes in a MS spectrum. Additional analyses, including prior GC or LC separation and MS fragmentation, are required for the differentiation.

The amino acid residues of τ-MeH and π-MeH metabolites have been qualitatively and quantitatively analysed previously using synthetic standards of the compounds and hyphenated mass spectrometry techniques such as GC-MS[44,45] and LC-MS. [46,47] The metabolites are usually present in plasma and urine hence chromatographic separation is often required to reduce sample complexity and interference from other compounds present in the matrices. As the position of the methyl group on the modified histidine cannot be assigned only based on MS screening, the isomeric products differentiation is then solely determined by the difference of retention times obtained from GC or LC during compound elution. Other methods also include derivatization of the metabolites prior to chromatographic analysis to aid separation and identification.[48]

These methods are useful for the separation of the target compounds from other substances in the matrices. However, optimisation of the gradient and lengthy chromatographic separation is often required to ensure adequate detection of the species of interest. The retention times of compounds are also unique to each of the chromatography system, indicating various results maybe obtain from different research laboratories which results in poor reproducibility. In addition, it is challenging for these methods to identify the specific methylation position on the histidine residue, which means that synthetic standards are commonly required. Chemical derivatization can improve the ionisation efficiency of the compounds of interest and prevent interferences from other metabolites in the media by shifting the *m/z* range to a higher *m/z* values. However, derivatization involves an additional step to the sample preparation and it may lead to mixtures of partially and fully derivatized analytes, hence the reaction would need to be carefully optimised.[49]

Tandem mass spectrometry (MS/MS) has been developed over the years with the successful implementation and application of a wide range of fragmentation methods. MS/MS spectra generated by different fragmentation techniques can provide valuable structural information that is useful for peptide sequencing, identification of PTMs, as well as PTM site localization.[50-53] Common MS/MS methods used in this work include collisionally activated dissociation (CAD), infrared multiphoton dissociation (IRMPD), ultraviolet photodissociation (UVPD), and electron-based dissociation (ExD) methods such as electron capture dissociation (ECD) and electron induced dissociation (EID).[54,55]

CAD is a slow-heating activation method, available on most mass spectrometers, which utilises inert gas molecules for collisions with the species of interest. ExD methods such as ECD and EID use low and high energy electrons respectively, which interact with trapped precursor ions resulting in fragmentation. Photodissociation methods such as IRMPD and UVPD depend on the absorption of IR or UV photons by the analyte ions for dissociation to occur. The ability to differentiate between isomers has been shown previously using MS/MS on molecules such as leucine and isoleucine using CAD[56] as well as the isomeric products of deamidation (aspartic acid and isoaspartic acid) generating characteristic fragments for isoaspartic acid (c_n + C_2O_2H and z_{i-n}-C_2O_2H) via electron transfer dissociation (ETD) and ECD, which allow for the reliable discrimination between the isomers.[57,58]

Herein, we demonstrate the application of various fragmentation methods on an FT-ICR mass spectrometer including slow heating activation, electron-mediated and photodissociation methods for the differentiation of the isomeric τ-MeH and π-MeH amino acid residues as well as peptides of actin. The applied fragmentation methods are optimised for the detection and assignment of diagnostic fragments. The different MS/MS methods are also evaluated for their performance in distinguishing between the isomeric τ-MeH and π-MeH peptides of actin. Mixtures of the standards τ-MeH and π-MeH peptides are prepared for relative quantification using the characteristic fragments observed for one of the isomeric peptides by the selected MS/MS method. Calibration curves are generated and the relative quantification method is used to determine the percentage of τ-MeH and π-MeH in the tryptic digested actin peptides derived from rabbit, chicken, bovine, human platelet, and human recombinant actin for isomeric product quantification. The experimental results showcase the applicability and need for MS/MS methods in distinguishing between the methylated histidine isomers of actin peptides using critical isomer-differentiating fragments, which can further be applied to the relative quantification of the isomers in biologically significant samples.

3.3. Experimental Section

Sample preparation for amino acids

τ-MeH and π-MeH amino acids (Sigma Aldrich Ltd, Dorset, England) were provided by collaborators from Warwick medical school. Samples were diluted to 10 µM with methanol and 0.1 % formic acid prior to direct infusion MS and MS/MS analysis.

Sample preparation for synthetic peptides

τ-MeH and π-MeH peptides with the sequence YPIEHGIVTNWDDMEK were synthesised by GenScript Biotech Corporation (Netherlands). RP-HPLC was carried out by the supplier to estimate the purity percentage of the synthetic peptides based on UV absorbance at a wavelength of 220 nm. The percentage peptide purities were determined to be 95.4% for the τ-MeH peptide with the sequence YPIEτ-MeHGIVTNWDDMEK and 95.6% for the π-MeH peptide with the sequence YPIEπ-MeHGIVTNWDDMEK.

The synthetic peptides were dissolved in Milli-Q (Direct-Q® 3 UV System, Millipore Corporation, US) H_2O (~pH 7) at a concentration of 200 µM for storage in -80 °C freezer. The 200 µM stock solutions of each peptide were then diluted to 5 µM with 75:25:0.1 water/acetonitrile/formic acid prior to the MS and MS/MS experiments.

The calibration curve for the relative quantification by CAD MS/MS was obtained by mixing the τ-MeH and π-MeH containing peptides at 20 %, 40 %, 50 %, 60 %, 80 %, and 100 % with concentration at 10 µM.

Sample preparation for actin digestion

Actin powder samples from bovine heart muscle, rabbit skeletal muscle, chicken gizzard smooth muscle and human platelet (>99% purity for each protein determined by scanning densitometry of Coomassie Blue stained SDS-PAGE gel from the supplier) (Cytoskeleton, Inc., Denver, USA) were dissolved in 100 mM ammonium bicarbonate (ABC) solution to 1 µg/µL. Disulphide bonds were then reduced using 50 mM dithiothreitol (DTT, Sigma Aldrich Company Ltd.) for 30 minutes at 60 °C, followed by alkylation with 100 mM of iodoacetamide (IAA, Sigma Aldrich Company Ltd.) and the samples were stored in the dark at room temperature for 1 hour. The solution was then tryptic digested with 1 mg/mL trypsin (Sigma Aldrich Company Ltd.) in 100 mM AmmBic at 37 °C for 16 hours. After the tryptic digestion, samples were desalted using

SOLAμ SPE C18 cartridges (ThermoFisher, Waltham, MA, USA) with the elution buffer 80 % ACN and 0.1 % formic acid. The desalted samples were further diluted with 20 % ACN and 0.1 % formic acid into final concentrations of 0.2 μg/μL for direct infusion MS and MS/MS analysis. Purified human recombinant actin was provided and prepared by the Warwick medical school as previously described.[59] The sample was then prepared for tryptic digestion and direct infusion MS and MS/MS analysis in an identical manner to the actin powder samples.

Nano-LC separation

A homemade C18 RP nano capillary trap column (3cm, 150 μm I.D., 3 μm particle size, 300 Å pore size) and an in-house packed C18 RP nano capillary analytical column (20cm, 75 μm I.D., 3 μm particle size, 300 Å pore size) using Jupiter C18 packing material (3 μm particle size, 300 Å pore size; Phenomenex, UK) was packed with an ultrahigh constant pressure pump (Teledyne SSI, US) operated at 5000 psi. The columns were connected to an EASY nLC II system (Proxeon, Hemel Hempstead, UK) for the online nLC separation of the tryptic digested actin samples. The sample was separated with solvent A (5% ACN, 0.1% formic acid in water) and solvent B (95% ACN, 0.1% formic acid in water) on the nLC system. 15 uL of a 2 μM solution of the tryptic digested human platelet and human recombinant actin samples were loaded onto the C18 RP column. The nLC gradient was optimised for the digested actin samples as follows: 0-10 min, 5% B, 500 nL/min; 10-70 min, 5-40% B, 300 nL/min; 70-75 min, 40% B, 300 nL/min; 75-90 min, 40-80% B, 300 nL/min; 90-135 min, 80% B, 600 nL/min; 135-136 min, 80-5% B, 600 nL/min; 136-150 min, 5% B, 600 nL/min. The total run time of one sample is 150 minutes which includes sample loading, nLC separation, and the analytical column equilibration time.

A 50/50 mixture of the synthetic τ-MeH and π-MeH target peptide was also analysed via nano-LC. The 200 μM stock solutions of each synthetic peptide was diluted with H_2O into 10 nM solutions to match the peptide intensity in the digested human actin samples. 50 % of 10 nM τ-MeH modified synthetic peptide was mixed with 10 nM π-MeH modified synthetic peptide. 15 uL of the mixed solution was loaded onto the C18 RP column. The sample nano-LC gradient optimised for the digested human actin samples was applied to the 50/50 τ-MeH and π-MeH target peptide mixture.

The nLC separation was automatically controlled by the Bruker Daltonics Hyster automation software (Bruker Daltonics, Bremen, Germany). The nLC platform was

coupled to the 12 Tesla Bruker SolariX FT-ICR-MS via a custom-made nanospray source using pre-cut conductively coated SilicaTipTM emitters of 5cm long, with a 360 μm tip O.D., 20 μm I.D., 10 μm tip I.D. (New Objective, MA, USA).

FT-ICR MS analysis

All the experiments were carried out using a 12 tesla (T) SolariX Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS; Bruker Daltonik GmbH, Bremen, Germany), equipped with a shielded superconducting magnet.

Direct infusion experiments

For the direct infusion experiments, the samples were loaded into borosilicate glass capillary tips (purchased from World Precision Instruments, Inc., Sarasota, FL, USA), which were pulled using a Sutter P-97 capillary Flaming/Brown micropipette puller instrument (Sutter instruments Co., Novato, CA, USA). The pulled tips were optimised for a low-flow nano-electrospray ionisation (nESI) experiments.

All samples were sprayed in positive ionisation mode. Mass spectra were acquired with a 4 mega-word (MW) data-points (32 bits) over a mass range of *m/z* 147 – 3,000 to produce a 1.68 s transient and ~460,000 resolving power at *m/z* 400.

Positively charged ions were transmitted through a glass capillary to a quadrupole and then externally accumulated in a hexapole collision cell for 0.35 s before transferred to an infinity cell for MS excitation and detection.

For CAD MS/MS experiments, 3+ precursor ions of the synthetic peptides and actin digested peptides were first quadrupole isolated at *m/z* 654.3 with an isolation window of 5 *m/z*. The ions then underwent collisions with argon gas in the collision cell. The optimised collision energy (CE) of 17 V was applied to the peptides. Fragments, together with the precursor ions, were then transferred to the infinity cell for mass detection.

For IRMPD MS/MS, precursor ions were isolated with the quadrupole and then transmitted to the ICR cell. The ions were then fragmented using a continuous-wave CO_2 laser (Synrad, Washington, USA) with an output wavelength of 10.6 μm. An optimised pulse length of 0.18 s and 50 % laser power from a 25 W laser was used for fragmentation. The CO_2 laser was introduced from the back of the ICR cell through a

BaF_2 window and precursor ions were fragmented inside the ICR cell before excitation and detection.

For ECD MS/MS experiments, after quadrupole isolation, ions were directly transferred to the ICR cell, where they were irradiated with low energy electrons emitted from an indirectly heated hollow dispenser cathode heated via 1.5 A continuous current for ion fragmentation. The optimised ECD parameters for the peptide fragmentation were electron irradiation time of 0.25 s, extraction lens at 3.0 V, and cathode potential of 1.5 V.

With the pre-existing IRMPD setup, a 193 nm ArF excimer laser beam (10 Hz; Coherent, UK) was also introduced from the back of the ICR cell through a BaF_2 window. Like the IRMPD MS/MS experiment, ions were first isolated in quadrupole, transmitted to the ICR cell, and eventually irradiated with 1 laser shot (~5 mJ/pulse measured at the laser head) to generate fragments.

For 213 nm UVPD fragmentation, a stable telescopic compact high energy Q-switched pulsed Nd:YAG laser with an output wavelength of 213 nm (5^{th} harmonic of the Nd:YAG laser) (10 Hz; Litron Lasers, UK) was also used and ions were irradiated with 15 laser shots (~1.5 mJ/pulse at the laser head).

nLC MS/MS experiment

Mass spectra were acquired with 1 M data-points with a 0.42 s transient and auto MS/MS was applied. All eluted peptides were analysed in positive ionisation mode with a capillary voltage of 1.3 kV and the source temperature at 180 ^{o}C. Ions were externally accumulated in a hexapole collision cell for 0.35 s before transferred to an infinity cell for MS excitation and detection.

Throughout the nLC experiment, one MS scan was followed with one auto MS/MS scan. 0.3 s accumulation time was used for MS scan. Auto MS/MS with 5 *m/z* isolation window was then applied to the peptides that were in the *m/z* range of 654 – 661 and the intensities were higher than 2E6. Collisionally activated dissociation tandem MS (CAD MS/MS) with a rolling energy curve was used to fragment the target peptides with an accumulation for 1.5 s. All the ions were then excited and detected in the infinity cell.

Data analysis

All spectra were analysed using DataAnalysis 4.3 (Bruker), internally calibrated and fragments were assigned manually with a mass error <3 ppm (supplementary table S3.3– S3.19).

All spectra were internally calibrated with known *m/z* fragmented peaks that contain minimum threshold of S/N >3 and peaks were picked with relative intensities higher than $1x10^6$ according to the Bruker FTMS peak picking algorithm.

The CAD fragments that were observed in both digested actin and mixed synthetic peptide solutions (highlighted in the assignment tables) were used for calculating the relative abundance of τ-MeH modified peptide present in the mixed solution.

7-point calibration curves were generated by mixing the ratio of the τ-MeH and π-MeH synthetic peptides (YPIEHGIVTNWDDMEK) according to the Table 3.1 for the CAD MS/MS quantification experiments.

Table 3. 1 Percentage of each synthetic peptide in mixtures for MS/MS quantification experiments.

Calibration Point	YPIE(τ-MeH)GIVTNWDDMEK (%)	YPIE(π -MeH)GIVTNWDDMEK (%)
1	0	100
2	20	80
3	40	60
4	50	50
5	60	40
6	80	20
7	100	0

The CAD fragments that were observed in both digested actin and mixed synthetic peptide solutions were used for calculating the relative abundance of τ-MeH modified peptide present in the mixed solution. Since the CAD MS/MS fragment *m/z* 327.67 (b_5^{2+}), obtained from the YPIEHGIVTNWDDMEK peptide could only be observed in the τ-MeH modified peptide, not in the π-MeH modified peptide; a calibration curve was then built by calculating the percentage of peak area of *m/z* 327.67

fragments against the peak area sum of 18 other fragments observed in the MS/MS spectra.

Hence the relative abundance of the τ-MeH peptides was estimated using the following equation [Eqn. 3.1]:

$$Relative\ percentage\ ratio\ of\ \tau - MeH\ (\%) = \frac{peak\ area\ of\ [b_5]^{2+} or\ [y_{11}]^{+}}{sum\ of\ all\ fragment\ peak\ areas} \times 100 \quad \text{[Eqn. 3.1]}$$

Each calibration point in the quantification experiment was calculated by averaging the peak area from 9 CAD MS/MS spectra. The R^2 of the calibration curve was then calculated by linear fitting of the curve. For the actin digested samples, 9 CAD MS/MS spectra were also used for calculating the relative abundance of τ-MeH peptide in the solution. The equation of the calibration curve was then used to determine the percentage of the τ-MeH modified peptide generated in the actin digest samples.

All calibration curves were plotted and a linear fit was applied using the software package of Origin 2019 (OriginLab Corporation, USA).

3.4. Results and Discussion

Direct infusion MS and EID MS/MS analysis of isomeric τ-MeH and π-MeH amino acid residues

Methods such as GC-MS and HPLC generally require derivatization of τ-MeH and π-MeH, focusing on the retention time differences in the chromatograms as a mean for differentiation. In this work, the methylhistidine isomers were first analysed at the amino acid level via direct infusion MS and EID MS/MS. For both τ-MeH and π-MeH, the protonated molecular ions were observed at *m/z* 170.092403 in the mass spectra, which were identical (Supplementary Figure S3.1) therefore τ-MeH and π-MeH cannot be distinguished from one another solely based on MS screening analysis. The molecular ion was quadrupole isolated (Supplementary Figure S3.2) and after optimisation of the electron energy, EID MS/MS spectra were generated (Figure 3.4). The EID MS/MS spectra for both modified amino acids resulted in 8 assignable fragments for π-MeH and 5 fragments were assigned for τ-MeH (Supplementary Table S3.1– S3.2).

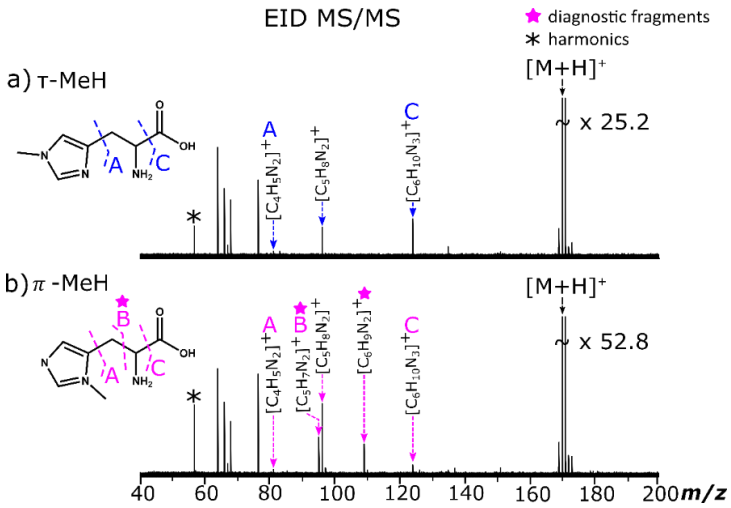


Figure 3. 4 EID spectra of a) τ-MeH and b) π-MeH. Peak assignment tables for the assigned EID MS/MS spectra are provided with absolute average mass errors approximately $< 0.3 \pm 0.4$ ppm for both amino acids (Supplementary Table S3.1– S3.2).

In principle, EID fragments are very similar to those generated by electron ionisation (EI) as both techniques use electrons with high energies to cause fragmentation. Previous experimental EI spectra obtained from GC-MS data did not demonstrate the differences in the non-derivatised methylhistidine isomers as focus was largely on the

differences observed in the retention time of the compounds that are normally derivatised.[48,60,61]

Predicted EI spectra available from the Human Metabolome Database (HMBD) suggest that the fragmentation patterns for the underivatized τ-MeH and π-MeH are generally the same with only a difference of 1 fragment noted in each EI spectra.[62] In the predicted EI spectrum for τ-MeH, a fragment at *m/z* 82.0525 ($C_4H_6N_2^{+\cdot}$) is present, which is absent in the π-MeH spectra whereas an EI fragment at *m/z* 141.06585 ($C_4H_6N_2O_2^{+}$) is only observed for π-MeH. The predicted diagnostic fragments are absent from the EID MS/MS spectra obtained in this work. However, in the EID MS/MS spectrum, 2 unique fragments were observed for π-MeH, which were absent in the τ-MeH spectrum. In the predicted EI spectra from the HMBD, fragment B in the EID spectrum of Figure 3.4b is expected to be present for both isomers and whereas the diagnostic fragment with the molecular formula $C_6H_9N_2$ for π-MeH in Figure 3.4b is absent in the predicted EI spectra.

However, it should be noted that even though we cannot determine which nitrogen of the imidazole ring the methyl group is attached to as there are no fragments formed due to cleavage across the imidazole ring, two diagnostic fragments (fragment B and the fragment with the molecular formula $C_6H_9N_2$) were observed only in the π-MeH EID spectrum, which can be used to differentiate between the isomeric τ-MeH and π-MeH amino acids based on the differences in their relative intensities. The peaks observed in the region *m/z* 60-70 may be designated as electronic noise peaks as they were present in the acquired spectrum in the absence of the ionised sample. The electronic noise peaks are also typically caused by radiofrequency interference (RFI), but they do not overlap with any fragments in these fragmentation spectra and are easily distinguishable from real peaks as they have no isotopes.[63]

Direct infusion MS and MS/MS analysis of isomeric τ-MeH and π-MeH synthetic actin peptides

The modified methylhistidine isomers of actin peptides were then placed under consideration as methylation at a specific position (H73) in actin has been linked to significant biological and functional effects of the protein as mentioned previously.[23,35,64-66] The peptide with the sequence YPIEH(Me)GIVTNWDDMEK from actin, in which methylation takes place at H73 is isolated and studied herein. The synthetic actin peptides have the same sequence and the only difference between them is the position of the methyl group on the nitrogen atom of the imidazole ring in histidine, where methylation occurs at either the τ or the π position resulting in the isomeric peptides; YPIEH(τ-Me)GIVTNWDDMEK and YPIEH(π-Me)GIVTNWDDMEK. The isomeric actin peptides were subjected to MS analysis, resulting in identical mass spectra (Supplementary Figure S3.3), thus discrimination between the τ-MeH and π-MeH peptides cannot be achieved simply based on the mass spectra. The triply charged molecular ion, $[M+3H]^{3+}$ was observed at *m/z* 654.308498 and the protonated doubly charged molecular ion, $[M+2H]^{2+}$ was observed at *m/z* 980.959108 in the mass spectra for both peptides.

The 3+ precursor ion at *m/z* 654.3 was quadrupole isolated for both peptides and CAD MS/MS was applied to each peptide individually. Figure 3.5 depicts the assigned CAD MS/MS spectra, where predominantly b/y fragments are observed as expected of the type of fragmentation method used and the cleavage coverage for both peptides is 100 %.

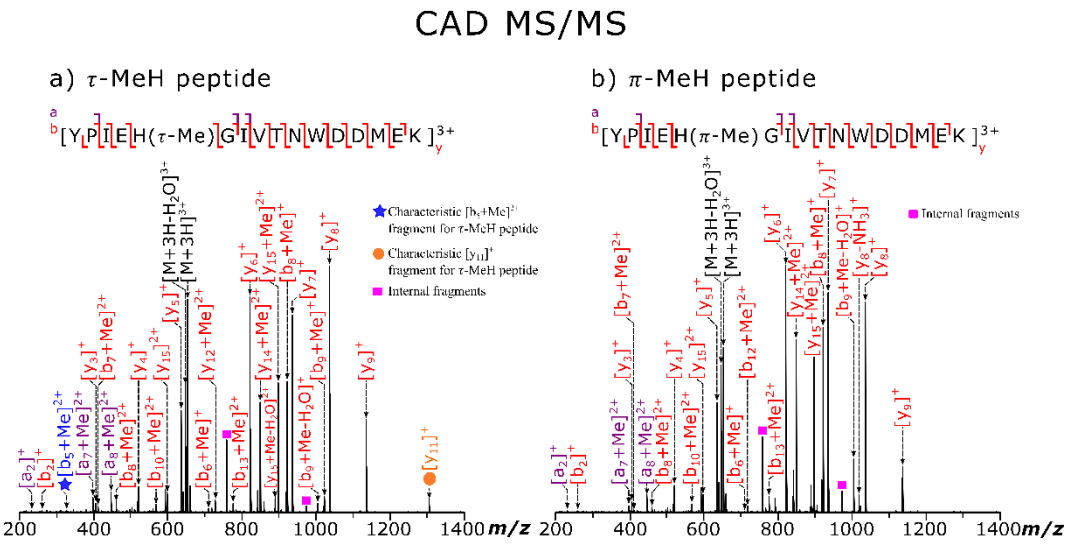


Figure 3. 5 CAD spectra of a) τ-MeH and b) π-MeH modified actin peptides.

After close inspection and comparison between the τ-MeH and π-MeH peptide CAD spectra, diagnostic fragments were detected for τ-MeH peptide, which were absent or at significantly lower intensity in the π-MeH peptide spectra as shown in Figure 3.6. The two diagnostic fragments detected for the τ-MeH peptide are the $[b_5+Me]^{2+}$ fragment ion at *m/z* 327.665830 and the $[y_{11}]^{+}$ fragment ion at *m/z* 1307.593630.

The relative intensity of the characteristic fragment ions to the precursor ion are 0.6 % and 6.1 % for the $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragments in τ-MeH CAD MS/MS spectrum, respectively (Figure 3.6). Both $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment peaks were easily observed and peak-picked in the τ-MeH CAD MS/MS spectrum as the S/N ratio were 395.4 and 2062.5 respectively after 200-scan accumulation. On the other hand, $[b_5+Me]^{2+}$ fragment was absent in the π-MeH CAD MS/MS spectrum (Figure 3.6a); while the relative peak intensity of $[y_{11}]^{+}$ fragment was significantly reduced to 0.06% compared to the precursor ion intensity after 200-scan accumulation (Figure 3.6b). Considering the relative peak intensity of other fragment peaks, including $[y_4]^{+}$, $[y_5]^{+}$, $[y_6]^{+}$, and $[y_7]^{+}$ (Supplementary Table S3.3), there are no significant difference between τ-MeH and π-MeH peptide CAD MS/MS spectra, indicating $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragments generated from the 3+ precursor ions were the characteristic fragments to differentiate τ-MeH and π-MeH peptides.

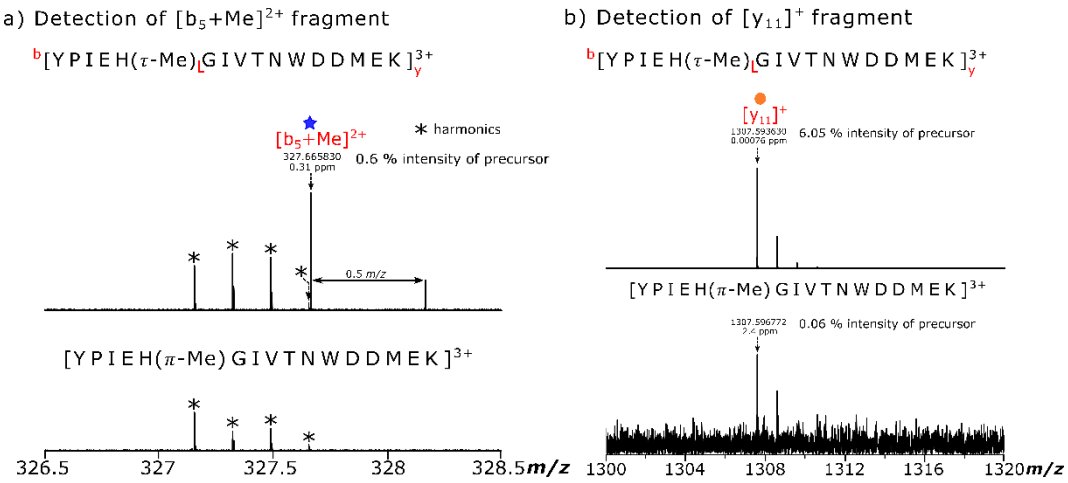


Figure 3. 6 Zoom in of the characteristic fragments at *m/z* 327.6 ($[b_5]^{2+}$) and *m/z* 1307 ($[y_{11}]^{+}$) present in the CAD MS/MS spectra of the τ-MeH synthetic peptide, which are absent or very low intensity for the π-MeH peptide.

To prevent the characteristic fragments were only generated under specific energy which results in poor reproducibility, various collision energies were used to fragment the τ-MeH and π-MeH peptides. Based on the experiments in the above section, the optimised CE was at 17V which resulted in the highest S/N ratio of most of the fragments. For $[b_5+Me]^{2+}$ fragment of τ-MeH peptide, the highest S/N ratio (S/N = 395) was obtained from the CAD MS/MS spectrum with 17V collision energy; while the under- (CE = 14V) and over-fragmentation (CE = 22 and 25) could still obtain a peak intensity with S/N over 100 (Figure 3.7a). In contrast, $[b_5+Me]^{2+}$ fragment has never appeared in the CAD MS/MS spectra π-MeH peptide regarding the fragmentation energies, indicating fragmentation energy was not a factor to generate the characteristic $[b_5+Me]^{2+}$ fragment. Similar result was observed from $[y_{11}]^{+}$ fragment between τ-MeH and π-MeH peptides CAD MS/MS spectra (Figure 3.7b). Even though $[y_{11}]^{+}$ fragment from τ-MeH peptide obtained the highest S/N ratio (S/N = 3366) when CE equalled to 22, the under- (CE = 14V and 17V) and over-fragmentation (CE = 25) could still generate S/N over 350. Comparing to τ-MeH peptide, $[y_{11}]^{+}$ fragment of π-MeH peptides were significant reduced with the highest S/N below 50, showing this characteristic fragment ($[y_{11}]^{+}$) can also be a potential marker for τ-MeH peptide regarding the change in collision energy.

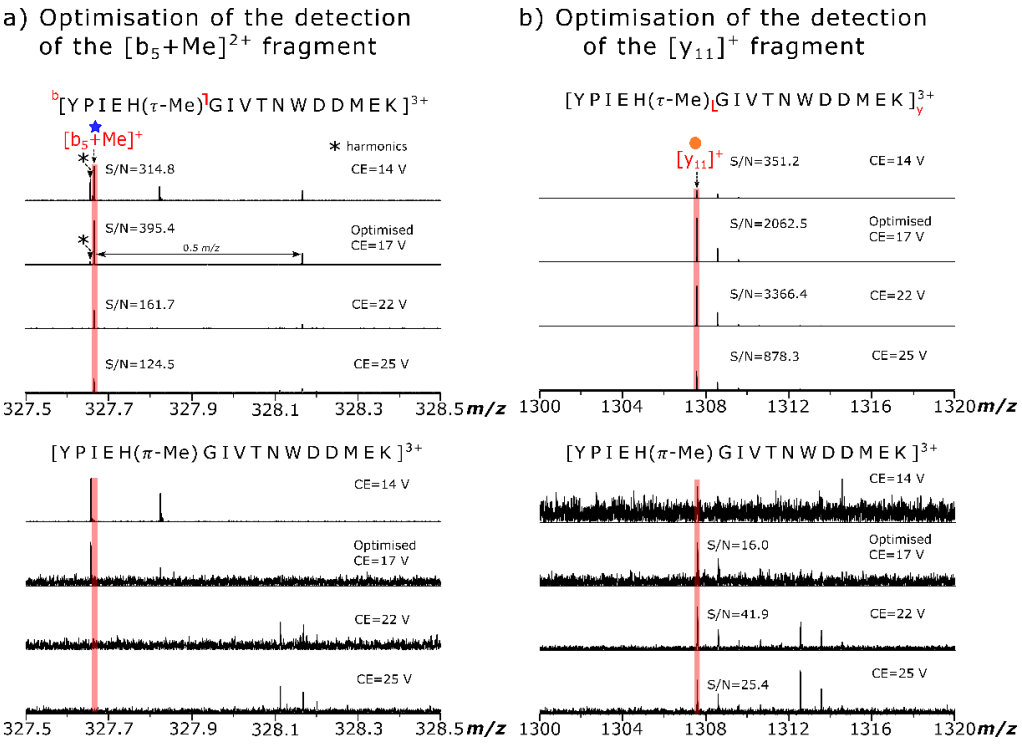


Figure 3. 7 Collision energy (CE) optimisation demonstrated for characteristic fragments at *m/z* 327.6 ($[b_5]^{2+}$) ion and at *m/z* 1307 ($[y_{11}]^{+}$) detection for the CAD MS/MS experiments.

As expected, application of IRMPD MS/MS to the isomeric peptides produced mainly b/y fragment ions and 100 % cleavage coverage of the peptide sequence, like CAD MS/MS (Figure 3.8). The characteristic fragments $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ were also observed in the τ-MeH IRMPD MS/MS spectrum, whilst absent in the π-MeH MS/MS spectrum (Supplementary Figure S3.4). With IRMPD MS/MS, the relative percentage intensity of the $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions to the 3+ precursor ion was 4.56 % and 0.22 %, respectively. Interestingly, when compared to CAD MS/MS, the relative percentage intensity of the $[b_5+Me]^{2+}$ to the precursor ion in the τ-MeH peptide IRMPD MS/MS spectrum was 7.6-fold higher in intensity whereas the relative intensity percentage ratio of the $[y_{11}]^{+}$ fragment to the precursor ion was approximately 28-fold lower in intensity compared to the relative fragment intensities generated from τ-MeH peptide obtained using CAD MS/MS, indicating the fragmentation mechanism may be different between CAD and IRMPD MS/MS despite similar fragmentation patterns were observed.

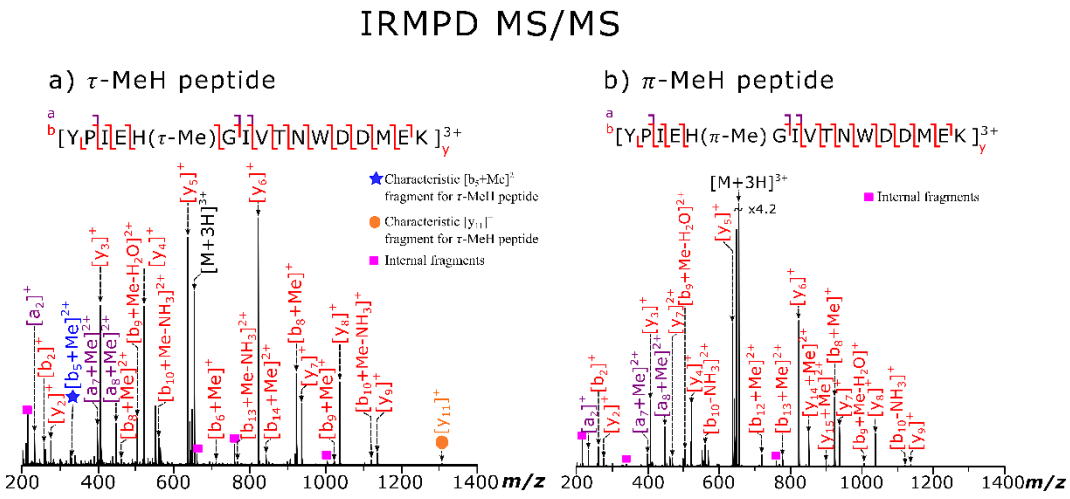


Figure 3. 8 IRMPD spectra of a) τ-MeH and b) π-MeH modified actin peptides.

Application of UVPD MS/MS resulted in extensive fragmentation, as 100 % cleavage coverage was easily obtained for both peptides with 193 nm UVPD MS/MS and a cleavage coverage of 94 % was obtained with 213 nm UVPD MS/MS (Figure 3.9 and Figure 3.10). Additional fragment types were also observed in the UVPD MS/MS spectra for the isomeric peptides. The b/y fragments dominate the spectra due to internal conversion of the photon energy into vibrational modes, which is generally observed with CAD MS/MS and IRMPD MS/MS. However, c/z, a/x, and internal fragments were also present, resulting in complex UVPD MS/MS spectra. A distribution of the different fragment types detected with all the fragmentation methods used for the synthetic τ-MeH and π-MeH peptides is shown in the Supplementary Figure S3.8. With 193 nm and 213 nm UVPD MS/MS, all fragment series were detected, although the fragmentation efficiency for both UVPD MS/MS methods were the lowest (approximately 30 % and lower) compared to CAD and IRMPD MS/MS (Supplementary Figure S3.7).

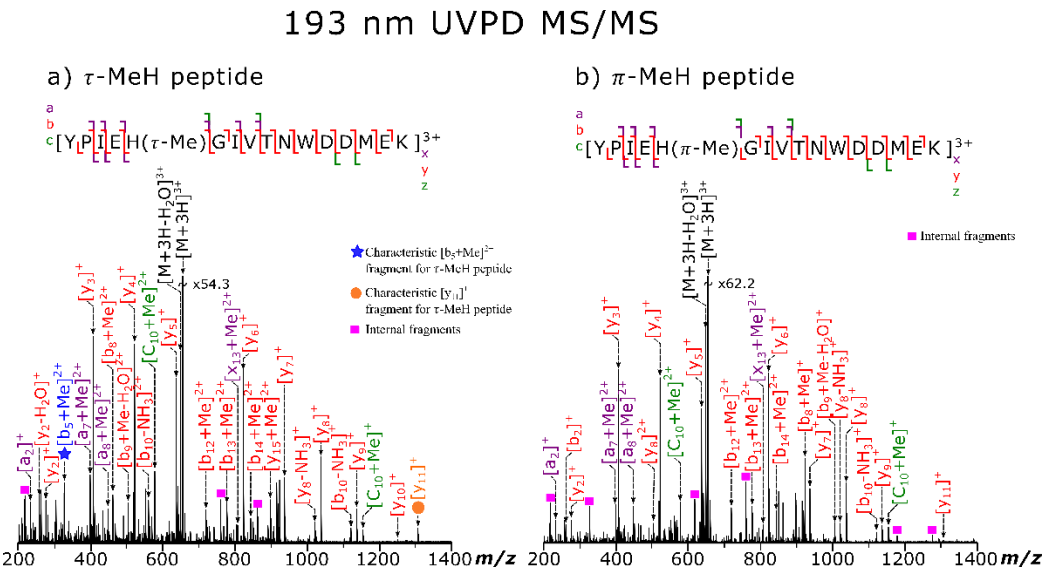


Figure 3. 9 The 193 nm UVPD spectra of a) τ-MeH and b) π-MeH modified actin peptides.

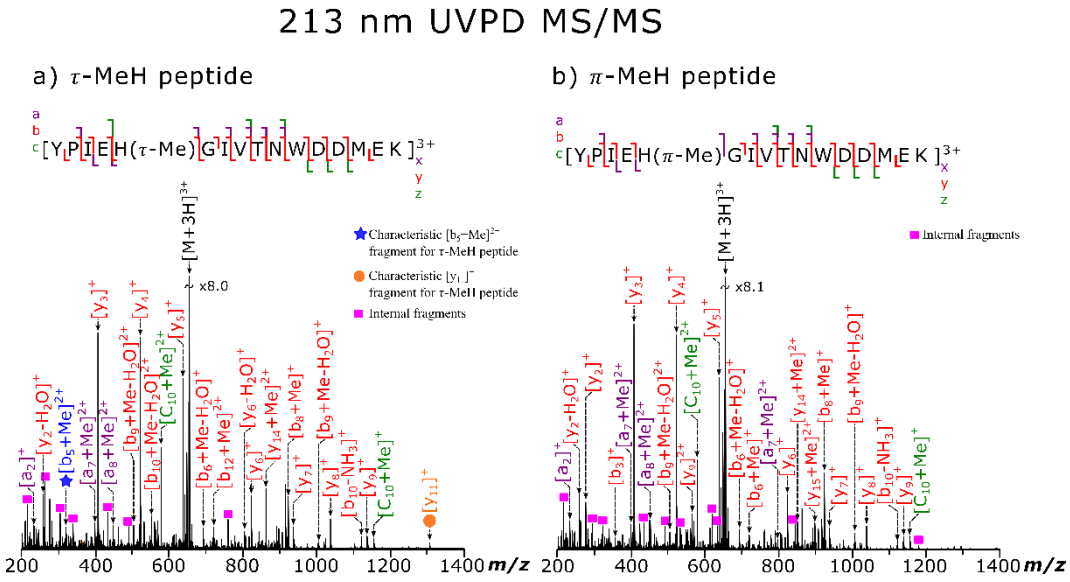


Figure 3. 10 The 213 nm UVPD spectra of a) τ-MeH and b) π-MeH modified actin peptides.

The characteristic $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions were detected when both 193 nm and 213 nm UVPD MS/MS were applied to the τ-MeH peptide and the absence of the characteristic fragments was also observed in the π-MeH peptide UVPD MS/MS spectra (Supplementary Figures S3.5 and S3.6). With 213 nm UVPD MS/MS, for the $[b_5+Me]^{2+}$ and the $[y_{11}]^{+}$ fragment ions, the relative percentage intensities to the precursor ion were approximately 3.8-fold and 2.5-fold higher in intensity than what was observed

with 193 nm UVPD MS/MS. With 193 nm UVPD MS/MS, as mentioned in the experimental section, 1 laser shot was applied at 5 mJ/shot whereas for 213 nm UVPD MS/MS, the isomeric peptides were subjected to 15 laser shots at ~1.5 mJ/pulse. Hence, the differences in intensities may be explained by the number of laser pulses that were used as increasing the number of laser shots can result in extensive fragmentation as well as improve the relative intensities of the observed peaks.

ExD methods were also applied to the τ-MeH and π-MeH modified synthetic peptides as shown by the ECD MS/MS spectra in Figure 3.11 and the EID MS/MS spectra in Figure 3.12. A cleavage coverage of 94 % was observed with ECD compared to 88 % cleavage coverage was observed with EID MS/MS, possibly due to lower fragmentation efficiency obtained with the EID MS/MS experiments. The calculated fragmentation efficiency was approximately 60 % for the τ-MeH peptide and 62 % for the π- with ECD, whereas the fragmentation efficiency was lower for EID, with values of 23 % and 25 % obtained for the τ-MeH and π- MeH peptides, respectively. However, in comparison to ECD MS/MS, as higher electron energies are used for EID MS/MS, denser and complex spectra can be obtained as depicted by Figure 3.12, where predominantly c/z fragments were observed, accompanied with side chain losses, b/y fragments and 15 assigned internal fragments compared to the 2 internal fragments assigned in the ECD MS/MS spectra.

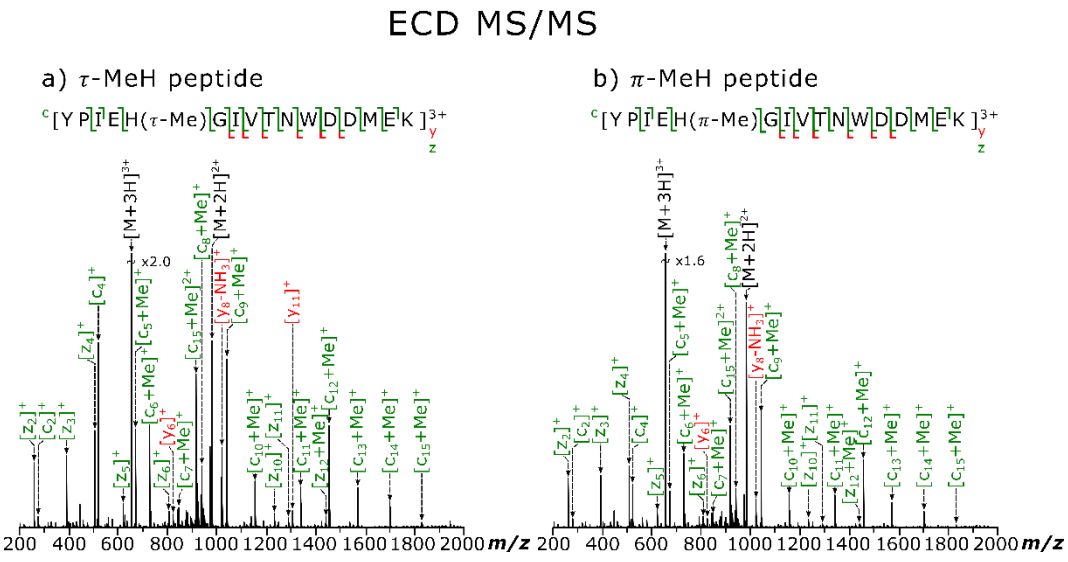


Figure 3. 11 ECD spectra of a) τ-MeH and b) π-MeH modified actin peptides.

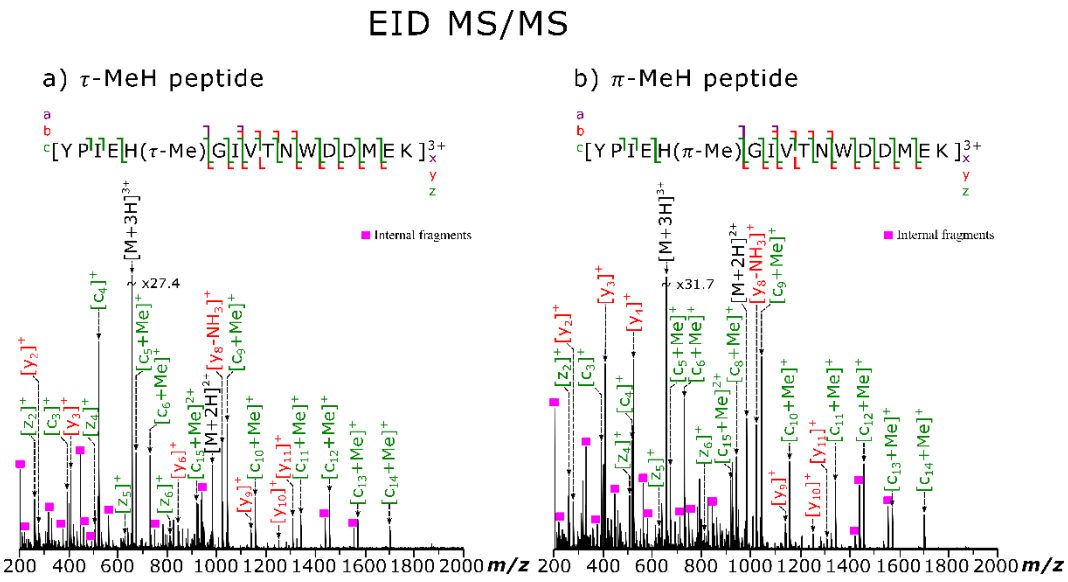


Figure 3. 12 EID spectra of a) τ-MeH and b) π-MeH modified actin peptides.

In the ECD and EID MS/MS spectra, the corresponding fragments to the diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragments previously observed in the CAD, IRMPD and UVPD MS/MS spectra for only the τ-MeH peptide, are the $[c_5+Me]^{2+}$ and $[z_{11}]^{+}$ fragments. However, unlike CAD and the photodissociation fragmentation methods, the $[c_5+Me]^{2+}$ and $[z_{11}]^{+}$ fragments can be observed in the ECD and EID MS/MS spectra for both peptides, hence it is not possible to differentiate between them based on the lack of diagnostic fragments.

In summary, application of all direct infusion MS/MS methods applied herein resulted in rich fragmentation spectra, demonstrating near or complete cleavage coverage of the isomeric peptides. The diagnostic fragments $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragments were observed in the τ-MeH synthetic peptide CAD, IRMPD and UVPD MS/MS spectra, which were absent in the π-MeH peptide spectra. Hence, the diagnostic fragments can be used to distinguish between the isomeric forms of the methylated target actin peptide. Since CAD MS/MS is the most general fragmentation technique applied in various types of MS instrument, the following complex sample experiments will solely focus on using CAD MS/MS to differentiate and quantify the τ-MeH and π-MeH peptides.

Detection of the target peptide and diagnostic τ-MeH peptide fragments in tryptic digested human actin samples using nano-LC coupled to FT-ICR MS

The synthetic τ-MeH and π-MeH actin peptides were analysed as 50/50 mixture under the same LC, MS, and CAD MS/MS conditions as the tryptic digested human actin samples. No differentiation between the τ-MeH and π-MeH peptides was observed in the retention time as the EIC of the 3+ target peptide precursor ion resulted in the presence of only one peak in the chromatogram at 21.25 mins, which is clearly demonstrated by Figure 3.13 below.

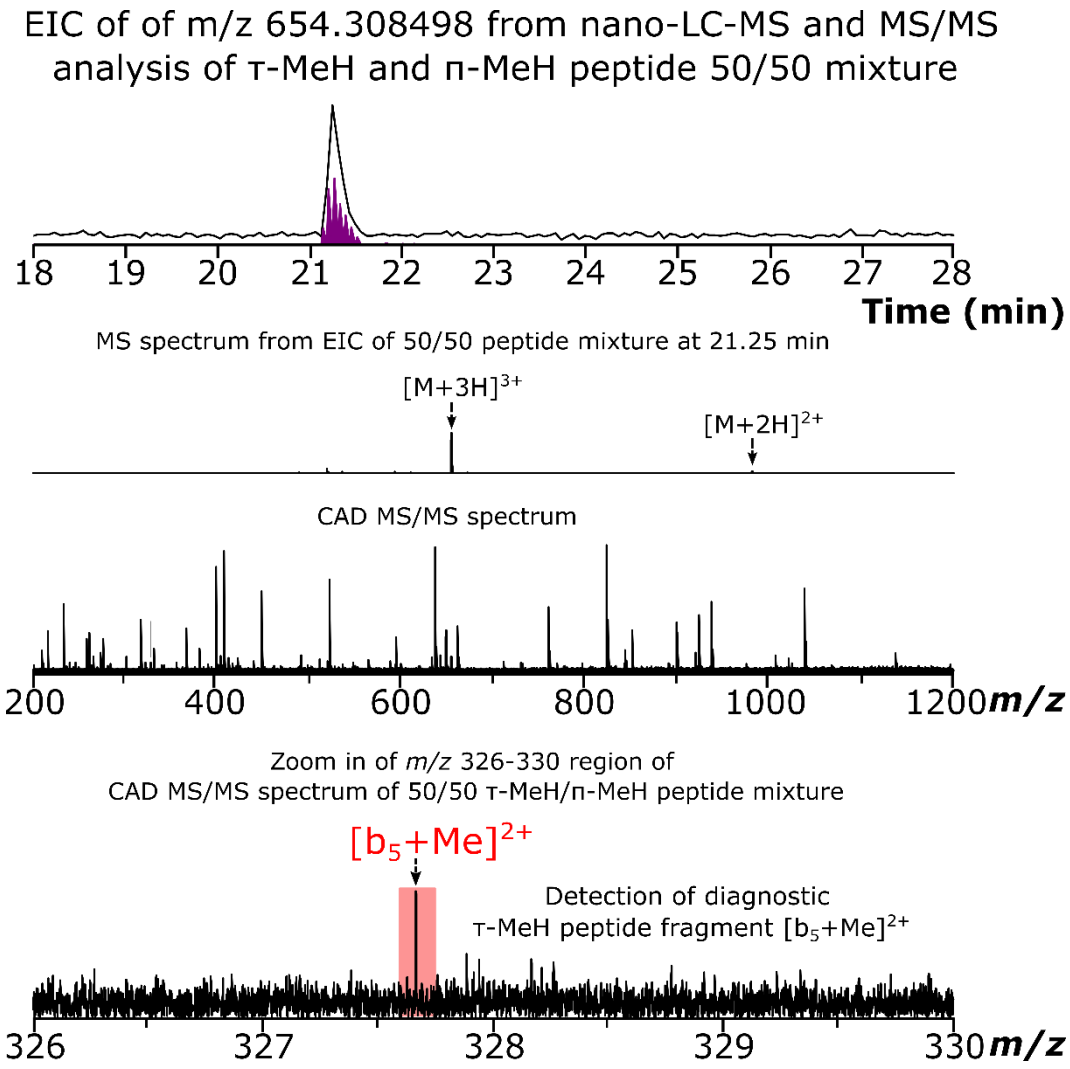


Figure 3. 13 The nLC-FT-ICR-MS and CAD MS/MS results for 50/50 mixture of synthetic τ-MeH and π-MeH actin peptides, depicting the MS obtained from the EIC of the target peptide, the CAD MS/MS spectrum, and the presence of the diagnostic $[b_5+Me]^{2+}$ fragment ion.

The mass spectrum in Figure 3.13 is identical to the mass spectra obtained for the individual τ-MeH and π-MeH synthetic peptides via direct infusion nanoESI-MS (Supplementary Figure S3.3) with the 3+ precursor ion present at *m/z* 654.308498 and the 2+ precursor ion present at *m/z* 980.959108. Application of CAD MS/MS resulted in the presence of the confirmatory diagnostic $[b_5+Me]^{2+}$ fragment ion (bottom spectrum in Figure 3.13), which can be easily attributed to the 50 % τ-MeH peptide in the synthetic peptide mixture.

Subsequently, the tryptic digested human platelet actin and human recombinant actin were analysed via nano-LC MS and CAD MS/MS. Although the target peptide was detected in both samples, a difference in the elution profile was noted as the target peptide eluted earlier for the human platelet actin (65.79 mins) than for the human recombinant actin sample (75.71 mins). The nLC-MS and CAD MS/MS results for the tryptic digested human platelet actin and human recombinant actin are shown in Figure 3.14 and Figure 3.15, respectively.

The TIC revealed 10 additional peaks in the chromatogram for the recombinant actin sample compared to 4 additional peaks in the TIC for the platelet actin. The MS acquired from the EIC of the target peptide (*m/z* 654.308498) at 75.71 mins was equally dense for the human recombinant actin sample as shown by Figure 3.14. The issue with peptide co-elution is also demonstrated herein and in Supplementary Figure S3.9 as another peak (3+ peptide ion at *m/z* 652.0263173) is present in the MS and close to the triply charged precursor ion of the target peptide (Figure 3.16).

Application of CAD MS/MS to the human actin samples clearly highlights the presence of the diagnostic τ-MeH peptide CAD fragment ion ($[b_5+Me]^{2+}$) in the bottom spectra of Figure 3.14 and Figure 3.15. However, the intensity of the diagnostic $[y_{11}]^+$ fragment ion in the MS/MS spectra was very low and below the chosen intensity threshold therefore the FTMS peak picking algorithm was unable to select them. Hence, the $[y_{11}]^+$ fragment ion was disregarded due to either very low intensity in the nanoLC-MS/MS spectra for the human actin tryptic digest samples or the absence in the synthetic τ-MeH peptide nanoLC-MS/MS spectra.

More significantly, the isomeric τ-MeH and π-MeH modified target peptides were not separated in the retention time for either of the human actin digest samples nor the 50/50 τ-MeH and π-MeH synthetic peptide mixture as shown by the EIC of the target peptide *m/z* in Figures 3.13, 3.14, and 3.15.

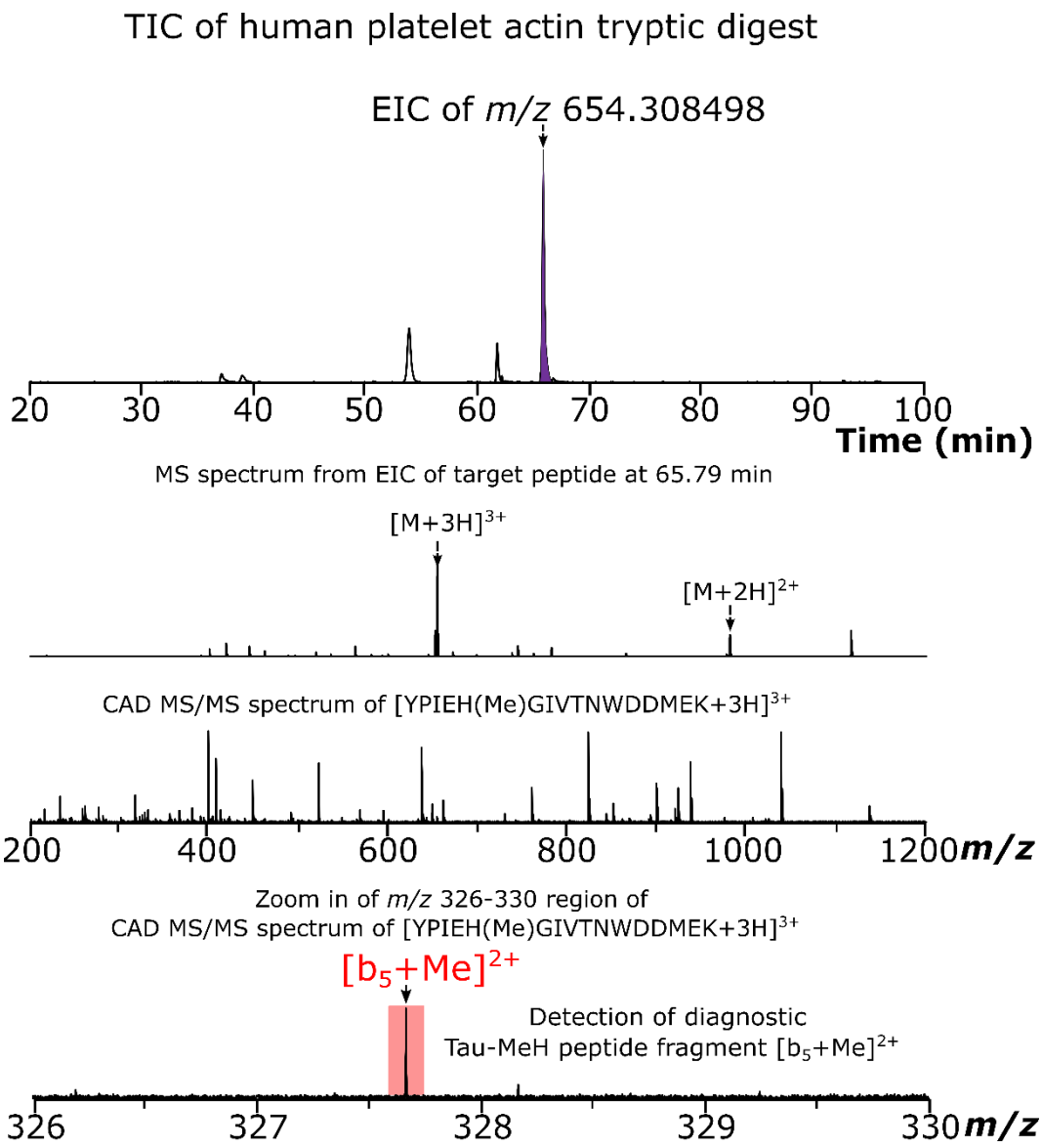


Figure 3. 14 The nLC-FT-ICR-MS and CAD MS/MS results for tryptic digested human platelet actin, depicting the MS obtained from the EIC of the target peptide, the CAD MS/MS spectrum, and the presence of the diagnostic $[b_5+Me]^{2+}$ fragment ion.

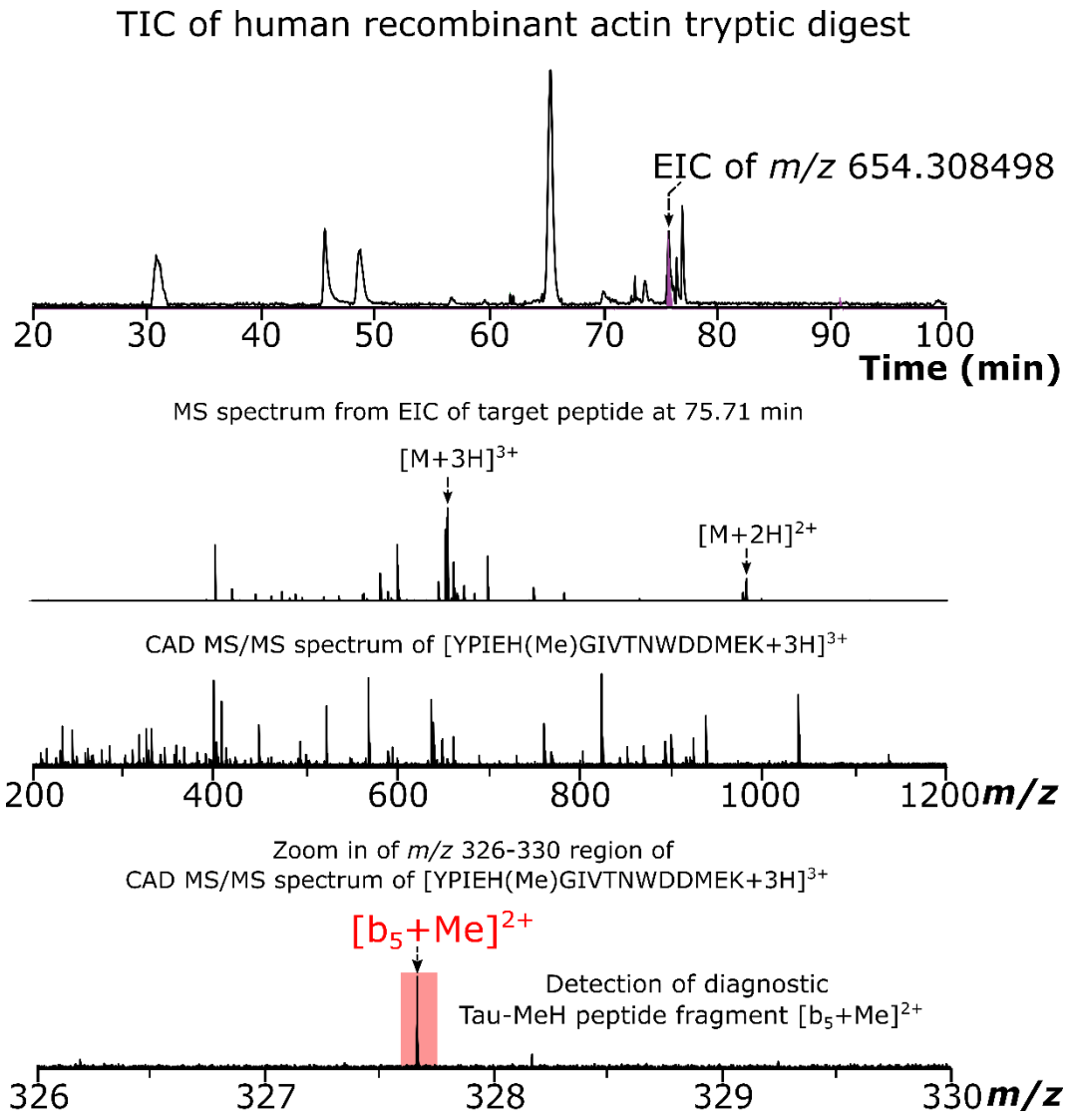


Figure 3. 15 The nLC-FT-ICR-MS and CAD MS/MS results for tryptic digested human recombinant actin, depicting the MS obtained from the EIC of the target peptide, the CAD MS/MS spectrum, and the presence of the diagnostic $[b_5+Me]^{2+}$ fragment ion.

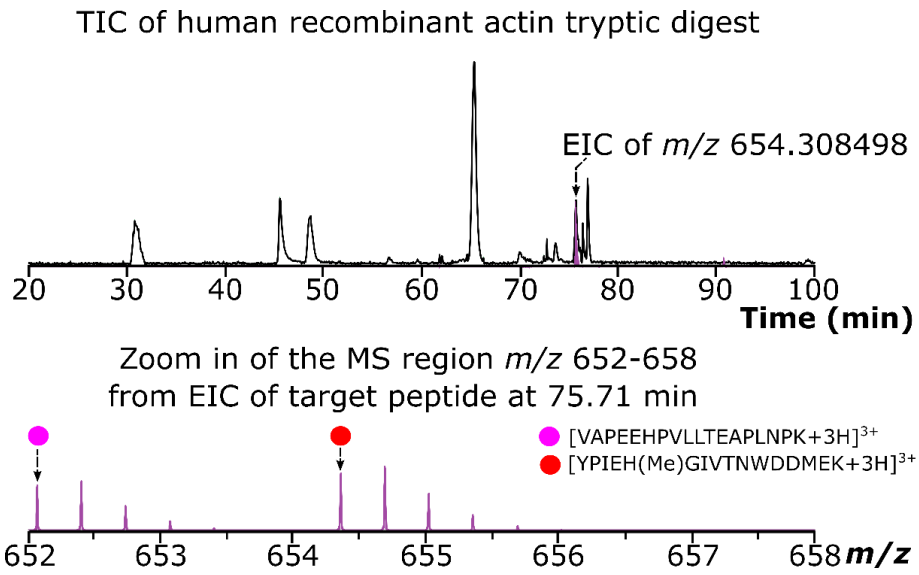


Figure 3. 16 The nLC-FT-ICR-MS results for human recombinant tryptic digest, depicting the MS obtained from the EIC of the target peptide, which highlights the triply charged precursor ion of the target peptide and a coeluting peptide in the *m/z* 652-658 region of the MS.

Due to the observations in Figure 3.16, which may affect the fragments observed in the MS/MS spectra, a direct infusion MS experiment was carried out on a mixture of tryptic digest of rabbit actin with the synthetic τ-MeH peptide as shown in Figure 3.17 below. As a result of the closeness in *m/z* of a triply charged peptide ion at *m/z* 652.0263173 from the rabbit actin tryptic digest to the target peptide at *m/z* 654.308498 from the τ-MeH synthetic peptide, a smaller isolation window of 2 *m/z* is required to prevent interference from the nearby peptide at *m/z* 652.0263173. Ultimately, as there was no retention time separation of the isomeric MeH peptides in the 50/50 synthetic peptide mixture or the digested human actin samples, peptide co-elution during the nano-LC experiment and the presence of other peptides aside from the target peptide during mass isolation, this may lead to under-estimation or over-estimation of the τ-MeH and π-MeH content in the actual sample mixture therefore direct infusion nano-ESI MS and MS/MS is preferred over nano-LC experiments.

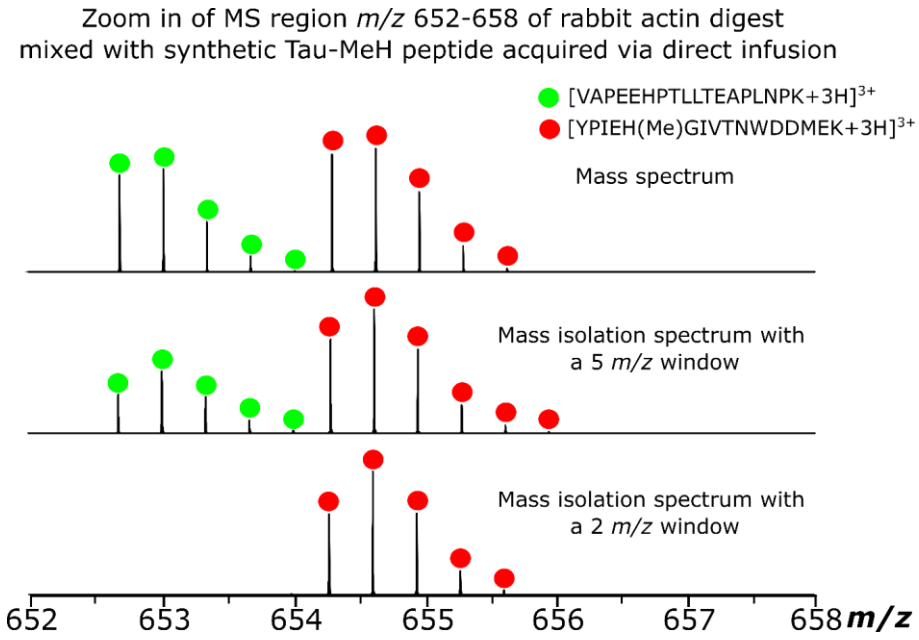


Figure 3. 17 Direct infusion nESI-FT-ICR-MS zoom in of *m/z* 652-658 region for rabbit actin tryptic digest mixed with synthetic tau-MeH peptide (top spectrum), with a 5 *m/z* mass isolation window applied (middle spectrum), and a 2 *m/z* mass isolation window applied (bottom spectrum).

Discussion of the formation of diagnostic fragments

The position of the τ-nitrogen and π-nitrogen on the imidazole ring of histidine may provide some valuable insight as to why the diagnostic fragments, particularly the $[b_5+Me]^{2+}$ fragment ion was only generated for the τ-MeH synthetic peptide. Both isomers contain strongly basic imidazole nitrogen groups with little difference between the methylated nitrogen pKa values; 9.25 and 9.43 for the τ-MeH isomer and the π-MeH isomer, respectively. Therefore, it may be of interest to consider other factors, which may influence the formation of the critical isomeric peptide differentiating fragments.

The mechanism for the formation of b fragment ions with CAD MS/MS resulting in a stable cyclic oxazolone structure is of interest due to the preferential fragmentation of the τ-MeH peptide at the methylated histidine residue in the target peptide sequence. The specific diagnostic $[b_5+Me]^{2+}$ fragment ion generated contains the methylhistidine modification site, whereas the τ-MeH peptide diagnostic $[y_{11}]^{+}$ fragment ion, resulting from cleavage at the glycine residue in the peptide sequence does not contain the modification site, hence the provided explanation is proposed for the $[b_5+Me]^{2+}$ fragment ion.

As demonstrated by Figure 3.18; two isomers are present, one isomer with the H on the imine nitrogen (τ-nitrogen), then it can share the proton with the oxazolone, whereas if the H is on the amine nitrogen (π-nitrogen), it cannot bind to the oxazolone nitrogen due to the conversion to a tetrahedral structure. The structural flexibility for normal backbone reorientation is hindered due to the saturation of the nitrogen with hydrogens in the oxazolone structure of Figure 3.19b). Therefore, the preferential formation of the $[b_5+Me]^{2+}$ fragment ion may be attributed to the steric hindrance caused by the location of the methylated π-nitrogen. This may be supported by the HDX results presented by Gucinski *et. al.*[67] who demonstrated that no HDX occurred at the π-methyl-substituted histidine side chain site because methylation of the π-nitrogen causes steric hindrance of any proton bridging between the imidazole ring of histidine and other heteroatoms.

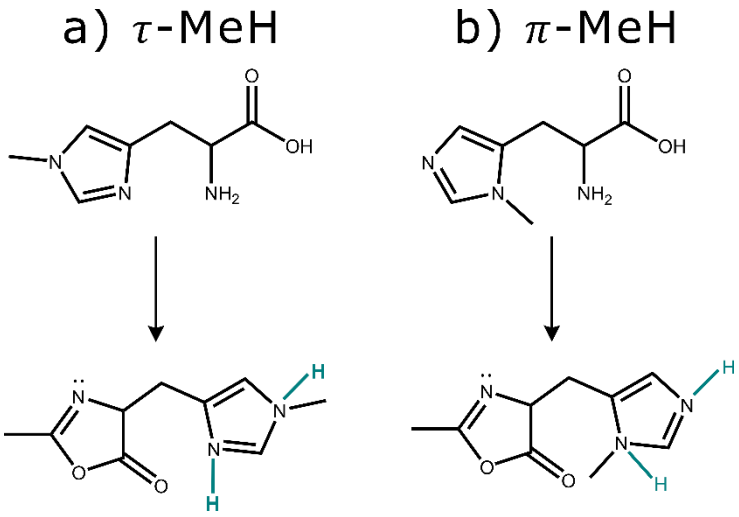


Figure 3. 18 Structures of a) τ-MeH and b) π-MeH and their respective oxazolone ring formations.

Computational methods such as ab initio and density functional theory calculations have also been studied and support the experimental findings of Gucinski *et. al.* regarding the effects and reactions of histidine and histidine containing peptides upon fragmentation. [68,69] However, the primary focus of this study was regarding the differentiation of the isomeric methylhistidine actin peptides and their relative quantification using MS/MS methods. The tentative mechanism presented herein is purely to aid understanding of the data presented and highlight the previous work allowing the current study to occur. The observations demonstrate the effects of histidine methylation as a cleavage-inhibiting PTM, resulting in the absence of the diagnostic

$[b_5+Me]^{2+}$ fragment ion in the CAD, IRMPD, and UVPD MS/MS spectra of the π-MeH synthetic peptide and the presence of this fragment in the τ-MeH synthetic peptide MS/MS spectra.

τ-MeH and π-MeH peptide quantification from different actin digest samples using CAD MS/MS

From the above sections, CAD MS/MS can generate characteristic fragments ($[b_5+Me]^{2+}$ and $[y_{11}]^+$) which can be potentially applied in relative quantification of τ-MeH and π-MeH peptides. Previous methods applied for relative isomeric peptide quantification using diagnostic fragments generated via MS/MS techniques were discussed in length in Chapter 2 in relation to isoD and D peptide quantification using diagnostic fragments generated via ECD MS/MS, however, the same principles and developed methods apply herein.

The synthetic isomeric τ-MeH and π-MeH peptides were mixed to obtain mixtures, in which the τ-MeH peptide percentage content varied from 0 % to 100 % in 20 % increments, with the inclusion of the 50% of τ-MeH and 50 % π-MeH peptide mixture. Figure 3.19 exhibits a direct proportional change in the intensity of the diagnostic fragments at *m/z* 327.665932 ($[b_5+Me]^{2+}$) and *m/z* 1307.593629 ($[y_{11}]^+$) generated by CAD MS/MS to the percentage of the τ-MeH peptide concentration in the isomeric peptide mixtures.

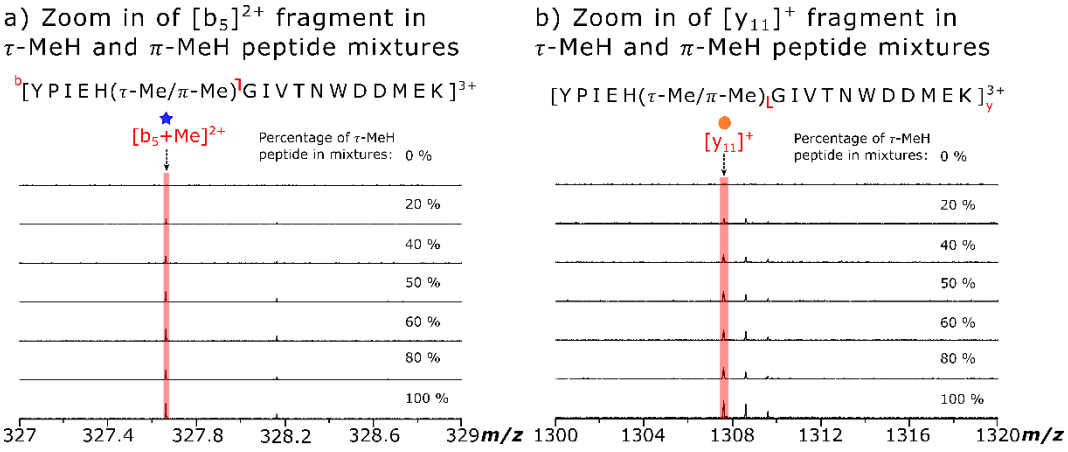


Figure 3. 19 Zoom in of characteristic fragments at *m/z* 327.6 ($[b_5+Me]^{2+}$) and *m/z* 1307 ($[y_{11}]^+$) in CAD MS/MS spectra of τ-MeH and π-MeH synthetic peptide mixtures (scaled to the same intensity at y-axis).

A ratio was taken from the peak area of the diagnostic $[b_5+Me]^{2+}$ or the $[y_{11}]^+$ fragment ion to the sum of all the common fragments present in the CAD MS/MS spectra for each isomeric τ-MeH and π-MeH peptide mixture. The relative peak area ratio was then plotted against the percentage content of the τ-MeH and π-MeH synthetic peptide mixtures. The plotted data demonstrates a linear trend between the peak area ratio of the diagnostic $[b_5+Me]^{2+}$ or the $[y_{11}]^+$ fragment ion and the percentage content/concentration of the τ-MeH peptide in the τ-MeH and π-MeH peptide mixtures as shown by Figure 3.20. Calibration curves were obtained with a good linearity (R^2>0.99) for both the τ-MeH peptide diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^+$ fragment ions.

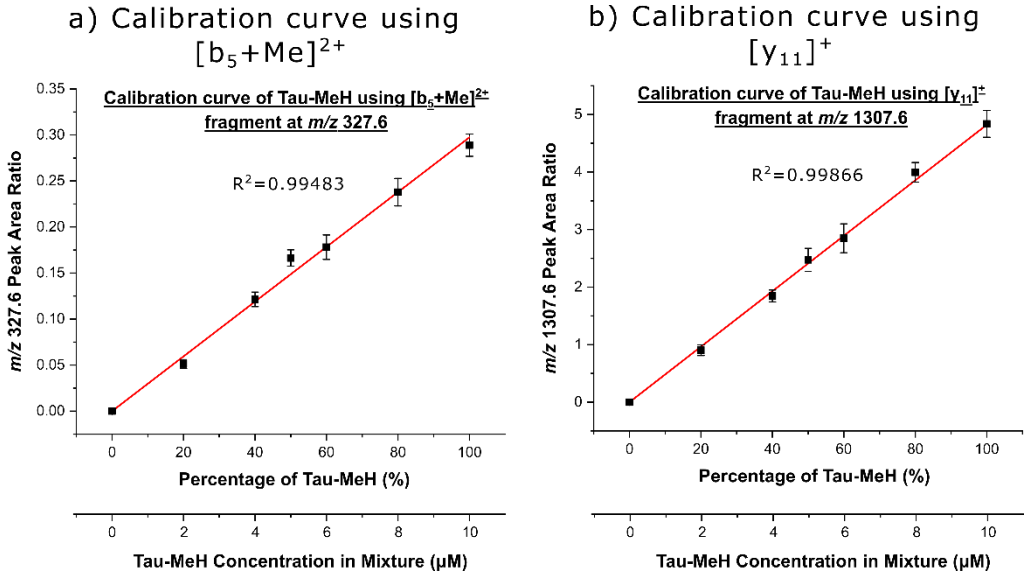


Figure 3. 20 Calibration curves composed of τ-MeH and π-MeH synthetic peptide mixtures using the a) $[b_5+Me]^{2+}$ at *m/z* 327.6 and b) $[y_{11}]^+$ at *m/z* 1307.6 fragments for quantification.

The methylated tryptic peptide from bovine, chicken, rabbit, human platelet, and human recombinant actin was fragmented using CAD MS/MS. Figure 3.21a) represents the CAD MS/MS spectra of the isolated methylated peptide from the different actin samples. When subjected to CAD MS/MS analysis, like the synthetic isomeric MeH peptides, 100 % cleavage coverage was also observed for the target peptide of the mammalian actin digest samples. Inserts of the zoomed in MS/MS region, where the diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^+$ fragment ions are present in the spectra are provided in Figure 3.21b). Thus, the percentage content of the of τ-MeH peptide in the tryptic digested actin samples can be determined using the detected diagnostic fragments and the highlighted common fragments in the CAD MS/MS spectra in Figure 3.21 a).

For clarification, the target peptide sequence differs by change of one amino acid (V76I) in the bovine, chicken, and rabbit actin compared to the synthetic peptides, human platelet actin, and human recombinant actin. This is demonstrated by the triply charged protonated molecular ion *m/z* shift to higher mass (*m/z* 658.980383) in the mass spectra for the animal actin samples (Supplementary Figure S3.9). The difference in the peptide sequences is also noted by the peptide cleavage diagrams in Figure 3.21a) and the apparent mass shift to higher *m/z* of the $[y_{11}]^+$ fragment ion in Figure 3.21b) generated from the methylated tryptic peptide of bovine, chicken, and rabbit actin.

To minimise the underestimation or overestimation of the τ-MeH and π-MeH content in the target peptide due to the minor difference in the peptide sequence between the animal actin and human actin samples, only the "common" fragments observed for both types of actin samples were used for the relative quantification of the τ-MeH and π-MeH peptides. The common fragments consist of the fragments present in the CAD MS/MS spectra of the target methylated tryptic peptide observed in all of actin samples studied, which are unaffected by or do not include the fragments generated via cleavage at V76I, where there is the amino acid change in the target peptide sequence.

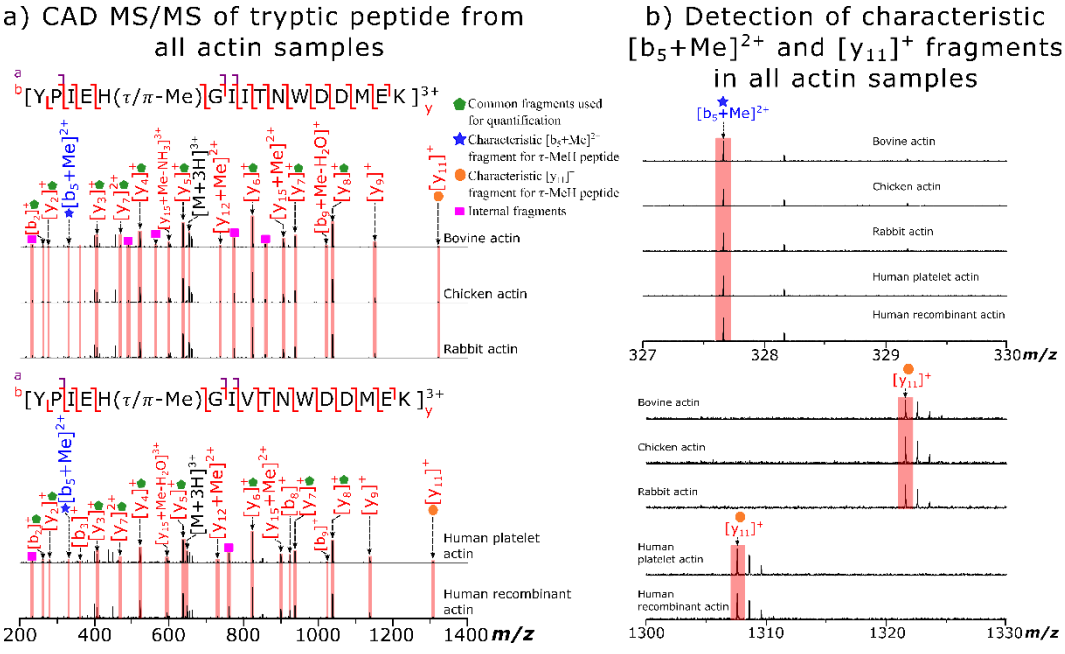


Figure 3. 21 CAD MS/MS spectra of the target methylated tryptic peptide from different actin samples and b) zoom in of the *m/z* region where the diagnostic $[b_5+Me]^{2+}$) and $[y_{11}]^+$ fragment ions are present for the tryptic peptide from the different actin digest samples (scaled to the same intensity at the y-axis).

Following the determination of the peak area ratios of the diagnostic $[b_5+Me]^{2+}$ or the $[y_{11}]^{+}$ fragment ion to the sum of the common fragments, the equations of the linear calibration curves generated (Figure 3.20) were used to determine the percentage content of τ-MeH in the different actin samples. The relative quantification results of τ-MeH in the target peptide from bovine, chicken, rabbit, human platelet, and human recombinant actin are shown in Figure 3.22.

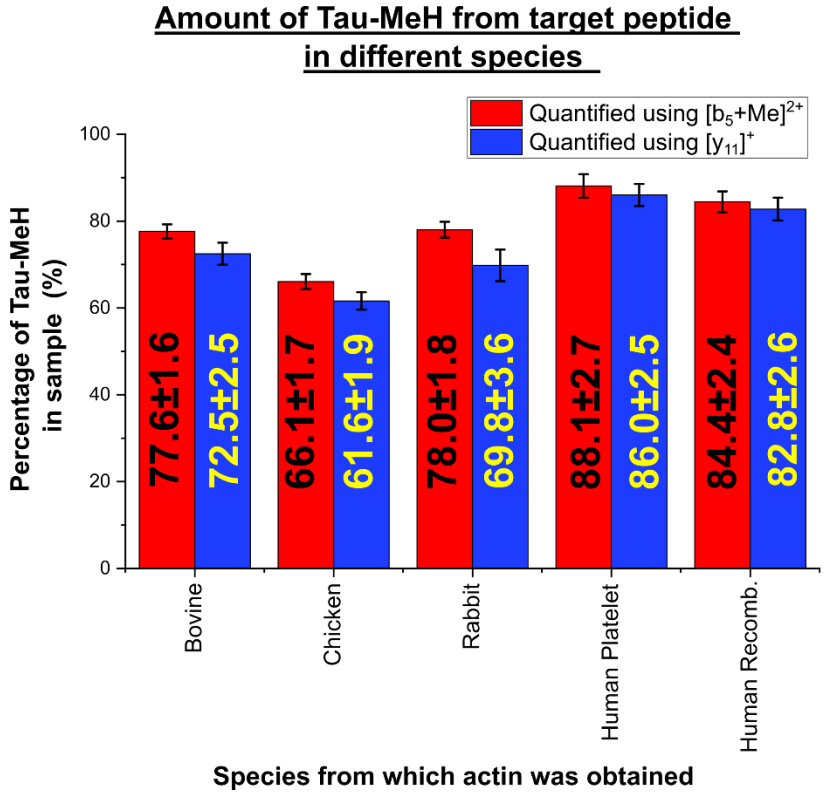


Figure 3. 22 Relative τ-MeH peptide quantification results for different actin samples.

Multiple observations can be made based on the relative quantification results displayed in Figure 3.22. Firstly, the τ-MeH percentage content in the target peptide of the human platelet actin and human recombinant actin was higher (> 80 %) compared to the τ-MeH content of the target peptide derived from the animal actin samples (< 80 %).

Secondly, differences in the τ-MeH percentage content for the animal actin samples are observed when quantification is carried out using the $[b_5+Me]^{2+}$ fragment compared to the $[y_{11}]^{+}$ fragment ion. For example, a higher percentage of τ-MeH in the target peptide is detected for the animal actin samples when quantifying using the $[b_5+Me]^{2+}$ fragment when including the amount of variation of the results. However, if the standard deviation of the calculated τ-MeH percentage content in the target peptide of

the human actin samples is included, little to no difference is observed between quantification using the $[b_5+Me]^{2+}$ fragment or the $[y_{11}]^{+}$ fragment ion.

Thirdly, it is difficult to directly compare the τ-MeH and π-MeH content values obtained in this work to the values obtained by Johnson *et. al.*[16] as the percentages mentioned Figure 3.22 are representative of the τ-MeH and pi-MeH content in the affected target peptide rather than the individual τ-MeH amino acid, which was quantified using electrophoresis in the work carried out by Johnson and co-workers. However, a significant trend is observed within the results that follow the ratio of τ-MeH to histidine in the order of the lowest ratio of τ-MeH to histidtine determined in the chicken actin (1:7.46), the rabbit actin (1:7.6), and the highest ratio of τ-MeH to histidine was observed in the human actin (1:8.62), which agree with the relative quantification results presented in this work.

3.5. Conclusions

The experimental results demonstrate the differentiation of the isomeric τ-MeH and π-MeH actin peptides via the generation of diagnostic MS/MS fragments. The $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions were observed only for the synthetic τ-MeH peptide in the CAD, IRMPD, and UVPD MS/MS spectra but were absent in the π-MeH peptide MS/MS spectra. Although the $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions are reliably produced for the τ-MeH peptide via the different MS/MS techniques, this means that the presence of the π-MeH peptide must be determined via the lack of the τ-MeH peptide diagnostic fragment ions as no fragments unique to the π-MeH peptide were detected.

However, based on the detection of the specific τ-MeH peptide fragments, a relative quantification method was successfully developed using CAD MS/MS. A linear trend was observed between the relative intensity of the diagnostic fragment ions and the τ-MeH content in the synthetic τ-MeH and π-MeH peptide mixtures. A ratio was taken of each diagnostic fragment peak area to the sum of the common fragments in the MS/MS spectra, which were plotted against the τ-MeH percentage content in the synthetic peptide mixtures. A good linearity ($R^2 > 0.99$) was achieved for the calibration curves using the $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions.

The equations of the generated calibration curves were then applied to quantify the τ-MeH and π-MeH content in the target peptide of 5 types of actin samples obtained from different species including bovine, chicken, rabbit, and human actin. The relative τ-MeH quantification results show that the τ-MeH form is the dominant isomeric form in all mammalian actin samples studied herein, with the highest percentage of the τ-MeH content in the target peptide detected in the human actin samples (> 80 %). The results also demonstrate the strong potential for direct relative quantification of isomeric species using fragmentation techniques, which can be applicable to complex and biologically significant samples.

3.6. References

(1) Lee, D. Y.; Teyssier, C.; Strahl, B. D.; Stallcup, M. R. Role of Protein Methylation in Regulation of Transcription. Endocr. Rev. 2005, 26 (2), 147-170.

(2) Alam, H.; Gu, B.; Lee, M. G. Histone methylation modifiers in cellular signaling pathways. Cell. Mol. Life Sci. 2015, 72 (23), 4577-4592.

(3) Wei, W.; Ji, X.; Guo, X.; Ji, S. Regulatory Role of N6-methyladenosine (m6A) Methylation in RNA Processing and Human Diseases. J. Cell. Biochem. 2017, 118 (9), 2534-2543.

(4) Clarke, S. Aging as war between chemical and biochemical processes: Protein methylation and the recognition of age-damaged proteins for repair. Ageing Res. Rev. 2003, 2 (3), 263-285.

(5) Shimizu, T.; Matsuoka, Y.; Shirasawa, T. Biological Significance of Isoaspartate and Its Repair System. Biol. Pharm. Bull. 2005, 28 (9), 1590-1596.

(6) Bedford, M. T.; Frankel, A.; Yaffe, M. B.; Clarke, S.; Leder, P.; Richard, S. Arginine Methylation Inhibits the Binding of Proline-rich Ligands to Src Homology 3, but Not WW, Domains *. J. Biol. Chem. 2000, 275 (21), 16030-16036.

(7) Hardy, M. F.; Perry, S. V. In vitro Methylation of Muscle Proteins. Nature 1969, 223 (5203), 300-302.

(8) Moore, L. D.; Le, T.; Fan, G. DNA Methylation and Its Basic Function. Neuropsychopharmacology 2013, 38 (1), 23-38.

(9) Greer, E. L.; Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. Nat. Rev. Genet. 2012, 13 (5), 343-357.

(10) Han, D.; Huang, M.; Wang, T.; Li, Z.; Chen, Y.; Liu, C.; Lei, Z.; Chu, X. Lysine methylation of transcription factors in cancer. Cell Death Dis. 2019, 10 (4), 290-290.

(11) Rowe, E. M.; Xing, V.; Biggar, K. K. Lysine methylation: Implications in neurodegenerative disease. Brain Res. 2019, 1707, 164-171.

(12) Paulsen, M.; Ferguson-Smith, A. C. DNA methylation in genomic imprinting, development, and disease. J. Pathol. 2001, 195 (1), 97-110.

(13) Kim, M.; Long, T. I.; Arakawa, K.; Wang, R.; Yu, M. C.; Laird, P. W. DNA Methylation as a Biomarker for Cardiovascular Disease Risk. PLoS One 2010, 5 (3), e9692.

(14) Clarke, S. Protein methylation. Curr. Opin. Cell Biol. 1993, 5 (6), 977-983.

(15) Cantoni, G. L. The Nature of the Active Methyl Donor Formed Enzymatically from L-Methionine and Adenosinetriphosphate. J. Am. Chem. Soc. 1952, 74 (11), 2942-2943.

(16) Johnson, P.; Harris, C. I.; Perry, S. V. 3-methylhistidine in actin and other muscle proteins. Biochem. J. 1967, 105 (1), 361-370.

(17) Zhang, L.; Ding, X.; Cui, J.; Xu, H.; Chen, J.; Gong, Y.-N.; Hu, L.; Zhou, Y.; Ge, J.; Lu, Q.; Liu, L.; Chen, S.; Shao, F. Cysteine methylation disrupts ubiquitin-chain sensing in NF-κB activation. Nature 2012, 481 (7380), 204-208.

(18) Atik, A. E.; Guray, M. Z.; Yalcin, T. Observation of the side chain O-methylation of glutamic acid or aspartic acid containing model peptides by electrospray ionization-mass spectrometry. J. Chromatogr. B 2017, 1047, 75-83.

(19) Huszar, G.; Elzinga, M. Amino acid sequence around the single 3-methylhistidine residue in rabbit skeletal muscle myosin. Biochemistry 1971, 10 (2), 229-236.

(20) Huszar, G.; Elzinga, M. Homologous Methylated and Nonmethylated Histidine Peptides in Skeletal and Cardiac Myosins. J. Biol. Chem. 1972, 247 (3), 745-753.

(21) Huxley, A. F.; Niedergerke, R. Structural Changes in Muscle During Contraction: Interference Microscopy of Living Muscle Fibres. Nature 1954, 173 (4412), 971-973.

(22) Huxley, H.; Hanson, J. Changes in the Cross-Striations of Muscle during Contraction and Stretch and their Structural Interpretation. Nature 1954, 173 (4412), 973-976.

(23) Clarke, M.; Spudich, J. A. Nonmuscle Contractile Proteins: The Role of Actin and Myosin in Cell Motility and Shape Determination. Annu. Rev. Biochem 1977, 46 (1), 797-822.

(24) Durham, A. C. H. A unified theory of the control of actin and myosin in nonmuscle movements. Cell 1974, 2 (3), 123-135.

(25) Sanger, J. M.; Mittal, B.; Dome, J. S.; Sanger, J. W. Analysis of cell division using fluorescently labeled actin and myosin in living PtK2 cells. Cell Motil. 1989, 14 (2), 201-219.

(26) Wu, S.-Z.; Bezanilla, M. Myosin VIII associates with microtubule ends and together with actin plays a role in guiding plant cell division. eLife 2014, 3, e03498.

(27) Tang, D. D.; Gerlach, B. D. The roles and regulation of the actin cytoskeleton, intermediate filaments and microtubules in smooth muscle cell migration. Respir. Res. 2017, 18 (1), 54.

(28) Searle, J.; Westall, R. The occurrence of free methylhistidine in urine. Biochem. J. 1951, 48 (4), 1.

(29) Tallan, H. H.; Stein, W. H.; Moore, S. 3-Methylhistidine, a new amino acid from human urine. J. Biol. Chem. 1954, 206 (2), 825-834.

(30) Ballard, F. J.; Tomas, F. M. 3-Methylhistidine as a Measure of Skeletal Muscle Protein Breakdown in Human Subjects: The Case for Its Continued Use. Clin. Sci. 1983, 65 (3), 209-215.

(31) Young, V. R.; Munro, H. N. Ntau-methylhistidine (3-methylhistidine) and muscle protein turnover: an overview. Fed. Proc. 1978, 37 (9), 2291-2300.

(32) Butt, J. H.; Fleshler, B. Anserine, A Source of 1-Methylhistidine in Urine of Man. Proc. Soc. Exp. Biol. Med. 1965, 118 (3), 722-725.

(33) Myint, T.; Fraser, G. E.; Lindsted, K. D.; Knutsen, S. F.; Hubbard, R. W.; Bennett, H. W. Urinary 1-Methylhistidine Is a Marker of Meat Consumption in Black and in White California Seventh-day Adventists. Am. J. Epidemiol. 2000, 152 (8), 752-755.

(34) Asatoor, A. M.; Armstrong, M. D. 3-Methylhistidine, a component of actin. Biochem. Biophys. Res. Commun. 1967, 26 (2), 168-174.

(35) Trayer, I. P.; Harris, C. I.; Perry, S. V. 3-Methyl Histidine and Adult and Foetal Forms of Skeletal Muscle Myosin. Nature 1968, 217 (5127), 452-453.

(36) Liebecq, C. Biochemical nomenclature and related documents, ed.; Portland Press, 1992; Vol., p.

(37) Yao, X.; Grade, S.; Wriggers, W.; Rubenstein, P. A. His(73), often methylated, is an important structural determinant for actin. A mutagenic analysis of HIS(73) of yeast actin. J. Biol. Chem. 1999, 274 (52), 37443-37449.

(38) Nyman, T.; Schüler, H.; Korenbaum, E.; Schutt, C. E.; Karlsson, R.; Lindberg, U. The role of MeH73 in actin polymerization and ATP hydrolysis11Edited by R. Huber. J. Mol. Biol. 2002, 317 (4), 577-589.

(39) Guo, Q.; Liao, S.; Kwiatkowski, S.; Tomaka, W.; Yu, H.; Wu, G.; Tu, X.; Min, J.; Drozak, J.; Xu, C. Structural insights into SETD3-mediated histidine methylation on β-actin. eLife 2019, 8, e43676.

(40) Kwiatkowski, S.; Seliga, A. K.; Vertommen, D.; Terreri, M.; Ishikawa, T.; Grabowska, I.; Tiebe, M.; Teleman, A. A.; Jagielski, A. K.; Veiga-da-Cunha, M.; Drozak, J. SETD3 protein is the actin-specific histidine N-methyltransferase. eLife 2018, 7, e37921.

(41) Shu, W.-J.; Du, H.-N. The methyltransferase SETD3-mediated histidine methylation: Biological functions and potential implications in cancers. Biochim. Biophys. Acta 2021, 1875 (1), 188465.

(42) Wilkinson, A. W.; Diep, J.; Dai, S.; Liu, S.; Ooi, Y. S.; Song, D.; Li, T. M.; Horton, J. R.; Zhang, X.; Liu, C.; Trivedi, D. V.; Ruppel, K. M.; Vilches-Moure, J. G.; Casey, K. M.; Mak, J.; Cowan, T.; Elias, J. E.; Nagamine, C. M.; Spudich, J. A.; Cheng, X., *et. al.* SETD3 is an actin histidine methyltransferase that prevents primary dystocia. Nature 2019, 565 (7739), 372-376.

(43) Zheng, Y.; Zhang, X.; Li, H. Molecular basis for histidine N3-specific methylation of actin H73 by SETD3. Cell Discov. 2020, 6 (1), 3.

(44) Rathmacher, J. A.; Link, G. A.; Flakoll, P. J.; Nissen, S. L. Gas chromatographic/mass spectrometric analysis of stable isotopes of 3-methylhistidine in biological fluids: Application to plasma kinetics in vivo. Biol. Mass Spectrom. 1992, 21 (11), 560-566.

(45) Thompson, M. G.; Palmer, R. M.; Thom, A.; Garden, K.; Lobley, G. E.; Calder, G. N tau-methylhistidine turnover in skeletal muscle cells measured by GC-MS. Am. J. Physiol. Cell Physiol. 1996, 270 (6), C1875-C1879.

(46) Ali Qureshi, G.; Van den berg, S.; Gutierrez, A.; Bergström, J. Determination of histidine and 3-methylhistidine in physiological fluids by high-performance liquid chromatography. J. Chromatogr. A 1984, 297, 83-89.

(47) Houweling, M.; van der Drift, S. G. A.; Jorritsma, R.; Tielens, A. G. M. Quantification of plasma 1- and 3-methylhistidine in dairy cows by high-performance liquid chromatography- tandem mass spectrometry. J. Dairy Sci. 2012, 95 (6), 3125-3130.

(48) Libera, L. D. Determination of 3-methylhistidine in hydrolysed proteins by fluorescamine derivatization and high-performance liquid chromatography. J. Chromatogr. A 1991, 536, 283-288.

(49) Halket, J. M.; Waterman, D.; Przyborowska, A. M.; Patel, R. K. P.; Fraser, P. D.; Bramley, P. M. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J. Exp. Bot. 2005, 56 (410), 219-243.

(50) Dančík, V.; Addona, T. A.; Clauser, K. R.; Vath, J. E.; Pevzner, P. A. De Novo Peptide Sequencing via Tandem Mass Spectrometry. J. Comput. Biol. 1999, 6 (3-4), 327-342.

(51) Hunt, D. F.; Buko, A. M.; Ballard, J. M.; Shabanowitz, J.; Giordani, A. B. Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. Biomed. Mass Spectrom. 1981, 8 (9), 397-408.

(52) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein sequencing by tandem mass spectrometry. Proc. Natl. Acad. Sci. U.S.A. 1986, 83 (17), 6233.

(53) Li, H.; Liu, C.; Burge, L.; Southerland, W. Identification of two post-translational modifications via tandem mass spectrometry. Int. J. Comput. Biol. Drug Des. 2012, 5 (3-4), 314-324.

(54) Qi, Y.; Volmer, D. A. Electron-based fragmentation methods in mass spectrometry: An overview. Mass Spectrom. Rev. 2017, 36 (1), 4-15.

(55) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. Chem. Rev. 2020, 120 (7), 3328-3380.

(56) Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Anal. Chem. 1987, 59 (21), 2621-2625.

(57) Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. Protein Sci. 2005, 14 (2), 452-463.

(58) O'Connor, P. B.; Cournoyer, J. J.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. Differentiation of Aspartic and Isoaspartic Acids Using Electron Transfer Dissociation. J. Am. Soc. Mass Spectrom. 2006, 17 (1), 15-19.

(59) Hatano, T.; Sivashanmugam, L.; Suchenko, A.; Hussain, H.; Balasubramanian, M. K. Pick-ya actin – a method to purify actin isoforms with bespoke key post-translational modifications. J. Cell Sci. 2020, 133 (2), jcs241406.

(60) Walker, V.; Mills, G. A. Quantitative Methods for Amino Acid Analysis in Biological Fluids. Ann. Clin. Biochem. 1995, 32 (1), 28-57.

(61) Min, S.; Yisheng, Y.; Lu, Y. Determination of urinary 3-methylhistidine by high-performance liquid chromatography with o-phthaldialdehyde precolumn derivatization. J. Chromatogr. B 1992, 581 (2), 272-276.

(62) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G., *et. al.* HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007, 35 (Database issue), D521-D526.

(63) Mathur, R.; O'Connor, P. B. Artifacts in Fourier transform mass spectrometry. Rapid Commun. Mass Spectrom. 2009, 23 (4), 523-529.

(64) Johnson, P.; Lobley, G. E.; Perry, S. V. Distribution and biological role of 3-methyl-histidine in actin and myosin. Biochem. J. 1969, 114 (2), 34P-34P.

(65) Johnson, P.; Perry, S. V. Biological activity and the 3-methylhistidine content of actin and myosin. Biochem. J. 1970, 119 (2), 293-298.

(66) Marx Jean, L. Actin and Myosin: Role in Nonmuscle Cells. Science 1975, 189 (4196), 34-37.

(67) Gucinski, A. C.; Chamot-Rooke, J.; Nicol, E.; Somogyi, Á.; Wysocki, V. H. Structural Influences on Preferential Oxazolone versus Diketopiperazine b2+ Ion Formation for Histidine Analogue-Containing Peptides. J. Phys. Chem. A 2012, 116 (17), 4296-4304.

(68) Tureček, F.; Chung, T. W.; Moss, C. L.; Wyer, J. A.; Ehlerding, A.; Holm, A. I. S.; Zettergren, H.; Nielsen, S. B.; Hvelplund, P.; Chamot-Rooke, J.; Bythell, B.; Paizs, B. The Histidine Effect. Electron Transfer and Capture Cause Different Dissociations and Rearrangements of Histidine Peptide Cation-Radicals. J. Am. Chem. Soc. 2010, 132 (31), 10728-10740.

(69) Nelson, C. R.; Abutokaikah, M. T.; Harrison, A. G.; Bythell, B. J. Proton Mobility in b2 Ion Formation and Fragmentation Reactions of Histidine-Containing Peptides. J. Am. Soc. Mass Spectrom. 2016, 27 (3), 487-497.

3.7 Supplementary Information

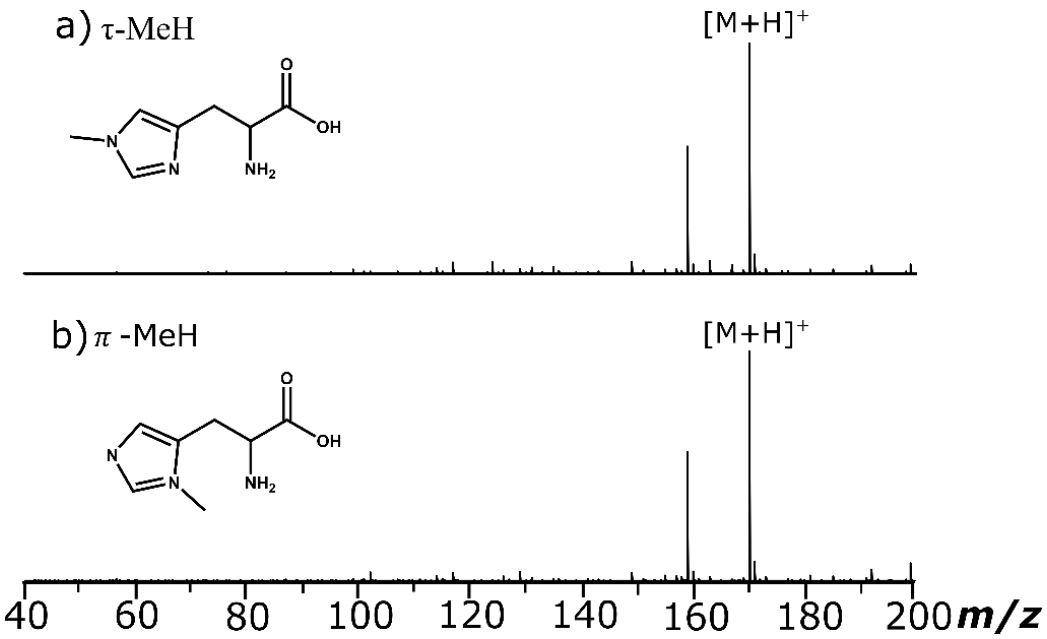


Figure S3. 1 Mass spectra of a) tele-methylhistidine (τ-MeH) and b) pros-methylhistidine (π-MeH) acquired via direct infusion on the 12 T FTICR-MS.

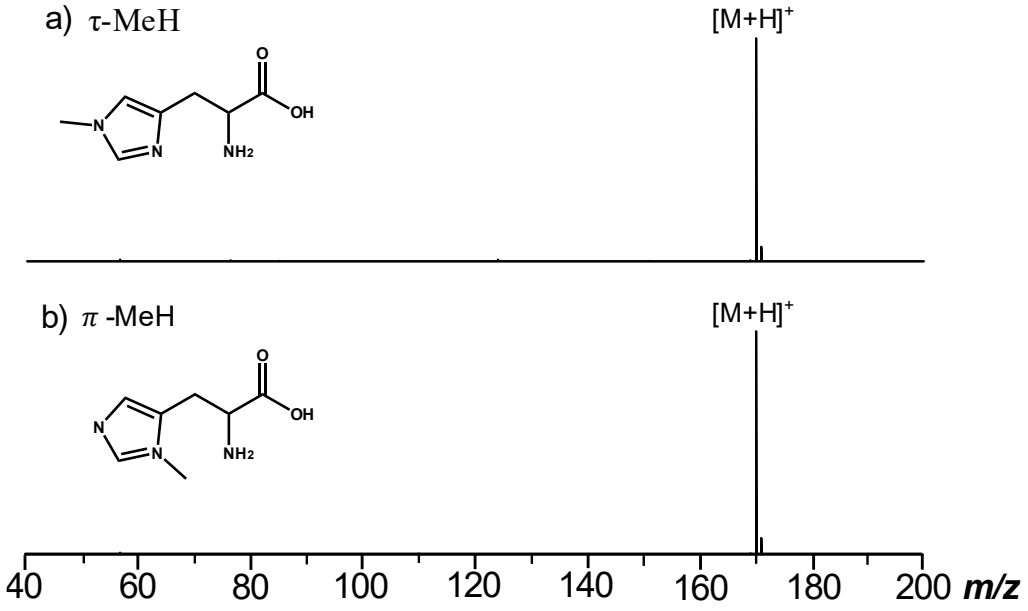


Figure S3. 2 Mass isolation spectra of a) tele-methylhistidine (τ-MeH) and b) pros-methylhistidine (π-MeH) amino acids.

Table S3. 1 Peak assignment table for the EID MS/MS of the protonated τ-MeH.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
A	$C_4H_5N_2$	2293231	537992	81.045273	81.045272	-0.01
	$C_4H_7N_2$	2731547	602952	83.060923	83.060924	0.01
	$C_5H_8N_2$	11096744	478647	96.068748	96.068788	0.42
C	$C_6H_{10}N_3$	13953325	383782	124.087472	124.087472	0.00
	$C_7H_{11}N_3O_2$	1142361856	275604	170.092403	170.092403	0.00
Average error						0.08
Absolute average error						0.09
Standard deviation						0.18

Table S3. 2 Peak assignment table for the EID MS/MS of the protonated π-MeH.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
A	$C_4H_5N_2$	2343732	552346	81.045273	81.045273	0.00
B	$C_5H_7N_2$	14158620	495849	95.060923	95.060921	-0.02
	$C_5H_8N_2$	25775054	472434	96.068748	96.068744	-0.04
	$C_5H_9N_2$	3339416	485356	97.076573	97.076575	0.02
	$C_6H_9N_2$	11569543	392502	109.076573	109.076529	-0.40
	$C_5H_8N_3$	2579067	497596	110.071822	110.071778	-0.40
C	$C_6H_{10}N_3$	4480046	418919	124.087472	124.087356	-0.93
	$C_6H_{12}N_3$	2565116	417394	126.103122	126.102999	-0.98
	$C_7H_{11}N_3O_2$	2409047808	275886	170.092403	170.092403	0.00
Average error						-0.31
Absolute average error						0.31
Standard deviation						0.40

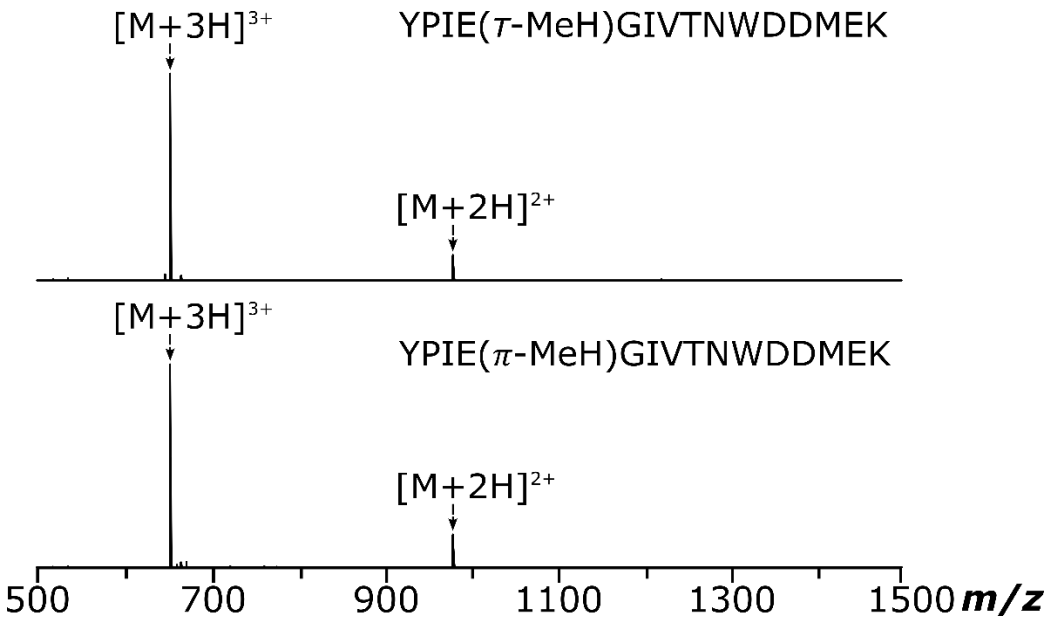


Figure S3. 3 Mass spectra of a) tele-methylhistidine (τ-MeH) and b) pros-methylhistidine (π-MeH) synthetic actin peptides.

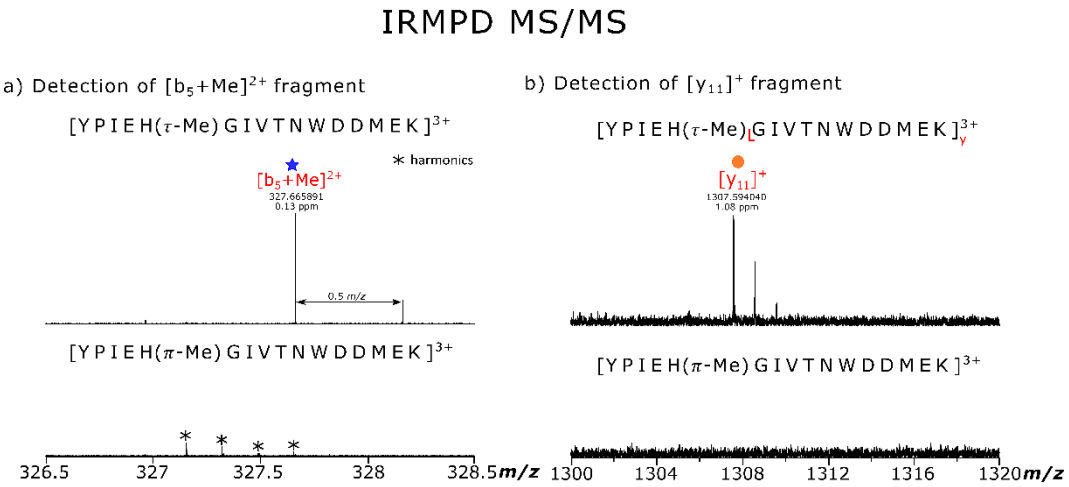


Figure S3. 4 Zoom in of the characteristic fragments at *m/z* 327.6 (b_5^{2+}) and *m/z* 1307 (y_{11}^{1+}) present in the IRMPD MS/MS spectra of the τ-MeH synthetic peptide, which are absent for the π-MeH peptide.

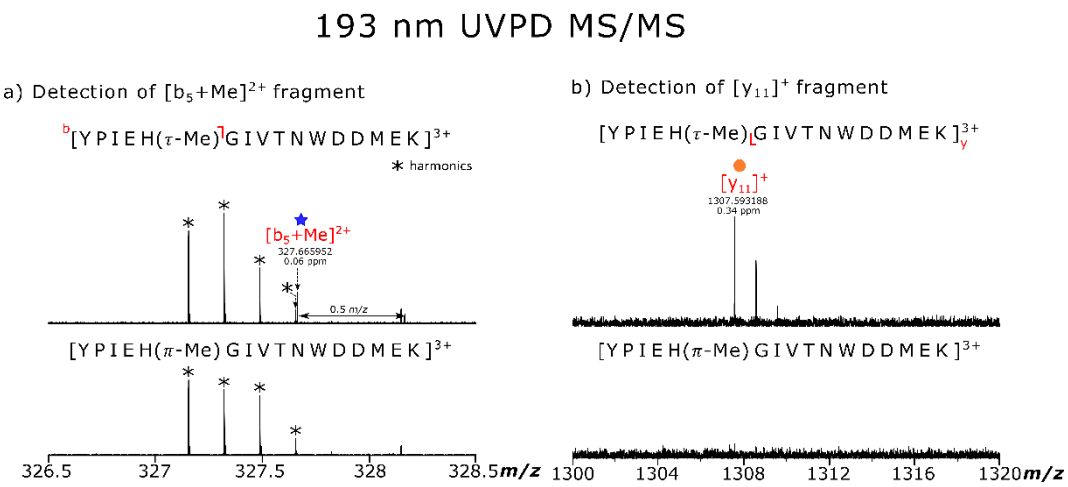


Figure S3. 5 Zoom in of the characteristic fragments at *m/z* 327.6 (b_5^{2+}) and *m/z* 1307 (y_{11}^{1+}) present in the 193 nm UVPD MS/MS spectra of the τ-MeH synthetic peptide, which are absent for the π-MeH peptide.

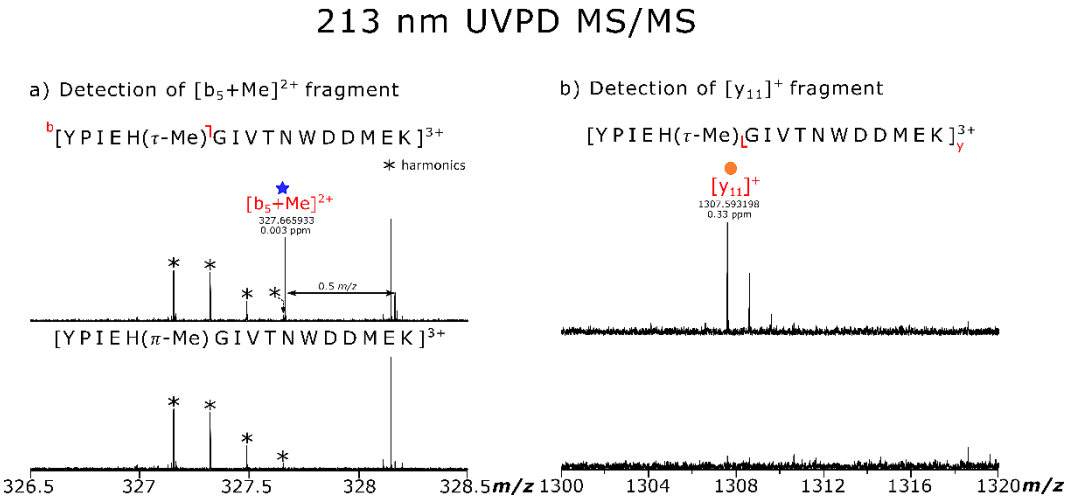


Figure S3. 6 Zoom in of the characteristic fragments at *m/z* 327.6 (b_5^{2+}) and *m/z* 1307 (y_{11}^{1+}) present in the 213 nm UVPD MS/MS spectra of the τ-MeH synthetic peptide, which are absent for the π-MeH peptide.

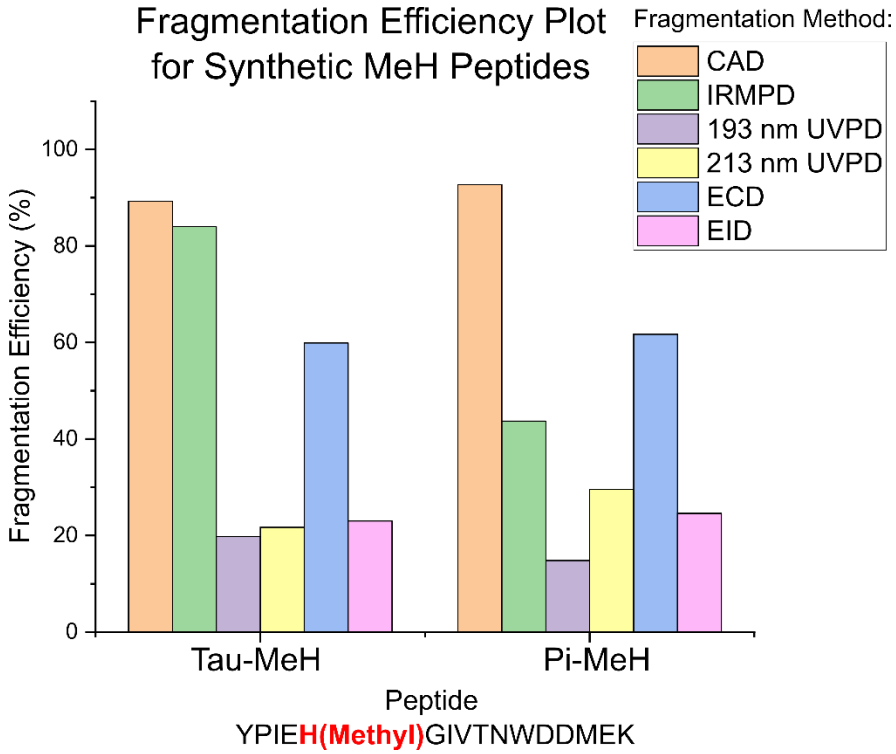


Figure S3. 7 Fragmentation efficiency plot for τ-MeH and π-MeH modified synthetic actin peptides using different fragmentation methods.

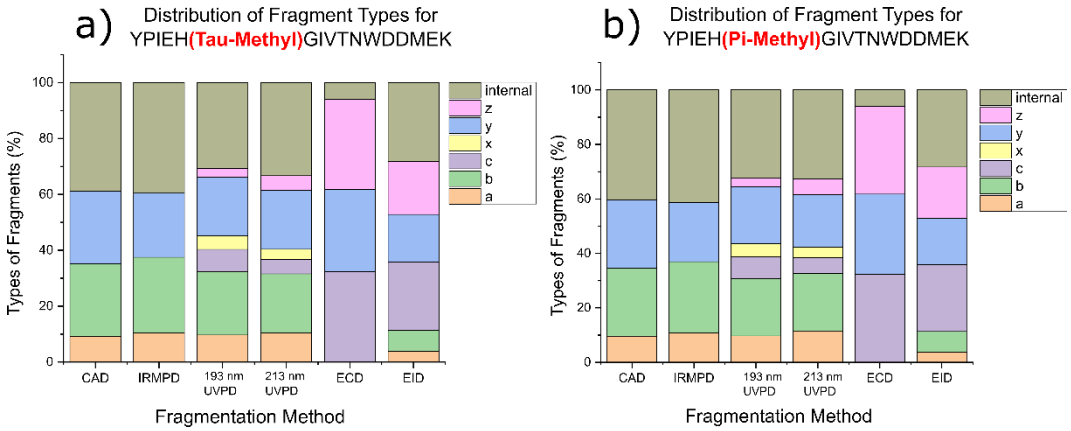


Figure S3. 8 Distribution of fragment types observed with different fragmentation methods for the synthetic actin peptides with the modification at a) τ-MeH and b) π-MeH in the peptide sequence.

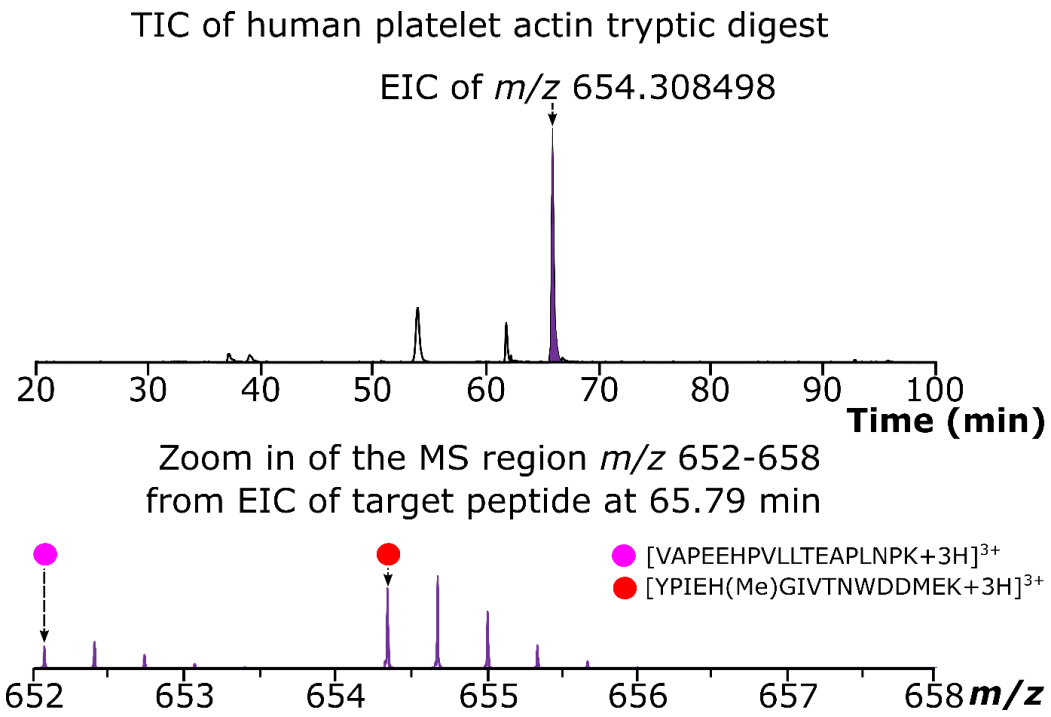


Figure S3. 9 The nLC-FT-ICR-MS results for human platelet tryptic digest, depicting the MS obtained from the EIC of the target peptide, which highlights the triply charged precursor ion of the target peptide and a coeluting peptide in the *m/z* 652-658 region of the MS.

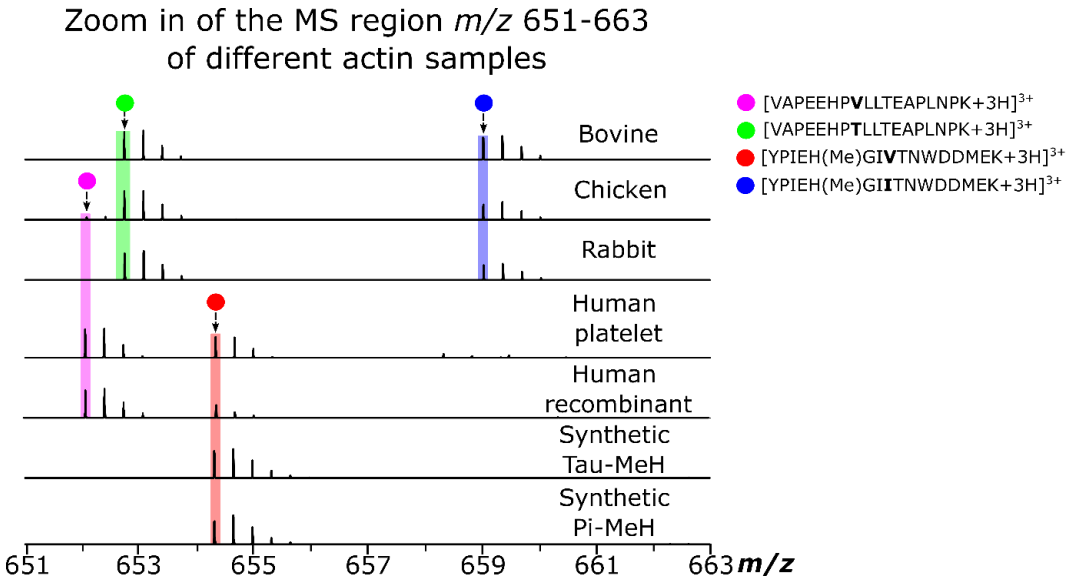


Figure S3. 10 Mass spectral region *m/z* 651-663 depicting the *m/z* shift of the triply charged protonated molecular ion of the target peptide for the animal actin samples due to the V76I amino acid change in the peptide sequence.

Table S3. 3 Comparison of the relative peak intensities (%) of the $[y_4]^+$-$[y_7]^+$ fragments in the synthetic τ-MeH peptide and π-MeH peptide CAD MS/MS spectra.

Fragment	Charge state	Normalised relative intensity (%)	
		τ-MeH peptide	π-MeH peptide
y_4	1+	4.73	4.41
y_5	1+	18.8	18.2
y_6	1+	40.0	40.9
y_7	1+	36.5	36.4

Table S3. 4 Peak assignment table for the CAD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(τ-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN-H_2O	1+	C_8 H_{12} N_3 O_3	198.087318	198.087464	0.74
PI	1+	C_{11} H_{19} N_2 O_2	211.144104	211.144194	0.43
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097971	0.41
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128512	0.25
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.144851	0.07
ME	1+	C_{10} H_{17} N_2 O_4 S_1	261.090354	261.090356	0.01
b_2	1+	C_{14} H_{17} N_2 O_3	261.123370	261.123371	0.00
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113859	-0.03
NW-CO	1+	C_{14} H_{17} N_4 O_2	273.134602	273.13459	-0.04
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155375	-0.08
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124386	-0.16
WD-H_2O	1+	C_{15} H_{14} N_3 O_3	284.102968	284.102927	-0.14
VTN-H_2O	1+	C_{13} H_{21} N_4 O_4	297.155732	297.155661	-0.24
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129437	-0.27
b_5+Me-CO	2+	C_{31} H_{44} N_7 O_7	313.668749	313.668364	-1.23
b_5+Me	2+	C_{32} H_{44} N_7 O_8	327.665932	327.665834	-0.30
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.186579	-0.35
b_6+Me-CO	2+	C_{33} H_{47} N_8 O_8	342.179203	342.179091	-0.33
b_6+Me-H_2O	2+	C_{34} H_{45} N_8 O_8	347.171382	347.171327	-0.16
b_6+Me	2+	C_{34} H_{47} N_8 O_9	356.176386	356.176601	0.60
TNW-CO	1+	C_{18} H_{24} N_5 O_4	374.182281	374.182029	-0.67
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.207276	-0.42
DME	1+	C_{14} H_{22} N_3 O_7 S_1	376.117297	376.117145	-0.40
a_7/b_8+Me-CO	2+	C_{39} H_{58} N_9 O_9	398.721239	398.721212	-0.07
y_6-NH_3	2+	C_{35} H_{48} N_7 O_{13} S_1	403.654904	403.654908	0.01
b_7+Me-H_2O	2+	C_{40} H_{56} N_9 O_9	403.713414	403.713216	-0.49
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.1958742	407.195792	-0.20
y_6	2+	C_{35} H_{51} N_8 O_{13} S_1	412.1681718	412.168069	-0.25
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718589	-0.26
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	421.25578	421.255665	-0.27
a_8/b_8+Me-CO	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.255446	0.00
y_7-NH_3	2+	C_{39} H_{54} N_9 O_{15} S_1	460.676368	460.676193	-0.38
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252815	-0.19
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.18951	-0.28
a_4	1+	C_{24} H_{35} N_4 O_6	475.255111	475.255026	-0.18

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
DDME	1+	C_{18} H_{27} N_4 O_{10} S_1	491.14424	491.144273	0.07
a_9/b_9+Me-CO	2+	C_{48} H_{74} N_{11} O_{12}	498.779285	498.779314	0.06
b_4	1+	C_{25} H_{35} N_4 O_7	503.250026	503.250218	0.38
b_9+Me-H_2O-CO	2+	C_{48} H_{72} N_{11} O_{11}	503.77146	503.771229	-0.46
y_4-H_2O	1+	C_{20} H_{34} N_5 O_8 S_1	504.21226	504.212265	0.01
y_4-NH_3	1+	C_{20} H_{33} N_4 O_9 S_1	505.196276	505.196199	-0.15
y_8-H_2O	2+	C_{43} H_{62} N_{11} O_{16} S_1	510.708199	510.707821	-0.74
y_8-NH_3	2+	C_{43} H_{61} N_{10} O_{17} S_1	511.200207	511.200282	0.15
y_8	2+	C_{43} H_{64} N_{11} O_{17} S_1	519.713482	519.713344	-0.27
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.222962	0.28
NWDD	1+	C_{23} H_{27} N_6 O_9	531.183403	531.183344	-0.11
a_{10}+Me-NH_3	2+	C_{52} H_{77} N_{12} O_{14}	547.287474	547.287366	-0.20
b_{10}+Me-CO	2+	C_{52} H_{80} N_{13} O_{14}	555.801023	555.800582	-0.79
y_9-H_2O	2+	C_{48} H_{71} N_{12} O_{17} S_1	560.242406	560.242404	0.00
y_9-NH_3	2+	C_{48} H_{70} N_{11} O_{18} S_1	560.734414	560.734237	-0.32
b_{10}+Me-H_2O	2+	C_{53} H_{78} N_{13} O_{14}	560.792924	560.792667	-0.46
b_{10}+Me-NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.284677	-0.45
y_{14}+Me	3+	C_{74} H_{114} N_{19} O_{25} S_1	567.603133	567.602767	-0.64
y_9	2+	C_{48} H_{73} N_{12} O_{18} S_1	569.2476838	569.247853	0.30
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.79827	0.11
y_{15}+Me -H_2O	3+	C_{79} H_{119} N_{20} O_{25} S_1	593.950533	593.950201	-0.56
y_{15}+Me	3+	C_{79} H_{121} N_{20} O_{26} S_1	599.954055	599.95414	0.14
y_5-H_2O	1+	C_{24} H_{39} N_6 O_{11} S_1	619.239203	619.239466	0.42
b_5+Me -CO	1+	$C_{31}H_{44}$ N_7 O_7	626.329673	626.330701	1.64
PIEH(Methyl)GI-CO	1+	C_{30} H_{49} N_8 O_7	633.372421	633.371859	-0.89
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249942	0.29
b_{11}+Me -CO	2+	$C_{63}H_{90}$ N_{15} O_{15}	648.840679	648.840124	-0.86
M+H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.30916	1.01
a_6+Me	1+	C_{33} H_{47} N_8 O_8	683.351137	683.351597	0.67
b_6+Me -H_2O	1+	C_{34} H_{45} N_8 O_8	693.335487	693.335809	0.46
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.346052	711.346049	0.00

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351548	0.30
y_{12}+Me	1+	C_{63} H_{96} N_{17} O_{21} S_1	729.837734	729.837735	0.00
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	760.435201	760.435248	0.06
b_{13}+Me -NH_3	2+	C_{72} H_{97} N_{16} O_{22}	769.351531	769.351892	0.47
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864655	-0.19
y_{13}+Me-H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	785.353748	785.35395	0.26
y_{13}+Me	2+	C_{68} H_{103} N_{18} O_{24} S_1	794.35903	794.359069	0.05
b_7+Me -CO	1+	C_{39} H_{58} N_9 O_9	796.43575	796.436446	0.87
y_6-H_2O	1+	C_{35} H_{49} N_8 O_{12} S_1	805.318516	805.319038	0.65
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.420115	0.70
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.330193	1.35
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430272	0.19
b_{14}+Me -H_2O	2+	C_{77} H_{107} N_{18} O_{22} S_1	834.379765	834.380296	0.64
b_{14}+Me -NH_3	2+	C_{77} H_{106} N_{17} O_{23} S_1	834.871773	834.872018	0.29
y_{14}+Me -H_2O	2+	C_{74} H_{112} N_{19} O_{24} S_1	841.89578	841.896303	0.62
y_{14}+Me -NH_3	2+	C_{74} H_{111} N_{18} O_{25} S_1	842.387787	842.387589	-0.24
b_{14}+Me	2+	C_{77} H_{109} N_{18} O_{23} S_1	843.385048	843.386344	1.54
PIEH(Methyl)GIVT-H_2O	1+	C_{40} H_{63} N_{10} O_{10}	843.472315	843.472533	0.26
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.902122	1.25
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.483302	0.49
y_{15}+Me -H_2O	2+	C_{79} H_{119} N_{20} O_{25} S_1	890.422162	890.421493	-0.75
y_{15}+Me -NH_3	2+	C_{79} H_{118} N_{19} O_{26} S_1	890.914169	890.914174	0.01
b_8+Me -CO	1+	C_{44} H_{67} N_{10} O_{10}	895.504163	895.503767	-0.44
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.42863	1.32
b_8-H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.48869	0.80
b_{15}+Me	2+	C_{82} H_{116} N_{19} O_{26} S_1	907.906344	907.906777	0.48
b_{15}+Me +H_2O	2+	C_{82} H_{118} N_{19} O_{27} S_1	916.911627	916.911537	-0.10
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.344953	-0.55

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.499766	1.34
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.373669	1.68
a_9+Me	1+	C_{48} H_{74} N_{11} O_{12}	996.551293	996.551874	0.58
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535459	-0.18
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.392833	-0.30
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546291	0.08
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.421751	2.00
a_{10}+Me -NH_3	1+	C_{52} H_{77} N_{12} O_{14}	1093.567672	1093.567194	-0.44
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488917	0.73
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.587322	-1.59
y_{10}-H_2O	1+	C_{54} H_{82} N_{13} O_{18} S_1	1232.5616	1232.562414	0.66
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.575974	3.05
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593628	0.00
Average error					0.11
Absolute average error					0.47
Standard deviation					0.49

Table S3. 5 Peak assignment table for the CAD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN-H_2O	1+	$C_8 H_{12} N_3 O_3$	198.087318	198.087430	0.57
PI	1+	$C_{11} H_{19} N_2 O_2$	211.144104	211.144210	0.50
TN	1+	$C_8 H_{14} N_3 O_4$	216.097882	216.097967	0.39
a_2	1+	$C_{13} H_{17} N_2 O_2$	233.128454	233.128507	0.23
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144843	0.04
ME	1+	$C_{10} H_{17} N_2 O_4 S_1$	261.090354	261.090361	0.03
b_2	1+	$C_{14} H_{17} N_2 O_3$	261.1233705	261.123371	0.00
NW-CO	1+	$C_{14} H_{17} N_4 O_2$	273.134602	273.134603	0.00
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155384	-0.05
WD-H_2O	1+	$C_{15} H_{14} N_3 O_3$	284.102968	284.102937	-0.11
VTN-H_2O	1+	$C_{13} H_{21} N_4 O_4$	297.155732	297.155685	-0.16
NW	1+	$C_{15} H_{17} N_4 O_3$	301.129517	301.129455	-0.21
PIE	1+	$C_{16} H_{26} N_3 O_5$	340.186697	340.186578	-0.35
b_6+Me	2+	$C_{34} H_{47} N_8 O_9$	356.176386	356.176584	0.56
TNW-CO	1+	$C_{18} H_{24} N_5 O_4$	374.182281	374.182088	-0.52
b_3	1+	$C_{20} H_{28} N_3 O_4$	374.207433	374.207292	-0.38
DME	1+	$C_{14} H_{22} N_3 O_7 S_1$	376.117297	376.117195	-0.27
a_7/b_8+Me -CO	2+	$C_{39} H_{58} N_9 O_9$	398.721239	398.721235	-0.01
y_6-NH_3	2+	$C_{35} H_{48} N_7 O_{13} S_1$	403.654904	403.654908	0.01
b_7+Me -H_2O	2+	$C_{40} H_{56} N_9 O_9$	403.713414	403.713267	-0.36
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.1958742	407.195807	-0.17
y_6	2+	$C_{35} H_{51} N_8 O_{13} S_1$	412.1681718	412.167986	-0.45
b_7+Me	2+	$C_{40} H_{58} N_9 O_{10}$	412.718696	412.718607	-0.22
H(Methyl)GIV	1+	$C_{20} H_{33} N_6 O_4$	421.25578	421.255515	-0.63
a_8/b_8+Me -CO	2+	$C_{44} H_{67} N_{10} O_{10}$	448.255446	448.255467	0.05
y_7-NH_3	2+	$C_{39} H_{54} N_9 O_{15} S_1$	460.676368	460.67611	-0.56
b_8+Me	2+	$C_{45} H_{68} N_{10} O_{11}$	462.252903	462.252774	-0.28
y_7	2+	$C_{39} H_{58} N_{10} O_{15} S_1$	469.189643	469.189622	-0.04
DDME	1+	$C_{18} H_{27} N_4 O_{10} S_1$	491.14424	491.143612	-1.28
a_9/b_9+Me -CO	2+	$C_{48} H_{74} N_{11} O_{12}$	498.779285	498.779283	0.00

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_4	1+	$C_{25}H_{35}N_4O_7$	503.250026	503.249795	-0.46
b_9+Me -H_2O-CO	2+	$C_{48}H_{72}N_{11}O_{11}$	503.77146	503.771188	-0.54
y_4-H_2O	1+	$C_{20}H_{34}N_5O_8S_1$	504.21226	504.211504	-1.50
y_4-NH_3	1+	$C_{20}H_{33}N_4O_9S_1$	505.196276	505.196265	-0.02
y_8-H_2O	2+	$C_{43}H_{62}N_{11}O_{16}S_1$	510.708199	510.708134	-0.13
y_8-NH_3	2+	$C_{43}H_{61}N_{10}O_{17}S_1$	511.200207	511.200335	0.25
y_8	2+	$C_{43}H_{64}N_{11}O_{17}S_1$	519.713482	519.713281	-0.39
y_4	1+	$C_{20}H_{36}N_5O_9S_1$	522.2228142	522.223005	0.37
NWDD	1+	$C_{23}H_{27}N_6O_9$	531.183403	531.184033	1.19
a_{10}+Me -NH_3	2+	$C_{52}H_{77}N_{12}O_{14}$	547.287474	547.287589	0.21
b_{10}+Me -CO	2+	$C_{52}H_{80}N_{13}O_{14}$	555.801023	555.800644	-0.68
y_9-H_2O	2+	$C_{48}H_{71}N_{12}O_{17}S_1$	560.242406	560.242537	0.23
y_9-NH_3	2+	$C_{48}H_{70}N_{11}O_{18}S_1$	560.734414	560.734446	0.06
b_{10}+Me -H_2O	2+	$C_{53}H_{78}N_{13}O_{14}$	560.792924	560.792697	-0.40
b_{10}+Me -NH_3	2+	$C_{53}H_{77}N_{12}O_{15}$	561.284931	561.28473	-0.36
y_{14}+Me	3+	$C_{74}H_{114}N_{19}O_{25}S_1$	567.603133	567.602639	-0.87
y_9	2+	$C_{48}H_{73}N_{12}O_{18}S_1$	569.2476838	569.248403	1.26
b_{10}+Me	2+	$C_{53}H_{80}N_{13}O_{15}$	569.798206	569.798246	0.07
y_{15}+Me -H_2O	3+	$C_{79}H_{119}N_{20}O_{25}S_1$	593.950533	593.950013	-0.88
y_{15}+Me	3+	$C_{79}H_{121}N_{20}O_{26}S_1$	599.954055	599.954112	0.10
y_5-H_2O	1+	$C_{24}H_{39}N_6O_{11}S_1$	619.239203	619.239538	0.54
PIEH(Methyl)GI-CO	1+	$C_{30}H_{49}N_8O_7$	633.372421	633.371569	-1.35
y_5	1+	$C_{24}H_{41}N_6O_{12}S_1$	637.2497542	637.250022	0.42
b_{11}+Me -CO	2+	$C_{63}H_{90}N_{15}O_{15}$	648.840679	648.840013	-1.03
M+H	3+	$C_{88}H_{132}N_{21}$	654.308498	654.309366	1.33

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
		O_{28} S_1			
b_6+Me -H_2O	1+	C_{34} H_{45} N_8 O_8	693.335487	693.335608	0.17
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.346052	711.347168	1.57
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351354	0.03
y_{12}+Me	1+	C_{63} H_{96} N_{17} O_{21} S_1	729.837734	729.83893	1.64
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	760.435201	760.434946	-0.34
b_{13}+Me -NH_3	2+	C_{72} H_{97} N_{16} O_{22}	769.351531	769.351307	-0.29
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864195	-0.78
y_{13}+Me -H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	785.353748	785.354361	0.78
y_{13}+Me	2+	C_{68} H_{103} N_{18} O_{24} S_1	794.35903	794.357908	-1.41
b_7-CO	1+	C_{39} H_{58} N_9 O_9	796.43575	796.435432	-0.40
y_6-H_2O	1+	C_{35} H_{49} N_8 O_{12} S_1	805.318516	805.318547	0.04
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.420557	1.25
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.33015	1.30
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.429725	-0.47
b_{14}+Me -NH_3	2+	C_{77} H_{106} N_{17} O_{23} S_1	834.871773	834.872182	0.49
y_{14}+Me -H_2O	2+	C_{74} H_{112} N_{19} O_{24} S_1	841.89578	841.895171	-0.72
y_{14}+Me -NH_3	2+	C_{74} H_{111} N_{18} O_{25} S_1	842.387787	842.3879	0.13
PIEH(Methyl)GIVT-H_2O	1+	C_{40} H_{63} N_{10} O_{10}	843.472315	843.472421	0.13
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.902415	1.59
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.483761	1.02
y_{15}+Me -H_2O	2+	C_{79} H_{119} N_{20} O_{25} S_1	890.422162	890.421159	-1.13
y_{15}+Me -NH_3	2+	C_{79} H_{118} N_{19} O_{26} S_1	890.914169	890.914266	0.11
b_8-CO	1+	C_{44} H_{67} N_{10} O_{10}	895.504163	895.502969	-1.33
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.428459	1.13
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.488045	0.09
b_{15}+Me	2+	C_{82} H_{116} N_{19}	907.906344	907.906983	0.70

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
		O_{26} S_1			
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.345257	-0.22
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.499758	1.33
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.373543	1.54
PIEH(Methyl)GIVTN	1+	C_{44} H_{71} N_{12} O_{13}	975.525807	975.526076	0.28
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.53514	-0.50
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.39264	-0.49
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.545769	-0.43
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.421324	1.59
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488103	0.01
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.588286	-0.75
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.575435	2.61
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593595	-0.03
Average error					0.03
Absolute average error					0.55
Standard deviation					0.52

Table S3. 6 Peak assignment table for the IRMPD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(τ-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN-H_2O	1+	C_8 H_{12} N_3 O_3	198.087318	198.087371	0.27
PI	1+	C_{11} H_{19} N_2 O_2	211.144104	211.144152	0.23
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097903	0.10
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128483	0.12
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.144858	0.10
ME	1+	C_{10} H_{17} N_2 O_4 S_1	261.090354	261.090358	0.02
b_2	1+	C_{14} H_{17} N_2 O_3	261.123370	261.123371	0.00
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113871	0.02
NW-CO	1+	C_{14} H_{17} N_4 O_2	273.134602	273.134621	0.07
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155396	0.00
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124403	-0.10
WD-H_2O	1+	C_{15} H_{14} N_3 O_3	284.102968	284.10297	0.01
VTN-H_2O	1+	C_{13} H_{21} N_4 O_4	297.155732	297.155627	-0.35
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.12949	-0.09
b_5+Me -CO	2+	C_{31} H_{44} N_7 O_7	313.668749	313.668378	-1.18
b_5+Me	2+	C_{32} H_{44} N_7 O_8	327.665932	327.665891	-0.13
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.186674	-0.07
b_6+Me	2+	C_{34} H_{47} N_8 O_9	356.176386	356.176675	0.81
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.207274	-0.42
DME	1+	C_{14} H_{22} N_3 O_7 S_1	376.117297	376.117227	-0.19
a_7/b_8+Me -CO	2+	C_{39} H_{58} N_9 O_9	398.721239	398.721206	-0.08
y_6-NH_3	2+	C_{35} H_{48} N_7 O_{13} S_1	403.654904	403.654936	0.08
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.1958742	407.195774	-0.25
y_6	2+	C_{35} H_{51} N_8 O_{13} S_1	412.1681718	412.168141	-0.07
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718611	-0.21
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	421.25578	421.255649	-0.31
a_8/b_8+Me -CO	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.255445	0.00
y_7-NH_3	2+	C_{39} H_{54} N_9 O_{15} S_1	460.676368	460.676175	-0.42
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252818	-0.18
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.189561	-0.17
DDME	1+	C_{18} H_{27} N_4 O_{10} S_1	491.14424	491.144003	-0.48
a_9/b_9+Me -CO	2+	C_{48} H_{74} N_{11} O_{12}	498.779285	498.779252	-0.07

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_4	1+	$C_{25}H_{35}N_4O_7$	503.250026	503.249939	-0.17
b_9+Me -H_2O-CO	2+	$C_{48}H_{72}N_{11}O_{11}$	503.77146	503.771364	-0.19
y_4-H_2O	1+	$C_{20}H_{34}N_5O_8S_1$	504.21226	504.212027	-0.46
y_4-NH_3	1+	$C_{20}H_{33}N_4O_9S_1$	505.196276	505.195871	-0.80
y_8-NH_3	2+	$C_{43}H_{61}N_{10}O_{17}S_1$	511.200207	511.200111	-0.19
y_8	2+	$C_{43}H_{64}N_{11}O_{17}S_1$	519.713482	519.713236	-0.47
y_4	1+	$C_{20}H_{36}N_5O_9S_1$	522.2228142	522.222688	-0.24
a_{10}+Me -NH_3	2+	$C_{52}H_{77}N_{12}O_{14}$	547.287474	547.287325	-0.27
b_{10}+Me -CO	2+	$C_{52}H_{80}N_{13}O_{14}$	555.801023	555.800441	-1.05
y_9-H_2O	2+	$C_{48}H_{71}N_{12}O_{17}S_1$	560.242406	560.242048	-0.64
b_{10}+Me -H_2O	2+	$C_{53}H_{78}N_{13}O_{14}$	560.792924	560.792704	-0.39
b_{10}+Me -NH_3	2+	$C_{53}H_{77}N_{12}O_{15}$	561.284931	561.284854	-0.14
y_9	2+	$C_{48}H_{73}N_{12}O_{18}S_1$	569.2476838	569.247537	-0.26
b_{10}+Me	2+	$C_{53}H_{80}N_{13}O_{15}$	569.798206	569.798099	-0.19
y_5-H_2O	1+	$C_{24}H_{39}N_6O_{11}S_1$	619.239203	619.23918	-0.04
PIEH(Methyl)GI-CO	1+	$C_{30}H_{49}N_8O_7$	633.372421	633.371787	-1.00
y_5	1+	$C_{24}H_{41}N_6O_{12}S_1$	637.2497542	637.249757	0.00
M+3H-2H_2O	3+	$C_{88}H_{126}N_{21}O_{26}S_1$	642.301455	642.301526	0.11
M+3H-H_2O	3+	$C_{88}H_{128}N_{21}O_{27}S_1$	648.304976	648.305162	0.29
M+3H	3+	$C_{88}H_{130}N_{21}O_{28}S_1$	654.308498	654.307913	-0.89
PIEH(Methyl)GI	1+	$C_{29}H_{49}N_8O_6$	661.366787	661.366768	-0.03
b_{11}+Me	2+	$C_{64}H_{91}N_{15}O_{16}$	662.837862	662.837084	-1.17
b_6+Me	1+	$C_{34}H_{47}N_8O_9$	711.346052	711.345691	-0.51
b_{12}+Me	2+	$C_{68}H_{95}N_{16}O_{19}$	720.351334	720.351358	0.03
y_{12}+Me	1+	$C_{63}H_{96}N_{17}O_{21}S_1$	729.837734	729.83809	0.49
PIEH(Methyl)GIV	1+	$C_{36}H_{58}N_9O_9$	760.435201	760.434953	-0.33
b_{13}+Me -NH_3	2+	$C_{72}H_{97}N_{16}O_{22}$	769.351531	769.351649	0.15
b_{13}+Me	2+	$C_{72}H_{100}N_{17}O_{22}$	777.864805	777.86466	-0.19
y_6-H_2O	1+	$C_{35}H_{49}N_8O_{12}S_1$	805.318516	805.318176	-0.42

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329475	0.48
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430079	-0.04
b_{14}+Me -NH_3	2+	C_{77} H_{106} N_{17} O_{23} S_1	834.871773	834.871448	-0.39
PIEH(Methyl)GIVT -H_2O	1+	C_{40} H_{63} N_{10} O_{10}	843.472315	843.47256	0.29
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.901451	0.46
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.482766	-0.13
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.488203	0.26
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.345379	-0.09
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.49906	0.57
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.372763	0.71
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535328	-0.31
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.393894	0.74
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546348	0.14
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.420354	0.65
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.562512	-0.07
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488088	0.00
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.588269	-0.76
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.59504	1.08
Average error					-0.09
Absolute average error					0.31
Standard deviation					0.29

Table S3. 7 Peak assignment table for the IRMPD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN-H_2O	1+	C_8 H_{12} N_3 O_3	198.087318	198.087307	-0.06
PI	1+	C_{11} H_{19} N_2 O_2	211.144104	211.144106	0.01
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097903	0.10
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128475	0.09
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.14478	-0.21
ME	1+	C_{10} H_{17} N_2 O_4 S_1	261.090354	261.090318	-0.14
b_2	1+	C_{14} H_{17} N_2 O_3	261.1233705	261.123371	0.00
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155408	0.04
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129494	-0.08
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.186614	-0.24
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.207267	-0.44
a_7+Me	2+	C_{39} H_{58} N_9 O_9	398.721239	398.721222	-0.04
y_6-NH_3	1+	C_{35} H_{48} N_7 O_{13} S_1	403.654904	403.654855	-0.12
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.1958742	407.19592	0.11
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	412.1681718	412.168167	-0.01
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718636	-0.15
a_8+Me	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.25557	0.28
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	460.676368	460.676095	-0.59
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252841	-0.13
y_7	1+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.189603	-0.09
a_9+Me	2+	C_{48} H_{74} N_{11} O_{12}	498.779285	498.779218	-0.13
b_4	1+	C_{25} H_{35} N_4 O_7	503.250026	503.249605	-0.84
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	503.77146	503.771297	-0.32
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	511.200207	511.20018	-0.05
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	519.713482	519.713232	-0.48
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.222823	0.02
a_{10}+Me -NH_3	2+	C_{52} H_{77} N_{12} O_{14}	547.287474	547.287037	-0.80
y_9-H_2O	1+	C_{48} H_{71} N_{12} O_{17} S_1	560.242406	560.242069	-0.60
b_{10}+Me -NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.284784	-0.26
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	569.2476838	569.247515	-0.30

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.798018	-0.33
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	599.954055	599.953938	-0.20
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249753	0.00
M+3H-2H_2O	3+	C_{88} $H_{126}N_{21}$ O_{26} S_1	642.301455	642.301161	-0.46
M+3H-H_2O	3+	C_{88} H_{128} N_{21} O_{27} S_1	648.304976	648.305392	0.64
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.309033	0.82
PIEH(Methyl)GI	1+	C_{31} H_{49} N_8 O_8	661.366787	661.365728	-1.60
b_{11}+Me	2+	C_{64} H_{90} N_{15} O_{16}	662.837862	662.837837	-0.04
b_6+Me	2+	C_{34} H_{47} N_8 O_9	711.346052	711.345745	-0.43
b_{12}+Me -NH_3	2+	C_{68} H_{92} N_{15} O_{19}	711.838059	711.837656	-0.57
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351638	0.42
b_{13}+Me -NH_3	2+	C_{72} H_{97} N_{16} O_{22}	769.351531	769.350739	-1.03
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864484	-0.41
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329726	0.78
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.429788	-0.40
PIEH(Methyl)GIVT-H_2O	1+	C_{40} H_{63} N_{10} O_{10}	843.472315	843.47121	-1.31
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.90148	0.49
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.426425	-1.13
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.344013	-1.57
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.498949	0.45
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.372608	0.55
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.53434	-1.29
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.41975	0.07
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.5609	-1.51
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.48808	-0.01
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.59108	1.71
Average error					-0.19
Absolute average error					0.44
Standard deviation					0.47

Table S3. 8 Peak assignment table for the 193 nm MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(τ-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
W	1+	C_{10} H_{11} N_2	159.091675	159.09178	0.66
PI-CO	1+	C_{10} H_{19} N_2 O_1	183.14919	183.149251	0.33
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097916	0.16
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128477	0.10
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.144854	0.08
b_2	1+	C_{14} H_{17} N_2 O_3	261.1233705	261.123371	0.00
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113883	0.06
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155409	0.04
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124448	0.06
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129514	-0.01
H(Methyl)GI	1+	C_{15} H_{24} N_5 O_3	322.187366	322.187362	-0.01
b_5+Me	2+	C_{32} H_{44} N_7 O_8	327.665932	327.665952	0.06
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.18669	-0.02
a_3	1+	C_{19} H_{28} N_3 O_3	346.212518	346.212498	-0.06
b_6+Me	2+	C_{34} H_{47} N_8 O_9	356.176386	356.176656	0.76
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.20742	-0.03
z_3	1+	C_{16} H_{29} N_3 O_6 S_1	391.177158	391.177145	-0.03
c_3	1+	C_{20} H_{31} N_4 O_4	391.233982	391.233926	-0.14
IEH(Methyl)	1+	C_{18} H_{28} N_5 O_5	394.208495	394.208496	0.00
a_7+Me	2+	C_{39} H_{58} N_9 O_9	398.721239	398.721234	-0.01
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.1958742	407.195842	-0.08
y_6	2+	C_{35} H_{51} N_8 O_{13} S_1	412.1681718	412.168169	-0.01
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718704	0.02
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	421.25578	421.255767	-0.03
IEH(Methyl)G-H_2O	1+	C_{20} H_{29} N_6 O_5	433.219394	433.219358	-0.08
a_8+Me	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.255388	-0.13
y_7-NH_3	2+	C_{39} H_{54} N_9 O_{15} S_1	460.676368	460.676279	-0.19
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252934	0.07
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.189575	-0.14
a_4	1+	C_{24} H_{35} N_4 O_6	475.255111	475.255044	-0.14
PIEH(Methyl)	1+	C_{23} H_{35} N_6 O_6	491.261259	491.261188	-0.14
b_4	1+	C_{25} H_{35} N_4 O_7	503.250026	503.249981	-0.09
b_9+Me -H_2O	2+	C_{49} H_{72} N_{11} O_{12}	503.77146	503.771525	0.13
z_4	1+	C_{20} H_{34} N_4 O_9 S_1	506.204101	506.204054	-0.09

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	511.200207	511.200362	0.30
b_9+Me	2+	C_{49} H_{74} N_{11} O_{13}	512.776742	512.776726	-0.03
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	519.713482	519.713469	-0.03
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.222798	-0.03
EH(Methyl)GIV-H_2O	1+	C_{25} H_{38} N_7 O_6	532.287808	532.287844	0.07
PIEH(Methyl)G	1+	C_{25} H_{38} N_7 O_7	548.282723	548.282702	-0.04
b_{10}+Me -NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.285026	0.17
y_9	2+	C_{48} H_{73} N_{12} O_{18} S_1	569.247689	569.247688	0.00
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.798258	0.09
c_{10}+Me	2+	C_{53} H_{83} N_{14} O_{15}	578.31148	578.311542	0.11
y_{15}+Me	3+	C_{79} H_{121} N_{20} O_{26} S_1	599.954055	599.954147	0.15
a_5+Me	1+	C_{31} H_{44} N_7 O_7	626.329673	626.329656	-0.03
b_5+Me -H_2O	1+	C_{32} H_{42} N_7 O_7	636.314023	636.314105	0.13
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249756	0.00
GIVTNW-CO	1+	C_{31} H_{47} N_8 O_7	643.356222	643.35612	-0.16
M+3H-H_2O	3+	C_{88} H_{128} N_{21} O_{27} S_1	648.304976	648.305091	0.18
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.307583	-1.40
PIEH(Methyl)GI	1+	C_{31} H_{49} N_8 O_8	661.366787	661.366605	-0.28
b_{11}+Me	2+	C_{64} H_{90} N_{15} O_{16}	662.837862	662.838152	0.44
c_5+Me	1+	C_{32} H_{47} N_8 O_8	671.351137	671.351108	-0.04
b_6+Me -H_2O	1+	C_{34} H_{45} N_8 O_8	693.335487	693.335456	-0.04
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.346052	711.345955	-0.14
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351286	-0.07
c_6+Me	1+	C_{34} H_{50} N_9 O_9	728.372601	728.372567	-0.05
y_{12}+Me	2+	C_{63} H_{96} N_{17} O_{21} S_1	729.837734	729.8378	0.09
x_{12}	2+	C_{64} H_{95} N_{17} O_{22} S_1	742.827366	742.827145	-0.30
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	760.435201	760.435123	-0.10
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864693	-0.14
y_{13}+Me -H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	785.353748	785.353911	0.21
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.434967	-0.29
y_6-H_2O	1+	C_{35} H_{49} N_8 O_{12} S_1	805.318516	805.31855	0.04

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.419551	0.00
x_{13}	2+	C_{69} H_{102} N_{18} O_{25} S_1	807.348662	807.34893	0.33
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329089	0.01
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430128	0.01
b_{14}+Me	2+	C_{77} H_{109} N_{18} O_{23} S_1	843.385048	843.385706	0.78
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.900851	-0.25
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.482681	-0.23
x_{14}	2+	C_{75} H_{113} N_{19} O_{26} S_1	863.890694	863.891243	0.64
y_{15}+Me -H_2O	2+	C_{79} H_{119} N_{20} O_{25} S_1	890.422162	890.422613	0.51
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.428128	0.76
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.487793	-0.19
b_{15}+Me	2+	C_{82} H_{116} N_{19} O_{26} S_1	907.906344	907.906079	-0.29
y_7-H_2O	1+	C_{39} H_{55} N_{10} O_{14} S_1	919.361444	919.361158	-0.31
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.345251	-0.23
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.498428	-0.11
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.372029	-0.07
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.524844	-0.25
PIEH(Methyl)GIVTN	1+	C_{44} H_{71} N_{12} O_{13}	975.525807	975.526052	0.25
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535391	-0.25
y_8-H_2O	1+	C_{43} H_{62} N_{11} O_{16} S_1	1020.409122	1020.408687	-0.43
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.392954	-0.18
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546135	-0.07
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419664	-0.01
y_9-H_2O	1+	C_{48} H_{71} N_{12} O_{17} S_1	1119.477536	1119.477597	0.05
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.562503	-0.07
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488073	-0.02

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.58885	-0.25
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615684	1155.615324	-0.31
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.572724	0.45
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593188	-0.34
Average error					-0.01
Absolute average error					0.17
Standard deviation					0.21

Table S3. 9 Peak assignment table for the 193 nm MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
W	1+	C_{10} H_{11} N_2	159.091675	159.091667	-0.05
PI-CO	1+	C_{10} H_{19} N_2 O_1	183.14919	183.149305	0.63
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097801	-0.37
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128452	-0.01
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.14484	0.03
b_2	1+	C_{14} H_{17} N_2 O_3	261.123370 5	261.123371	0.00
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113881	0.05
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155352	-0.16
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124426	-0.02
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129392	-0.42
H(Methyl)GI	1+	C_{15} H_{24} N_5 O_3	322.187366	322.18734	-0.08
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.18667	-0.08
a_3	1+	C_{19} H_{28} N_3 O_3	346.212518	346.212471	-0.14
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.207225	-0.56
z_3	1+	C_{16} H_{29} N_3 O_6 S_1	391.177158	391.177023	-0.35
IEH(Methyl)	1+	C_{18} H_{28} N_5 O_5	394.208495	394.208495	0.00
a_7+Me	1+	C_{39} H_{58} N_9 O_9	398.721239	398.721143	-0.24
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.195874 2	407.195871	-0.01
y_6	2+	C_{35} H_{51} N_8 O_{13} S_1	412.168171 8	412.168089	-0.20
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718706	0.02

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	421.25578	421.255719	-0.14
IEH(Methyl)G-H_2O	1+	C_{20} H_{29} N_6 O_5	433.219394	433.219336	-0.13
a_8+Me	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.255399	-0.10
y_7-NH_3	2+	C_{39} H_{54} N_9 O_{15} S_1	460.676368	460.67623	-0.30
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252843	-0.13
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.189533	-0.23
a_4	1+	C_{24} H_{35} N_4 O_6	475.255111	475.254945	-0.35
PIEH(Methyl)	1+	C_{23} H_{35} N_6 O_6	491.261259	491.261196	-0.13
b_4	1+	C_{25} H_{35} N_4 O_7	503.250026	503.249855	-0.34
b_9+Me -H_2O	2+	C_{49} H_{72} N_{11} O_{12}	503.77146	503.771445	-0.03
z_4	1+	C_{20} H_{34} N_4 O_9 S_1	506.204101	506.203867	-0.46
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	511.200207	511.200203	-0.01
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	512.776742	512.776624	-0.23
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	519.713482	519.713461	-0.04
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.22282	0.01
EH(Methyl)GIV-H_2O	1+	C_{25} H_{38} N_7 O_6	532.287808	532.287675	-0.25
PIEH(Methyl)G	1+	C_{25} H_{38} N_7 O_7	548.282723	548.282531	-0.35
b_{10}+Me -NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.284854	-0.14
y_9	2+	C_{48} H_{73} N_{12} O_{18} S_1	569.247689	569.24763	-0.10
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.798197	-0.02
c_{10}+Me	2+	C_{53} H_{83} N_{14} O_{15}	578.31148	578.31148	0.00
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	599.954055	599.954395	0.57
a_5+Me	1+	C_{31} H_{44} N_7 O_7	626.329673	626.329626	-0.08
b_5+Me -H_2O	1+	C_{32} H_{42} N_7 O_7	636.314023	636.313986	-0.06
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249876	0.19
GIVTNW-CO	1+	C_{31} H_{47} N_8 O_7	643.356222	643.35644	0.34
M+3H-H_2O	3+	C_{88} H_{128} N_{21} O_{27} S_1	648.304976	648.30529	0.48
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.306831	-2.55

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
PIEH(Methyl)GI	1+	C_{31} H_{49} N_8 O_8	661.366787	661.366865	0.12
b_{11}+Me	2+	C_{64} H_{90} N_{15} O_{16}	662.837862	662.837833	-0.04
c_5+Me	1+	C_{32} H_{47} N_8 O_8	671.351137	671.351159	0.03
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.346052	711.3458	-0.35
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351649	0.44
c_6+Me	1+	C_{34} H_{50} N_9 O_9	728.372601	728.372222	-0.52
y_{12}+Me	2+	C_{63} H_{96} N_{17} O_{21} S_1	729.837734	729.838107	0.51
x_{12}	2+	C_{64} H_{95} N_{17} O_{22} S_1	742.827366	742.82722	-0.20
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	760.435201	760.435269	0.09
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864824	0.02
y_{13}+Me -H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	785.353748	785.353689	-0.08
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.434924	-0.35
y_6-H_2O	1+	C_{35} H_{49} N_8 O_{12} S_1	805.318516	805.318189	-0.41
x_{13}	2+	C_{69} H_{102} N_{18} O_{25} S_1	806.419551	806.419569	0.31
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	807.348662	807.348914	0.02
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329326	0.30
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430208	0.11
b_{14}+Me	2+	C_{77} H_{109} N_{18} O_{23} S_1	843.385048	843.384241	-0.96
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.901291	0.27
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.482385	-0.57
x_{14}	2+	C_{75} H_{113} N_{19} O_{26} S_1	863.890694	863.890991	0.34
y_{15}+Me -H_2O	2+	C_{79} H_{119} N_{20} O_{25} S_1	890.422162	890.422119	-0.05
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.427824	0.42
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.487673	-0.32
b_{15}+Me	2+	C_{82} H_{116} N_{19} O_{26} S_1	907.906344	907.906553	0.23
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.34538	-0.09
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.498854	0.35

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.372011	-0.09
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.525239	0.17
PIEH(Methyl)GIVTN	1+	C_{44} H_{71} N_{12} O_{13}	975.525807	975.52586	0.05
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535636	-0.01
y_8-H_2O	1+	C_{43} H_{62} N_{11} O_{16} S_1	1020.409122	1020.408512	-0.60
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.393268	0.13
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546244	0.04
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419896	0.21
y_9-H_2O	1+	C_{48} H_{71} N_{12} O_{17} S_1	1119.477536	1119.478152	0.55
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488085	-0.01
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.589166	0.03
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615684	1155.61568	0.00
Average error					-0.09
Absolute average error					0.24
Standard deviation					0.32

Table S3. 10 Peak assignment table for the 213 nm MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(τ-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
W	1+	C_{10} H_{11} N_2	159.091675	159.091675	0.00
PI-CO	1+	C_{10} H_{19} N_2 O_1	183.14919	183.149196	0.03
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097884	0.01
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.144834	0.00
b_2	1+	C_{14} H_{17} N_2 O_3	261.1233705	261.123367	-0.01
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113877	0.04
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155405	0.03
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124441	0.04
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129507	-0.03
H(Methyl)GI	1+	C_{15} H_{24} N_5 O_3	322.187366	322.187391	0.08
b_5+Me	2+	C_{32} H_{44} N_7 O_8	327.665932	327.665933	0.00
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.186724	0.08
b_3	1+	C_{20} H_{28} N_3 O_4	374.207433	374.207461	0.07
z_3	1+	C_{16} H_{29} N_3 O_6 S_1	391.177158	391.177196	0.10
IEH(Methyl)	1+	C_{18} H_{28} N_5 O_5	394.208495	394.208521	0.07
a_7+Me	2+	C_{39} H_{58} N_9 O_9	398.721239	398.721237	-0.01
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.1958742	407.195873	0.00
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.718696	412.718706	0.02
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	421.25578	421.255835	0.13
IEH(Methyl)G-H_2O	1+	C_{20} H_{29} N_6 O_5	433.219394	433.219441	0.11
a_8+Me	2+	C_{44} H_{67} N_{10} O_{10}	448.255446	448.255463	0.04
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	462.252903	462.252916	0.03
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	469.189643	469.189645	0.00
c_8+Me	2+	C_{45} H_{70} N_{11} O_{11}	470.766178	470.766195	0.04
PIEH(Methyl)	1+	C_{23} H_{35} N_6 O_6	491.261259	491.261334	0.15
a_9+Me	2+	C_{48} H_{74} N_{11} O_{12}	498.779285	498.779298	0.03
b_4	1+	C_{25} H_{35} N_4 O_7	503.250026	503.250083	0.11
b_9+Me -H_2O	2+	C_{49} H_{72} N_{11} O_{12}	503.77146	503.771493	0.07
z_4	1+	C_{20} H_{34} N_4 O_9 S_1	506.204101	506.204153	0.10
b_9+Me	2+	C_{49} H_{74} N_{11}	512.776742	512.776756	0.03

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
		O_{13}			
c_4	1+	C_{25} H_{38} N_5 O_7	520.276575	520.276598	0.04
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.222838	0.05
EH(Methyl)GIV-H_2O	1+	C_{25} H_{38} N_7 O_6	532.287808	532.287866	0.11
PIEH(Methyl)G	1+	C_{25} H_{38} N_7 O_7	548.282723	548.282811	0.16
a_{10}+Me	2+	C_{52} H_{80} N_{13} O_{14}	555.800749	555.800834	0.15
b_{10}+Me -NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.28498	0.09
y_9	2+	C_{48} H_{73} N_{12} O_{18} S_1	569.247689	569.24769	0.00
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.798267	0.11
c_{10}+Me	2+	C_{53} H_{83} N_{14} O_{15}	578.31148	578.311496	0.03
z_5	1+	C_{24} H_{39} N_5 O_{12} S_1	621.231044	621.231019	-0.04
a_5+Me	1+	C_{31} H_{44} N_7 O_7	626.329673	626.329807	0.21
b_5+Me -H_2O	1+	C_{32} H_{42} N_7 O_7	636.314023	636.314094	0.11
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249794	0.06
GIVTNW-CO	1+	C_{31} H_{47} N_8 O_7	643.356222	643.356278	0.09
M+3H-H_2O	3+	C_{88} H_{128} N_{21} O_{27} S_1	648.304976	648.304893	-0.13
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.307622	-1.34
PIEH(Methyl)GI	1+	C_{31} H_{49} N_8 O_8	661.366787	661.366846	0.09
b_{11}+Me	2+	C_{64} H_{90} N_{15} O_{16}	662.837862	662.838092	0.35
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.346052	711.34611	0.08
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	720.351334	720.351289	-0.06
y_{12}+Me	2+	C_{63} H_{96} N_{17} O_{21} S_1	729.837734	729.837737	0.00
x_{12}	2+	C_{64} H_{95} N_{17} O_{22} S_1	742.827366	742.827318	-0.06
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	760.435201	760.4353	0.13
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	777.864805	777.864751	-0.07
y_{13}+Me -H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	785.353748	785.353739	-0.01
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.435154	-0.06
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.419621	0.09

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
x_{13}	2+	C_{69} H_{102} N_{18} O_{25} S_1	807.348662	807.349682	1.26
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.328988	-0.11
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430295	0.22
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.900879	-0.22
PIEH(Methyl)GIVT	1+	C_{40} H_{65} N_{10} O_{11}	861.482879	861.483007	0.15
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.427984	0.60
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	905.487965	905.488022	0.06
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.34546	920.345438	-0.02
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.498578	0.05
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.372008	-0.10
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.525202	0.13
PIEH(Methyl)GIVTN	1+	C_{44} H_{71} N_{12} O_{13}	975.525807	975.525821	0.01
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.53566	0.02
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.392824	-0.31
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.545208	1024.546313	0.10
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.4196	-0.07
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.562601	0.01
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488072	-0.02
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.589095	-0.04
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615684	1155.615552	-0.11
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593198	-0.33
Average error					0.02
Absolute average error					0.10
Standard deviation					0.17

Table S3. 11 Peak assignment table for the 213 nm MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
W	1+	C_{10} H_{11} N_2	159.091675	159.091677	0.01
PI-CO	1+	C_{10} H_{19} N_2 O_1	183.14919	183.149194	0.02
TN	1+	C_8 H_{14} N_3 O_4	216.097882	216.097884	0.01
a_2	1+	C_{13} H_{17} N_2 O_2	233.128454	233.128456	0.01
y_2-H_2O	1+	C_{11} H_{20} N_3 O_4	258.144833	258.144834	0.00
b_2	1+	C_{14} H_{17} N_2 O_3	261.1233705	261.123372	0.01
EH(Methyl)-H_2O	1+	C_{12} H_{15} N_4 O_3	263.113867	263.113887	0.08
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155403	0.02
EH(Methyl)	1+	C_{12} H_{17} N_4 O_4	281.124431	281.124454	0.08
NW	1+	C_{15} H_{17} N_4 O_3	301.129517	301.129499	-0.06
H(Methyl)GI	1+	C_{15} H_{24} N_5 O_3	322.187366	322.187394	0.09
PIE	1+	C_{16} H_{26} N_3 O_5	340.186697	340.186715	0.05
b_3	1+	C_{20} H_{28} N_3 O_4	346.212518	374.207457	0.06
z_3	1+	C_{16} H_{29} N_3 O_6 S_1	374.207433	391.177177	0.05
IEH(Methyl)	1+	C_{18} H_{28} N_5 O_5	391.177158	394.208531	0.09
a_7+Me	2+	C_{39} H_{58} N_9 O_9	394.208495	398.72125	0.03
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	398.721239	407.195871	-0.01
y_6	2+	C_{35} H_{51} N_8 O_{13} S_1	407.1958742	412.168226	0.13
b_7+Me	2+	C_{40} H_{58} N_9 O_{10}	412.1681718	412.718694	0.00
H(Methyl)GIV	1+	C_{20} H_{33} N_6 O_4	412.718696	421.25583	0.12
IEH(Methyl)G-H_2O	1+	C_{20} H_{29} N_6 O_5	421.25578	433.21945	0.13
a_8+Me	2+	C_{44} H_{67} N_{10} O_{10}	433.219394	448.255438	-0.02
y_7-NH_3	2+	C_{39} H_{54} N_9 O_{15} S_1	448.255446	460.676394	0.06
b_8+Me	2+	C_{45} H_{68} N_{10} O_{11}	460.676368	462.252907	0.01
y_7	2+	C_{39} H_{58} N_{10} O_{15} S_1	462.252903	469.189618	-0.05
PIEH(Methyl)	1+	C_{23} H_{35} N_6 O_6	469.189643	491.261334	0.15
a_9+Me	2+	C_{48} H_{74} N_{11} O_{12}	475.255111	498.779322	0.07
b_4	1+	C_{25} H_{35} N_4 O_7	491.261259	503.250057	0.06
b_9+Me -H_2O	2+	C_{49} H_{72} N_{11} O_{12}	503.250026	503.77148	0.04
z_4	1+	C_{20} H_{34} N_4 O_9 S_1	503.77146	506.20416	0.12
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	506.204101	511.200202	-0.01
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	511.200207	512.776744	0.00
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	512.776742	519.713478	-0.01
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	519.713482	522.222831	0.03
EH(Methyl)GIV-	1+	C_{25} H_{38} N_7 O_6	522.222814	532.287858	0.09

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
H_2O			2		
PIEH(Methyl)G	1+	C_{25} H_{38} N_7 O_7	532.287808	548.282825	0.19
a_{10}+Me	2+	C_{52} H_{80} N_{13} O_{14}	548.282723	555.800798	0.09
b_{10}-NH_3	2+	C_{53} H_{77} N_{12} O_{15}	561.284931	561.284943	0.02
y_9	2+	C_{48} H_{73} N_{12} O_{18} S_1	569.247689	569.247649	-0.07
b_{10}+Me	2+	C_{53} H_{80} N_{13} O_{15}	569.798206	569.798216	0.02
c_{10}+Me	2+	C_{53} H_{83} N_{14} O_{15}	578.31148	578.311472	-0.01
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	599.954055	599.954193	0.23
z_5	1+	C_{24} H_{39} N_5 O_{12} S_1	626.329673	621.231057	0.02
a_5+Me	1+	C_{31} H_{44} N_7 O_7	636.314023	626.329666	-0.01
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249768	0.02
M+3H-H_2O	3+	C_{88} H_{128} N_{21} O_{27} S_1	643.356222	648.304874	-0.16
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	648.304976	654.307797	-1.07
PIEH(Methyl)GI	1+	C_{31} H_{49} N_8 O_8	654.308498	661.366809	0.03
b_{11}+Me	2+	C_{64} H_{90} N_{15} O_{16}	661.366787	662.83818	0.48
c_5+Me	1+	C_{32} H_{47} N_8 O_8	662.837862	671.351176	0.06
b_6+Me	1+	C_{34} H_{47} N_8 O_9	671.351137	711.346041	-0.02
b_{12}+Me	2+	C_{68} H_{95} N_{16} O_{19}	711.346052	720.351272	-0.09
c_6+Me	1+	C_{34} H_{50} N_9 O_9	720.351334	728.372632	0.04
PIEH(Methyl)GIV	1+	C_{36} H_{58} N_9 O_9	728.372601	760.435251	0.07
b_{13}+Me	2+	C_{72} H_{100} N_{17} O_{22}	729.837734	777.864734	-0.09
x_{12}	2+	C_{64} H_{95} N_{17} O_{22} S_1	742.827366	742.827004	-0.49
y_{13}+Me -H_2O	2+	C_{68} H_{101} N_{18} O_{23} S_1	760.435201	785.353765	0.02
a_7+Me	1+	C_{39} H_{58} N_9 O_9	777.864805	796.435281	0.10
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	785.353748	806.41951	-0.05
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	796.435201	823.328962	-0.14
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	805.318516	824.430166	0.06
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	806.419551	850.900971	-0.11
x_{13}	2+	C_{69} H_{102} N_{18} O_{25} S_1	807.348662	807.348837	0.22
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	823.329081	899.427932	0.54
b_8+Me -H_2O	1+	C_{45} H_{65} N_{10} O_{10}	824.430116	905.488163	0.22
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	843.385048	920.345458	0.00
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	850.901062	923.498529	0.00
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	861.482879	937.372048	-0.05
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	890.422162	940.525028	-0.05
PIEH(Methyl)GIVTN	1+	C_{44} H_{71} N_{12} O_{13}	899.427444	975.525874	0.07
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	905.487965	1006.535642	0.00
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	907.906344	1021.39322	0.08

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
				4	
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	920.34546	1024.546371	0.16
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	923.498529	1038.419501	-0.17
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	937.3720971	1137.488067	-0.02
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	940.525079	1138.589076	-0.05
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	975.525807	1155.615594	-0.08
Average error					0.02
Absolute average error					0.09
Standard deviation					0.15

Table S3. 12 Peak assignment table for the ECD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(τ-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
VT	1+	C_9 H_{17} N_2 O_3	201.123369	201.123386	0.08
z_2	1+	C_{11} H_{20} N_2 O_5	260.136678	260.136673	-0.02
y_2	1+	C_{11} H_{22} N_3 O_5	276.155397	276.155432	0.13
c_2	1+	C_{14} H_{20} N_3 O_3	278.149918	278.149952	0.12
z_3-C_3H_6S	1+	C_{13} H_{23} N_3 O_6	317.158138	317.158178	0.13
z_3	1+	C_{16} H_{29} N_3 O_6 S_1	391.177159	391.177158	0.00
c_3	1+	C_{20} H_{31} N_4 O_4	391.233982	391.234009	0.07
y_3	1+	C_{16} H_{31} N_4 O_6 S_1	407.195882	407.19591	0.07
z_4-CO_2	1+	C_{19} H_{34} N_4 O_7 S_1	462.214273	462.214336	0.14
z_4	1+	C_{20} H_{34} N_4 O_9 S_1	506.204101	506.204102	0.00
c_4	1+	C_{25} H_{38} N_5 O_7	520.276575	520.276556	-0.04
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.222827	522.222951	0.24
z_5-CO_2	1+	C_{23} H_{39} N_5 O_{10} S_1	577.241217	577.2414	0.32
z_5	1+	C_{24} H_{39} N_5 O_{12} S_1	621.231044	621.231167	0.20
a_5+Me	1+	C_{31} H_{44} N_7 O_7	626.329673	626.329903	0.37
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.249771	637.24989	0.19
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.308453	-0.07
c_5+Me	1+	C_{32} H_{47} N_8 O_8	671.351138	671.351148	0.01
c_6+Me	1+	C_{34} H_{50} N_9 O_9	728.372602	728.37274	0.19
z_6-C_3H_6S	1+	C_{32} H_{43} N_7 O_{13}	733.291339	733.291638	0.41
z_6-CO_2	1+	C_{34} H_{49} N_7 O_{11} S_1	763.32053	763.320694	0.21

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_7-C_9H_7N	1+	C_{30} H_{48} N_8 O_{15} S_1	792.295439	792.295879	0.56
y_{13}+Me	2+	C_{68} H_{103} N_{18} O_{24} S_1	794.359033	794.358949	-0.11
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.435441	0.30
z_6	1+	C_{35} H_{49} N_7 O_{13} S_1	807.310357	807.310613	0.32
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329297	0.26
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430633	0.63
c_7+Me	1+	C_{40} H_{61} N_{10} O_{10}	841.456666	841.456959	0.35
z_{14}+Me	1+	C_{74} H_{112} N_{18} O_{25} S_1	842.891702	842.892216	0.61
z_7-C_3H_6S	1+	C_{36} H_{49} N_9 O_{15}	847.334267	847.334503	0.28
c_{14}+Me	2+	C_{77} H_{112} N_{19} O_{23} S_1	851.898324	851.898949	0.73
z_7-CO_2	1+	C_{38} H_{55} N_9 O_{12} S_1	877.363458	877.36382	0.41
c_{15}+Me	2+	C_{82} H_{119} N_{20} O_{26} S_1	916.419619	916.419752	0.15
z_7	1+	C_{39} H_{55} N_9 O_{15} S_1	921.353288	921.353668	0.41
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.499135	0.66
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.525366	0.31
z_8-C_3H_6S	1+	C_{40} H_{62} N_{10} O_{17} S_1	948.381946	948.382353	0.43
VTNWDDME-CO	1+	C_{41} H_{59} N_{10} O_{15} S_1	963.387658	963.388112	0.47
M+2H	2+	C_{88} H_{130} N_{21} O_{28} S_1	980.959108	980.959565	0.47
b_9+Me-H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535893	0.25
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.393022	-0.11
z_8	1+	C_{43} H_{62} N_{10} O_{17} S_1	1022.400963	1022.402793	1.79
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.547555	1.31
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.420377	0.67
c_9+Me	1+	C_{49} H_{77} N_{12} O_{13}	1041.572757	1041.572799	0.04
z_9	1+	C_{48} H_{71} N_{11} O_{18} S_1	1121.469377	1121.470662	1.15
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.48903	0.83
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.590541	1.23
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615687	1155.61621	0.45
z_{10}-CO_2	1+	C_{53} H_{82} N_{12} O_{17} S_1	1190.563615	1190.564609	0.83
z_{10}	1+	C_{54} H_{82} N_{12} O_{19} S_1	1234.553445	1234.554359	0.74
z_{11}-CO_2	1+	C_{55} H_{85} N_{13} O_{18} S_1	1247.585079	1247.585909	0.67
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.573004	0.67

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_{11}	1+	C_{56} H_{85} N_{13} O_{20} S_1	1291.574904	1291.576298	1.08
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.595466	1.40
c_{11}+Me	1+	C_{64} H_{93} N_{16} O_{16}	1341.694997	1341.696397	1.04
z_{12}-CO_2+Me	1+	C_{62} H_{94} N_{16} O_{19} S_1	1398.659641	1398.660457	0.58
z_{12}+Me	1+	C_{63} H_{94} N_{16} O_{21} S_1	1442.649466	1442.650464	0.69
c_{12}+Me	1+	C_{68} H_{98} N_{17} O_{19}	1456.721944	1456.722134	0.13
c_{13}+Me	1+	C_{72} H_{103} N_{18} O_{22}	1571.748884	1571.750554	1.06
c_{14}+Me	1+	C_{77} H_{112} N_{19} O_{23} S_1	1702.789373	1702.788642	-0.43
Average error					0.43
Absolute average error					0.45
Standard deviation					0.40

Table S3. 13 Peak assignment table for the ECD MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
VT	1+	$C_9 H_{17} N_2 O_3$	201.123369	201.12337	0.00
z_2	1+	$C_{11} H_{20} N_2 O_5$	260.136678	260.136673	-0.03
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155425	0.12
c_2	1+	$C_{14} H_{20} N_3 O_3$	278.149918	278.149948	0.12
z_3-C_3H_6S	1+	$C_{13} H_{23} N_3 O_6$	317.158138	317.158166	0.10
z_3	1+	$C_{16} H_{29} N_3 O_6 S_1$	391.177159	391.177158	0.00
c_3	1+	$C_{20} H_{31} N_4 O_4$	391.233982	391.233994	0.02
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.195882	407.195896	0.04
z_4-CO_2	1+	$C_{19} H_{34} N_4 O_7 S_1$	462.214273	462.214286	0.04
z_4	1+	$C_{20} H_{34} N_4 O_9 S_1$	506.204101	506.204079	-0.04
c_4	1+	$C_{25} H_{38} N_5 O_7$	520.276575	520.276614	0.07
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.222827	522.222895	0.14
z_5-CO_2	1+	$C_{23} H_{39} N_5 O_{10} S_1$	577.241217	577.241296	0.14
z_5	1+	$C_{24} H_{39} N_5 O_{12} S_1$	621.231044	621.231051	0.01
a_5+Me	1+	$C_{31} H_{44} N_7 O_7$	626.329673	626.329997	0.52
y_5	1+	$C_{24} H_{41} N_6 O_{12} S_1$	637.249771	637.249785	0.01
M+H	3+	$C_{88} H_{132} N_{21} O_{28} S_1$	654.308498	654.308157	-0.52
c_5+Me	1+	$C_{32} H_{47} N_8 O_8$	671.351138	671.351102	-0.06
c_6+Me	1+	$C_{34} H_{50} N_9 O_9$	728.372602	728.372557	-0.06
z_6-C_3H_6S	1+	$C_{32} H_{43} N_7 O_{13}$	733.291339	733.291416	0.11
z_6-CO_2	1+	$C_{34} H_{49} N_7 O_{11} S_1$	763.32053	763.320576	0.07
z_7-C_9H_7N	1+	$C_{30} H_{48} N_8 O_{15} S_1$	792.295439	792.29566	0.28
y_{13}+Me	2+	$C_{68} H_{103} N_{18} O_{24} S_1$	794.359033	794.359363	0.41
a_7+Me	1+	$C_{39} H_{58} N_9 O_9$	796.435201	796.435472	0.34
z_6	1+	$C_{35} H_{49} N_7 O_{13} S_1$	807.310357	807.310397	0.05
y_6	1+	$C_{35} H_{51} N_8 O_{13} S_1$	823.329081	823.32907	-0.01
b_7+Me	1+	$C_{40} H_{58} N_9 O_{10}$	824.430116	824.430327	0.26
c_7+Me	1+	$C_{40} H_{61} N_{10} O_{10}$	841.456666	841.456663	-0.01
z_{14}+Me	1+	$C_{74} H_{112} N_{18} O_{25} S_1$	842.891702	842.891945	0.29
z_7-C_3H_6S	1+	$C_{36} H_{49} N_9 O_{15}$	847.334267	847.334228	-0.04
c_{14}+Me	2+	$C_{77} H_{112} N_{19} O_{23} S_1$	851.898324	851.898568	0.29
z_7-CO_2	1+	$C_{38} H_{55} N_9 O_{12} S_1$	877.363458	877.363478	0.03
c_{15}+Me	2+	$C_{82} H_{119} N_{20} O_{26} S_1$	916.419619	916.41932	-0.33
z_7	1+	$C_{39} H_{55} N_9 O_{15} S_1$	921.353288	921.353292	0.00
b_8+Me	1+	$C_{45} H_{67} N_{10} O_{11}$	923.498529	923.498841	0.34
c_8+Me	1+	$C_{45} H_{70} N_{11} O_{11}$	940.525079	940.524966	-0.12
z_8-C_3H_6S	1+	$C_{40} H_{62} N_{10} O_{17} S_1$	948.381946	948.381936	-0.01

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
VTNWDDME-CO	1+	C_{41} H_{59} N_{10} O_{15} S_1	963.387658	963.387889	0.24
M+$_2$H	2+	C_{88} H_{130} N_{21} O_{28} S_1	980.959108	980.95889	-0.22
b_9+Me-H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.536283	0.63
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.392927	-0.20
z_8	1+	C_{43} H_{62} N_{10} O_{17} S_1	1022.400963	1022.402086	1.10
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546513	0.29
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419879	0.20
c_9+Me	1+	C_{49} H_{77} N_{12} O_{13}	1041.572757	1041.572208	-0.53
z_9	1+	C_{48} H_{71} N_{11} O_{18} S_1	1121.469377	1121.469918	0.48
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.488278	0.17
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.589655	0.45
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615687	1155.615393	-0.26
z_{10}-CO_2	1+	C_{53} H_{82} N_{12} O_{17} S_1	1190.563615	1190.563754	0.11
z_{10}	1+	C_{54} H_{82} N_{12} O_{19} S_1	1234.553445	1234.553403	-0.04
z_{11}-CO_2	1+	C_{55} H_{85} N_{13} O_{18} S_1	1247.585079	1247.585448	0.30
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.572134	-0.03
z_{11}	1+	C_{56} H_{85} N_{13} O_{20} S_1	1291.574904	1291.575752	0.66
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.594565	0.71
c_{11}+Me	1+	C_{64} H_{93} N_{16} O_{16}	1341.694997	1341.695088	0.07
z_{12}-CO_2+Me	1+	C_{62} H_{94} N_{16} O_{19} S_1	1398.659641	1398.659378	-0.19
z_{12}+Me	1+	C_{63} H_{94} N_{16} O_{21} S_1	1442.649466	1442.649551	0.06
c_{12}+Me	1+	C_{68} H_{98} N_{17} O_{19}	1456.721944	1456.720666	-0.87
c_{13}+Me	1+	C_{72} H_{103} N_{18} O_{22}	1571.748884	1571.74889	0.00
c_{14}+Me	1+	C_{77} H_{112} N_{19} O_{23} S_1	1702.78937	1702.78731	-1.21

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
			3	7	
Average error					0.07
Absolute average error					0.23
Standard deviation					0.27

Table S3. 14 Peak assignment table for the EID MS/MS spectrum of the synthetic peptide with the sequence $[YPIEH(\tau\text{-Me})GIVTNWDDMEK +3H]^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
VT	1+	$C_9 H_{17} N_2 O_3$	201.123369	201.123381	0.06
PI	1+	$C_{11} H_{19} N_2 O_2$	211.144104	211.14413	0.12
IE	1+	$C_{11} H_{19} N_2 O_4$	243.133934	243.133965	0.13
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144869	0.14
z_2	1+	$C_{11} H_{20} N_2 O_5$	260.136678	260.136678	0.00
EH(Methyl)-H_2O	1+	$C_{12} H_{15} N_4 O_3$	263.113867	263.113901	0.13
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155412	0.05
c_2	1+	$C_{14} H_{20} N_3 O_3$	278.149918	278.149926	0.03
EH(Methyl)	1+	$C_{12} H_{17} N_4 O_4$	281.124431	281.12448	0.17
z_3-C_3H_6S	1+	$C_{13} H_{23} N_3 O_6$	317.158138	317.158172	0.11
EH(Methyl)G-H_2O	1+	$C_{14} H_{18} N_5 O_4$	320.135331	320.13541	0.25
PIE	1+	$C_{16} H_{26} N_3 O_5$	340.186697	340.186705	0.02
z_3	1+	$C_{16} H_{29} N_3 O_6 S_1$	391.177159	391.177171	0.03
c_3	1+	$C_{20} H_{31} N_4 O_4$	391.233982	391.23399	0.02
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.195882	407.195894	0.03
H(Methyl)GIV	1+	$C_{20} H_{33} N_6 O_4$	421.25578	421.255844	0.15
IEH(Methyl)G-H_2O	1+	$C_{20} H_{29} N_6 O_5$	433.219394	433.219525	0.30
z_4-CO_2	1+	$C_{19} H_{34} N_4 O_7 S_1$	462.214273	462.214384	0.24
PIEH(Methyl)	1+	$C_{23} H_{35} N_6 O_6$	491.261259	491.261318	0.12
H(Methyl)GIVT-H_2O	1+	$C_{24} H_{38} N_7 O_5$	504.292894	504.293019	0.25
z_4	1+	$C_{20} H_{34} N_4 O_9 S_1$	506.204101	506.204023	-0.15
c_4	1+	$C_{25} H_{38} N_5 O_7$	520.276575	520.276535	-0.08
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.222827	522.22283	0.01
EH(Methyl)GIV-H_2O	1+	$C_{25} H_{38} N_7 O_6$	532.287808	532.287954	0.27
EH(Methyl)GIV	1+	$C_{25} H_{40} N_7 O_7$	550.298373	550.298696	0.59
z_5-CO_2	1+	$C_{23} H_{39} N_5 O_{10} S_1$	577.241217	577.241103	-0.20

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_5	1+	C_{24} H_{39} N_5 O_{12} S_1	621.231044	621.230943	-0.16
a_5+Me	1+	C_{31} H_{44} N_7 O_7	626.329673	626.329627	-0.07
y_5	1+	C_{24} H_{41} N_6 O_{12} S_1	637.249771	637.249828	0.09
M+3H	3+	C_{88} H_{132} N_{21} O_{28} S_1	654.308498	654.307111	-2.12
c_5+Me	1+	C_{32} H_{47} N_8 O_8	671.351138	671.351186	0.07
IVTNWD-H_2O	1+	C_{34} H_{47} N_8 O_9	711.346052	711.346249	0.28
c_6+Me	1+	C_{34} H_{50} N_9 O_9	728.372602	728.372588	-0.02
z_6-C_3H_6S $_1$+	1+	C_{32} H_{43} N_7 O_{13}	733.291339	733.291327	-0.02
z_7-C_9H_7N	1+	C_{30} H_{48} N_8 O_{15} S_1	792.295439	792.295349	-0.11
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.43525	0.06
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.419898	0.43
z_6	1+	C_{35} H_{49} N_7 O_{13} S_1	807.310357	807.310517	0.20
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329109	0.03
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.429993	-0.15
c_7+Me	1+	C_{40} H_{61} N_{10} O_{10}	841.456666	841.456721	0.07
z_7-C_3H_6S	1+	C_{36} H_{49} N_9 O_{15}	847.334267	847.334272	0.01
z_7-CO_2	1+	C_{38} H_{55} N_9 O_{12} S_1	877.363458	877.363417	-0.05
c_{15}+Me	2+	C_{82} H_{119} N_{20} O_{26} S_1	916.419619	916.419199	-0.46
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.345459	920.345255	-0.22
z_7	1+	C_{39} H_{55} N_9 O_{15} S_1	921.353288	921.353157	-0.14
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.498484	-0.05
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.52506	-0.02
VTNWDDME-CO	1+	C_{41} H_{59} N_{10} O_{15} S_1	963.387658	963.387841	0.19
M+H-H_2O	2+	C_{88} H_{128} N_{21} O_{27} S_1	971.953826	971.953183	-0.66
M+H	2+	C_{88} H_{130} N_{21} O_{28} S_1	980.959108	980.958428	-0.69
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535648	0.00
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.393274	0.13
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.545998	-0.20
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419214	-0.45
c_9+Me	1+	C_{49} H_{77} N_{12} O_{13}	1041.572757	1041.57257	-0.18
z_9	1+	C_{48} H_{71} N_{11} O_{18} S_1	1121.469377	1121.468447	-0.83
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.561721	-0.77
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.487905	-0.16
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.588931	-0.18
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615687	1155.615605	-0.07
z_{10}-CO_2	1+	C_{53} H_{82} N_{12} O_{17} S_1	1190.563615	1190.563945	0.28

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_{10}	1+	C_{54} H_{82} N_{12} O_{19} S_1	1234.553445	1234.552382	-0.86
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.571757	-0.33
z_{11}	1+	C_{56} H_{85} N_{13} O_{20} S_1	1291.574904	1291.574644	-0.20
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.592854	-0.59
c_{11}+Me	1+	C_{64} H_{93} N_{16} O_{16}	1341.694997	1341.694643	-0.26
z_{12}+Me	1+	C_{63} H_{94} N_{16} O_{21} S_1	1442.649466	1442.650279	0.56
c_{12}+Me	1+	C_{68} H_{98} N_{17} O_{19}	1456.721944	1456.721231	-0.49
c_{13}+Me	1+	C_{72} H_{103} N_{18} O_{22}	1571.748884	1571.747918	-0.61
c_{14}+Me	1+	C_{77} H_{112} N_{19} O_{23} S_1	1702.789373	1702.789408	0.02
Average error					-0.08
Absolute average error					0.24
Standard deviation					0.31

Table S3. 15 Peak assignment table for the EID MS/MS spectrum of the synthetic peptide with the sequence [YPIEH(π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
VT	1+	$C_9 H_{17} N_2 O_3$	201.123369	201.123386	0.08
PI	1+	$C_{11} H_{19} N_2 O_2$	211.144104	211.144125	0.10
IE	1+	$C_{11} H_{19} N_2 O_4$	243.133934	243.134005	0.29
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144872	0.15
z_2	1+	$C_{11} H_{20} N_2 O_5$	260.136678	260.136678	0.00
EH(Methyl)-H_2O	1+	$C_{12} H_{15} N_4 O_3$	263.113867	263.113942	0.29
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155419	0.08
c_2	1+	$C_{14} H_{20} N_3 O_3$	278.149918	278.149941	0.08
EH(Methyl)	1+	$C_{12} H_{17} N_4 O_4$	281.124431	281.124498	0.24
z_3-C_3H_6S	1+	$C_{13} H_{23} N_3 O_6$	317.158138	317.158183	0.14
EH(Methyl)G-H_2O	1+	$C_{14} H_{18} N_5 O_4$	320.135331	320.135432	0.32
PIE	1+	$C_{16} H_{26} N_3 O_5$	340.186697	340.186742	0.13
z_3	1+	$C_{16} H_{29} N_3 O_6 S_1$	391.177159	391.177184	0.06
c_3	1+	$C_{20} H_{31} N_4 O_4$	391.233982	391.23402	0.10
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.195882	407.195916	0.08
H(Methyl)GIV	1+	$C_{20} H_{33} N_6 O_4$	421.25578	421.25589	0.26
IEH(Methyl)G-H_2O	1+	$C_{20} H_{29} N_6 O_5$	433.219394	433.219525	0.30
z_4-CO_2	1+	$C_{19} H_{34} N_4 O_7 S_1$	462.214273	462.214257	-0.03
PIEH(Methyl)	1+	$C_{23} H_{35} N_6 O_6$	491.261259	491.261385	0.26
H(Methyl)GIVT-H_2O	1+	$C_{24} H_{38} N_7 O_5$	504.292894	504.293055	0.32
z_4	1+	$C_{20} H_{34} N_4 O_9 S_1$	506.204101	506.204097	-0.01
c_4	1+	$C_{25} H_{38} N_5 O_7$	520.276575	520.276576	0.00
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.222827	522.222829	0.00
EH(Methyl)GIV-H_2O	1+	$C_{25} H_{38} N_7 O_6$	532.287808	532.288041	0.44
EH(Methyl)GIV	1+	$C_{25} H_{40} N_7 O_7$	550.298373	550.298545	0.31
z_5-CO_2	1+	$C_{23} H_{39} N_5 O_{10} S_1$	577.241217	577.241096	-0.21
z_5	1+	$C_{24} H_{39} N_5 O_{12} S_1$	621.231044	621.231083	0.06
a_5+Me	1+	$C_{31} H_{44} N_7 O_7$	626.329673	626.329792	0.19
y_5	1+	$C_{24} H_{41} N_6 O_{12} S_1$	637.249771	637.249821	0.08
M+H	3+	$C_{88} H_{132} N_{21} O_{28} S_1$	654.308498	654.307089	-2.15
c_5+Me	1+	$C_{32} H_{47} N_8 O_8$	671.351138	671.351193	0.08
IVTNWD-H_2O	1+	$C_{34} H_{47} N_8 O_9$	711.346052	711.346289	0.33
c_6+Me	1+	$C_{34} H_{50} N_9 O_9$	728.372602	728.372595	-0.01
z_6-C_3H_6S	1+	$C_{32} H_{43} N_7 O_{13}$	733.291339	733.29123	-0.15

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
z_7-C_9H_7N	1+	C_{30} H_{48} N_8 O_{15} S_1	792.295439	792.295452	0.02
a_7+Me	1+	C_{39} H_{58} N_9 O_9	796.435201	796.43535	0.19
b_7+Me -H_2O	1+	C_{40} H_{56} N_9 O_9	806.419551	806.419783	0.29
z_6	1+	C_{35} H_{49} N_7 O_{13} S_1	807.310357	807.310236	-0.15
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.329146	0.08
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.430133	0.02
c_7+Me	1+	C_{40} H_{61} N_{10} O_{10}	841.456666	841.456701	0.04
z_7-C_3H_6S	1+	C_{36} H_{49} N_9 O_{15}	847.334267	847.334276	0.01
z_7-CO_2	1+	C_{38} H_{55} N_9 O_{12} S_1	877.363458	877.363632	0.20
c_{15}+Me	2+	C_{82} H_{119} N_{20} O_{26} S_1	916.419619	916.419245	-0.41
y_7-NH_3	1+	C_{39} H_{54} N_9 O_{15} S_1	920.345459	920.345504	0.05
z_7	1+	C_{39} H_{55} N_9 O_{15} S_1	921.353288	921.353484	0.21
c_8+Me	1+	C_{45} H_{70} N_{11} O_{11}	940.525079	940.525054	-0.03
VTNWDDME-CO	1+	C_{41} H_{59} N_{10} O_{15} S_1	963.387658	963.387677	0.02
M+H-H_2O	2+	C_{88} H_{128} N_{21} O_{27} S_1	971.953826	971.953425	-0.41
M+H	2+	C_{88} H_{130} N_{21} O_{28} S_1	980.959108	980.958715	-0.40
b_9+Me -H_2O	1+	C_{49} H_{72} N_{11} O_{12}	1006.535643	1006.535708	0.06
y_8-NH_3	1+	C_{43} H_{61} N_{10} O_{17} S_1	1021.393138	1021.393295	0.15
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.546282	0.07
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419752	0.07
c_9+Me	1+	C_{49} H_{77} N_{12} O_{13}	1041.572757	1041.572596	-0.15
z_9	1+	C_{48} H_{71} N_{11} O_{18} S_1	1121.469377	1121.469085	-0.26
b_{10}+Me -NH_3	1+	C_{53} H_{77} N_{12} O_{15}	1121.562586	1121.56157	-0.91
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.487944	-0.13
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.589109	-0.02
c_{10}+Me	1+	C_{53} H_{83} N_{14} O_{15}	1155.615687	1155.61567	-0.01
z_{10}-CO_2	1+	C_{53} H_{82} N_{12} O_{17} S_1	1190.563615	1190.56367	0.05
z_{10}	1+	C_{54} H_{82} N_{12} O_{19} S_1	1234.553445	1234.553447	0.00
y_{10}	1+	C_{54} H_{84} N_{13} O_{19} S_1	1250.572165	1250.572073	-0.07
z_{11}	1+	C_{56} H_{85} N_{13} O_{20} S_1	1291.57490	1291.57162	-2.54

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
			4		
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593255	-0.29
c_{11}+Me	1+	C_{64} H_{93} N_{16} O_{16}	1341.694997	1341.694531	-0.35
z_{12}+Me	1+	C_{63} H_{94} N_{16} O_{21} S_1	1442.649466	1442.649802	0.23
c_{12}+Me	1+	C_{68} H_{98} N_{17} O_{19}	1456.721944	1456.721684	-0.18
c_{13}+Me	1+	C_{72} H_{103} N_{18} O_{22}	1571.748884	1571.748052	-0.53
c_{14}+Me	1+	C_{77} H_{112} N_{19} O_{23} S_1	1702.789373	1702.789403	0.02
Average error					-0.04
Absolute average error					0.23
Standard deviation					0.40

Table S3. 16 Peak assignment table for the CAD MS/MS spectrum of the tryptic peptide from bovine actin with the sequence [YPIEH(τ/π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN	1+	$C_8 H_{14} N_3 O_4$	216.097882	216.097882	0.00
a_2	1+	$C_{13} H_{17} N_2 O_2$	233.128454	233.128491	0.16
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144746	-0.34
b_2	1+	$C_{14} H_{17} N_2 O_3$	261.1233705	261.123293	-0.30
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155374	-0.08
b_5+Me	2+	$C_{32} H_{44} N_7 O_8$	327.665932	327.665914	-0.05
b_6+Me	2+	$C_{34} H_{47} N_8 O_9$	356.176386	356.17647	0.24
a_7/b_8-CO	2+	$C_{39} H_{58} N_9 O_9$	398.721239	398.721704	1.17
y_6-NH_3	2+	$C_{35} H_{48} N_7 O_{13} S_1$	403.654904	403.654834	-0.17
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.1958742	407.195872	-0.01
y_6	2+	$C_{35} H_{51} N_8 O_{13} S_1$	412.1681718	412.16773	-1.07
b_7+Me	2+	$C_{40} H_{58} N_9 O_{10}$	412.718696	412.718715	0.05
EH(Methyl)GI	1+	$C_{20} H_{31} N_6 O_6$	451.229959	451.22955	-0.91
a_8+Me	2+	$C_{45} H_{69} N_{10} O_{10}$	455.263271	455.263717	0.98
y_7-NH_3	2+	$C_{39} H_{54} N_9 O_{15} S_1$	460.676368	460.675837	-1.15
y_7	2+	$C_{39} H_{58} N_{10} O_{15} S_1$	469.189643	469.18926	-0.82
b_8+Me	2+	$C_{46}H_{70}N_{10}O_{11}$	469.260729	469.260318	-0.88
PIEH(Methyl)	1+	$C_{23} H_{35} N_6 O_6$	491.261259	491.261075	-0.37
y_8-NH_3	2+	$C_{43} H_{61} N_{10} O_{17} S_1$	511.200207	511.200388	0.35
y_8	2+	$C_{43} H_{64} N_{11} O_{17} S_1$	519.713482	519.713483	0.00
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.2228142	522.223552	1.41
EH(Methyl)GII	1+	$C_{26} H_{42} N_7 O_7$	564.314023	564.313605	-0.74
b_{10}+Me	2+	$C_{54}H_{83}N_{13}O_{15}$	576.806032	576.805378	-1.13
y_{15}-NH_3	3+	$C_{80} H_{120} N_{19} O_{26} S_1$	598.950422	598.95055	0.21
y_{15}+Me	3+	$C_{80}H_{125}N_{20}O_{26}S_1$	604.62594	604.625927	-0.02
y_5	1+	$C_{24} H_{41} N_6 O_{12} S_1$	637.2497542	637.250649	1.40
M+$_3$H-H_2O	3+	$C_{89}H_{132}N_{21}O_{27}S_1$	652.976861	652.976677	-0.28
M+$_3$H	3+	$C_{89}H_{134}N_{21}O_{28}S_1$	658.980383	658.980804	0.64
PIEH(Methyl)GI	1+	$C_{31} H_{49} N_8 O_8$	661.366787	661.367427	0.97
b_6+Me	1+	$C_{34} H_{47} N_8 O_9$	711.3461052	711.344565	-2.17
y_{12}+Me	2+	$C_{64} H_{98} N_{17} O_{21} S_1$	736.845558	736.843823	-2.35
PIEH(Methyl)GII	1+	$C_{37} H_{60} N_9 O_9$	774.450851	774.451381	0.68
y_6	1+	$C_{35}H_{51}N_8O_{13}S_1$	823.329084	823.330823	2.11
b_7+Me	1+	$C_{40}H_{58}N_9O_{10}$	824.430117	824.431836	2.09
y_{14}+Me	2+	$C_{75}H_{117}N_{19}O_{25}S_1$	857.908889	857.908909	0.02
y_{15}+Me	2+	$C_{80} H_{123} N_{20} O_{26} S_1$	906.435269	906.433823	-1.60
y_7	1+	$C_{39}H_{57}N_{10}O_{15}S_1$	937.372012	937.374074	2.20
b_8+Me	1+	$C_{46}H_{69}N_{10}O_{11}$	937.514181	937.513597	-0.62

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_8	1+	$C_{43}H_{64}N_{11}O_{17}S_1$	1038.419691	1038.421612	1.85
y_9	1+	$C_{49}H_{75}N_{12}O_{18}S_1$	1151.503751	1151.502296	-1.26
y_{11}	1+	$C_{57}H_{89}N_{14}O_{20}S_1$	1321.609283	1321.609251	-0.02
Average error					0.00
Absolute average error					0.80
Standard deviation					0.72

Table S3. 17 Peak assignment table for the CAD MS/MS spectrum of the tryptic peptide from chicken actin with the sequence [YPIEH(τ/π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN	1+	$C_8H_{14}N_3O_4$	216.097882	216.097882	0.00
a_2	1+	$C_{13}H_{17}N_2O_2$	233.128454	233.128437	-0.07
y_2-H_2O	1+	$C_{11}H_{20}N_3O_4$	258.144833	258.144685	-0.57
b_2	1+	$C_{14}H_{17}N_2O_3$	261.1233705	261.123204	-0.64
y_2	1+	$C_{11}H_{22}N_3O_5$	276.155397	276.155296	-0.37
b_5+Me	2+	$C_{32}H_{44}N_7O_8$	327.665932	327.665757	-0.53
b_6+Me	2+	$C_{34}H_{47}N_8O_9$	356.176386	356.17616	-0.63
a_7+Me/b_8+Me-CO	2+	$C_{39}H_{58}N_9O_9$	398.721239	398.721328	0.22
y_6-NH_3	2+	$C_{35}H_{48}N_7O_{13}S_1$	403.654904	403.65431	-1.47
y_3	1+	$C_{16}H_{31}N_4O_6S_1$	407.1958742	407.195526	-0.86
y_6	2+	$C_{35}H_{51}N_8O_{13}S_1$	412.1681718	412.16714	-2.50
b_7+Me	2+	$C_{40}H_{58}N_9O_{10}$	412.718696	412.718362	-0.81
EH(Methyl)GI	1+	$C_{20}H_{31}N_6O_6$	451.229959	451.229184	-1.72
a_8+Me	2+	$C_{45}H_{69}N_{10}O_{10}$	455.263271	455.263272	0.00
y_7-NH_3	2+	$C_{39}H_{54}N_9O_{15}S_1$	460.676368	460.675797	-1.24
y_7	2+	$C_{39}H_{58}N_{10}O_{15}S_1$	469.189643	469.189165	-1.02
b_8+Me	2+	$C_{46}H_{70}N_{10}O_{11}$	469.260729	469.26001	-1.53
PIEH(Methyl)	1+	$C_{23}H_{35}N_6O_6$	491.261259	491.260734	-1.07
y_8-NH_3	2+	$C_{43}H_{61}N_{10}O_{17}S_1$	511.200207	511.199922	-0.56
y_8	2+	$C_{43}H_{64}N_{11}O_{17}S_1$	519.713482	519.713218	-0.51
y_4	1+	$C_{20}H_{36}N_5O_9S_1$	522.2228142	522.222968	0.29
EH(Methyl)GII	1+	$C_{26}H_{42}N_7O_7$	564.314023	564.313004	-1.81
b_{10}+Me	2+	$C_{54}H_{83}N_{13}O_{15}$	576.806032	576.804836	-2.07
y_{15}-NH_3	3+	$C_{80}H_{120}N_{19}O_{26}S_1$	598.950422	598.950121	-0.50
y_{15}+Me	3+	$C_{80}H_{125}N_{20}O_{26}S_1$	604.62594	604.625059	-1.46
y_5	1+	$C_{24}H_{41}N_6O_{12}S_1$	637.2497542	637.249914	0.25
b_{16}+Me	2+	$C_{89}H_{132}N_{21}O_{27}S_1$	652.976861	652.976786	-0.11

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
$M+_3H$	3+	$C_{89}H_{134}N_{21}O_{28}S_1$	658.980383	658.980443	0.09
PIEH(Methyl)GI	1+	$C_{31}\ H_{49}\ N_8\ O_8$	661.366787	661.367047	0.39
b_6+Me	1+	$C_{34}\ H_{47}\ N_8\ O_9$	711.3461052	711.344158	-2.74
y_{12}+Me	2+	$C_{64}\ H_{98}\ N_{17}\ O_{21}\ S_1$	736.845558	736.843613	-2.64
PIEH(Methyl)GII	1+	$C_{37}\ H_{60}\ N_9\ O_9$	774.450851	774.450844	-0.01
y_6	1+	$C_{35}H_{51}N_8O_{13}S_1$	823.329084	823.330012	1.13
b_7+Me	1+	$C_{40}H_{58}N_9O_{10}$	824.430117	824.431474	1.65
y_{14}+Me	2+	$C_{75}H_{117}N_{19}O_{25}S_1$	857.908889	857.909101	0.25
y_{15}+Me	2+	$C_{80}\ H_{123}\ N_{20}\ O_{26}\ S_1$	906.435269	906.432991	-2.51
y_7	1+	$C_{39}H_{57}N_{10}O_{15}S_1$	937.372012	937.373883	2.00
b_8+Me	1+	$C_{46}H_{69}N_{10}O_{11}$	937.514181	937.51297	-1.29
y_8	1+	$C_{43}H_{64}N_{11}O_{17}S_1$	1038.419691	1038.419834	0.14
y_9	1+	$C_{49}\ H_{75}\ N_{12}\ O_{18}\ S_1$	1151.503751	1151.50227	-1.29
y_{11}	1+	$C_{57}H_{89}N_{14}O_{20}S_1$	1321.609283	1321.609295	0.01
Average error					-0.64
Absolute average error					0.95
Standard deviation					0.81

Table S3. 18 Peak assignment table for the CAD MS/MS spectrum of the tryptic peptide from rabbit actin with the sequence $[YPIEH(\tau/\pi\text{-Me})GIVTNWDDMEK +3H]^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN	1+	$C_8\ H_{14}\ N_3\ O_4$	216.097882	216.097882	0.00
a_2	1+	$C_{13}\ H_{17}\ N_2\ O_2$	233.128454	233.128412	-0.18
y_2-H_2O	1+	$C_{11}\ H_{20}\ N_3\ O_4$	258.144833	258.144595	-0.92
b_2	1+	$C_{14}\ H_{17}\ N_2\ O_3$	261.1233705	261.123139	-0.89
y_2	1+	$C_{11}\ H_{22}\ N_3\ O_5$	276.155397	276.155205	-0.70
b_5	2+	$C_{32}\ H_{44}\ N_7\ O_8$	327.665932	327.665581	-1.07
b_6	2+	$C_{34}\ H_{47}\ N_8\ O_9$	356.176386	356.176048	-0.95
a_7/b_8-CO	2+	$C_{39}\ H_{58}\ N_9\ O_9$	398.721239	398.721218	-0.05
y_6-NH_3	2+	$C_{35}\ H_{48}\ N_7\ O_{13}\ S_1$	403.654904	403.654268	-1.58
y_3	1+	$C_{16}\ H_{31}\ N_4\ O_6\ S_1$	407.1958742	407.19534	-1.31
y_6	2+	$C_{35}\ H_{51}\ N_8\ O_{13}\ S_1$	412.1681718	412.167736	-1.06
b_7+Me	2+	$C_{40}\ H_{58}\ N_9\ O_{10}$	412.718696	412.71825	-1.08
EH(Methyl)GI	1+	$C_{20}\ H_{31}\ N_6\ O_6$	451.229959	451.229226	-1.62
a_8+Me	2+	$C_{45}\ H_{69}\ N_{10}\ O_{10}$	455.263271	455.263124	-0.32
y_7-NH_3	2+	$C_{39}\ H_{54}\ N_9\ O_{15}\ S_1$	460.676368	460.675613	-1.64
y_7	2+	$C_{39}\ H_{58}\ N_{10}\ O_{15}\ S_1$	469.189643	469.188853	-1.68
b_8+Me	2+	$C_{46}H_{70}N_{10}O_{11}$	469.260729	469.259706	-2.18

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
PIEH(Methyl)	1+	C_{23} H_{35} N_6 O_6	491.261259	491.260426	-1.70
y_8-NH_3	2+	C_{43} H_{61} N_{10} O_{17} S_1	511.200207	511.199622	-1.14
y_8	2+	C_{43} H_{64} N_{11} O_{17} S_1	519.713482	519.712887	-1.14
y_4	1+	C_{20} H_{36} N_5 O_9 S_1	522.2228142	522.222797	-0.03
EH(Methyl)GII	1+	C_{26} H_{42} N_7 O_7	564.314023	564.312924	-1.95
b_{10}+Me	2+	$C_{54}H_{83}N_{13}O_{15}$	576.806032	576.804549	-2.57
y_{15}+Me	3+	$C_{80}H_{125}N_{20}O_{26}S_1$	604.62594	604.624793	-1.90
y_5	3+	C_{24} H_{41} N_6 O_{12} S_1	637.2497542	637.249758	0.01
M+$_3$H-H_2O	3+	$C_{89}H_{132}N_{21}O_{27}S_1$	652.976861	652.976707	-0.24
M+$_3$H	3+	$C_{89}H_{134}N_{21}O_{28}S_1$	658.980383	658.980568	0.28
PIEH(Methyl)GI	3+	C_{31} H_{49} N_8 O_8	661.366787	661.366814	0.04
b_6+Me	1+	C_{34} H_{47} N_8 O_9	711.3461052	711.345408	-0.98
y_{12}+Me	1+	C_{64} H_{98} N_{17} O_{21} S_1	736.845558	736.844006	-2.11
PIEH(Methyl)GI I	2+	C_{37} H_{60} N_9 O_9	774.450851	774.450867	0.02
y_6	1+	$C_{35}H_{51}N_8O_{13}S_1$	823.329084	823.329994	1.11
b_7+Me	1+	$C_{40}H_{58}N_9O_{10}$	824.430117	824.431169	1.28
y_{14}+Me	1+	$C_{75}H_{117}N_{19}O_{25}S_1$	857.908889	857.909587	0.81
y_{15}+Me	2+	C_{80} H_{123} N_{20} O_{26} S_1	906.435269	906.433263	-2.21
y_7	2+	$C_{39}H_{57}N_{10}O_{15}S_1$	937.372012	937.374664	2.83
b_8+Me	1+	$C_{46}H_{69}N_{10}O_{11}$	937.514181	937.513082	-1.17
y_8	1+	$C_{43}H_{64}N_{11}O_{17}S_1$	1038.419691	1038.42061	0.88
y_9	1+	C_{49} H_{75} N_{12} O_{18} S_1	1151.503751	1151.501808	-1.69
y_{11}	1+	$C_{57}H_{89}N_{14}O_{20}S_1$	1321.609283	1321.609231	-0.04
Average error					-0.72
Absolute average error					1.08
Standard deviation					0.77

Table S3. 19 Peak assignment table for the CAD MS/MS spectrum of the tryptic peptide from human platelet actin with the sequence [YPIEH(τ/π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN	1+	$C_8 H_{14} N_3 O_4$	216.097882	216.097882	0.00
a_2	1+	$C_{13} H_{17} N_2 O_2$	233.128454	233.128483	0.12
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144803	-0.12
b_2	1+	$C_{14} H_{17} N_2 O_3$	261.1233705	261.123344	-0.10
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155331	-0.24
b_5+Me	2+	$C_{32} H_{44} N_7 O_8$	327.665932	327.66587	-0.19
b_6+Me	2+	$C_{34} H_{47} N_8 O_9$	356.176386	356.17658	0.54
a_7+Me/b_8+Me-CO	2+	$C_{39} H_{58} N_9 O_9$	398.721239	398.721435	0.49
y_6-NH_3	2+	$C_{35} H_{48} N_7 O_{13} S_1$	403.654904	403.654664	-0.59
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.1958742	407.195818	-0.14
y_6	2+	$C_{35} H_{51} N_8 O_{13} S_1$	412.1681718	412.167716	-1.11
b_7+Me	2+	$C_{40} H_{58} N_9 O_{10}$	412.718696	412.718563	-0.32
a_8+Me	2+	$C_{44} H_{67} N_{10} O_{10}$	448.255446	448.255509	0.14
y_7-NH_3	2+	$C_{39} H_{54} N_9 O_{15} S_1$	460.676368	460.676244	-0.27
b_8+Me	2+	$C_{45} H_{68} N_{10} O_{11}$	462.252903	462.252677	-0.49
y_7	2+	$C_{39} H_{58} N_{10} O_{15} S_1$	469.189643	469.189013	-1.34
a_9+Me/b_9+Me-CO	2+	$C_{48} H_{74} N_{11} O_{12}$	498.779285	498.778812	-0.95
b_9+Me-H_2O-CO	2+	$C_{48} H_{72} N_{11} O_{11}$	503.77146	503.771152	-0.61
y_8	2+	$C_{43} H_{64} N_{11} O_{17} S_1$	519.713482	519.713483	0.00
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.2228142	522.223344	1.01
y_9	2+	$C_{48} H_{73} N_{12} O_{18} S_1$	569.2476838	569.247062	-1.09
b_{10}+Me	2+	$C_{53} H_{80} N_{13} O_{15}$	569.798206	569.797544	-1.16
y_{15}+Me-NH_3	2+	$C_{79} H_{118} N_{19} O_{26} S_1$	594.278538	594.278348	-0.32
y_{15}+Me	3+	$C_{79} H_{121} N_{20} O_{26} S_1$	599.954055	599.953397	-1.10
y_5	1+	$C_{24} H_{41} N_6 O_{12} S_1$	637.2497542	637.250441	1.08
M+$_3$H-NH_3	3+	$C_{88} H_{127} N_{20} O_{28} S_1$	648.632981	648.634	1.57
M+$_3$H	3+	$C_{88} H_{132} N_{21} O_{28} S_1$	654.308498	654.309429	1.42
PIEH(Methyl)GI	1+	$C_{31} H_{49} N_8 O_8$	661.366787	661.367293	0.77
b_6+Me	1+	$C_{34} H_{47} N_8 O_9$	711.346052	711.344665	-1.95
y_{12}+Me	1+	$C_{63} H_{96} N_{17} O_{21} S_1$	729.837734	729.837714	-0.03
PIEH(Methyl)GI V	1+	$C_{36} H_{58} N_9 O_9$	760.435201	760.436635	1.89

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_6	1+	C_{35} H_{51} N_8 O_{13} S_1	823.329081	823.33027	1.44
b_7+Me	1+	C_{40} H_{58} N_9 O_{10}	824.430116	824.431024	1.10
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.900892	-0.20
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.427208	-0.26
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.497943	-0.63
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.373923	1.95
b_9+Me	1+	C_{49} H_{74} N_{11} O_{13}	1024.546208	1024.54601	-0.19
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.421644	1.89
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.486324	-1.55
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.590029	0.79
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593655	0.02
Average error					0.03
Absolute average error					0.74
Standard deviation					0.61

Table S3. 20 Peak assignment table for the CAD MS/MS spectrum of the tryptic peptide from human recombinant actin with the sequence [YPIEH(τ/π-Me)GIVTNWDDMEK +3H]$^{3+}$.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
TN	1+	$C_8 H_{14} N_3 O_4$	216.097882	216.097882	0.00
a_2	1+	$C_{13} H_{17} N_2 O_2$	233.128454	233.128426	-0.12
y_2-H_2O	1+	$C_{11} H_{20} N_3 O_4$	258.144833	258.144703	-0.50
b_2	1+	$C_{14} H_{17} N_2 O_3$	261.1233705	261.123181	-0.73
y_2	1+	$C_{11} H_{22} N_3 O_5$	276.155397	276.155282	-0.42
NW	1+	$C_{15} H_{17} N_4 O_3$	301.129517	301.129351	-0.55
b_5+Me	2+	$C_{32} H_{44} N_7 O_8$	327.665932	327.665811	-0.37
b_6+Me	2+	$C_{34} H_{47} N_8 O_9$	356.176386	356.176174	-0.60
a_7+Me/b_8+Me-CO	2+	$C_{39} H_{58} N_9 O_9$	398.721239	398.721385	0.37
y_6-NH_3	2+	$C_{35} H_{48} N_7 O_{13} S_1$	403.654904	403.654412	-1.22
y_3	1+	$C_{16} H_{31} N_4 O_6 S_1$	407.1958742	407.195596	-0.68
y_6	2+	$C_{35} H_{51} N_8 O_{13} S_1$	412.1681718	412.167339	-2.02
b_7+Me	2+	$C_{40} H_{58} N_9 O_{10}$	412.718696	412.718438	-0.63
a_8+Me	2+	$C_{44} H_{67} N_{10} O_{10}$	448.255446	448.255449	0.01
b_8+Me	2+	$C_{45} H_{68} N_{10} O_{11}$	462.252903	462.252445	-0.99
y_7	2+	$C_{39} H_{58} N_{10} O_{15} S_1$	469.189643	469.188951	-1.47
a_9+Me/b_9+Me-CO	2+	$C_{48} H_{74} N_{11} O_{12}$	498.779285	498.778705	-1.16
b_9+Me-H_2O-CO	2+	$C_{48} H_{72} N_{11} O_{11}$	503.77146	503.771078	-0.76
y_8-NH_3	2+	$C_{43} H_{61} N_{10} O_{17} S_1$	511.200207	511.19999	-0.42
y_8	2+	$C_{43} H_{64} N_{11} O_{17} S_1$	519.713482	519.713061	-0.81
y_4	1+	$C_{20} H_{36} N_5 O_9 S_1$	522.2228142	522.223148	0.64
b_{10}+Me-NH_3	2+	$C_{53} H_{77} N_{12} O_{15}$	561.284931	561.284688	-0.43
b_{10}+Me	2+	$C_{53} H_{80} N_{13} O_{15}$	569.798206	569.797109	-1.93
y_{15}+Me-H_2O	3+	$C_{79} H_{119} N_{20} O_{25} S_1$	593.950533	593.950701	0.28
y_5	1+	$C_{24} H_{41} N_6 O_{12} S_1$	637.2497542	637.250162	0.64
M+$_3$H-NH_3	3+	$C_{88} H_{127} N_{20} O_{28} S_1$	648.632981	648.634636	2.55
M+$_3$H	3+	$C_{88} H_{132} N_{21} O_{28} S_1$	654.308498	654.309578	1.65
PIEH(Methyl)GI	1+	$C_{31} H_{49} N_8 O_8$	661.366787	661.367421	0.96
b_6+Me	1+	$C_{34} H_{47} N_8 O_9$	711.346052	711.344128	-2.70
y_{12}+Me	2+	$C_{63} H_{96} N_{17} O_{21} S_1$	729.837734	729.83783	0.13
PIEH(Methyl)GIV	1+	$C_{36} H_{58} N_9 O_9$	760.435201	760.437183	2.61
y_6	1+	$C_{35} H_{51} N_8 O_{13} S_1$	823.329081	823.330419	1.63
b_7+Me	1+	$C_{40} H_{58} N_9 O_{10}$	824.430116	824.430948	1.01

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_{14}+Me	2+	C_{74} H_{114} N_{19} O_{25} S_1	850.901062	850.901036	-0.03
y_{15}+Me-NH_3	2+	C_{79} H_{118} N_{19} O_{26} S_1	890.914169	890.910849	-3.73
y_{15}+Me	2+	C_{79} H_{121} N_{20} O_{26} S_1	899.427444	899.427193	-0.28
b_8+Me	1+	C_{45} H_{67} N_{10} O_{11}	923.498529	923.497787	-0.80
y_7	1+	C_{39} H_{57} N_{10} O_{15} S_1	937.3720971	937.373878	1.90
y_8	1+	C_{43} H_{64} N_{11} O_{17} S_1	1038.419677	1038.419509	-0.16
y_9	1+	C_{48} H_{73} N_{12} O_{18} S_1	1137.488091	1137.486792	-1.14
b_{10}+Me	1+	C_{53} H_{80} N_{13} O_{15}	1138.589135	1138.587269	-1.64
y_{11}	1+	C_{56} H_{87} N_{14} O_{20} S_1	1307.593629	1307.593669	0.03
Average error					-0.28
Absolute average error					0.97
Standard deviation					0.85

4. Exploring Tandem Mass Spectrometry Methods for the Analysis of Dihydroxylated Vitamin D_3 Isomers

Vitamin D compounds are a group of steroids derived from cholesterol, which are vital for maintaining bone health in humans, especially the biologically active hormone 1,25-dihydroxyvitamin D_3, which can be difficult to distinguish from the inactive isomeric form, 24,25-dihydroxyvitamin D_3 as both metabolites have the exact same mass. Therefore, this chapter demonstrates the capabilities of tandem mass spectrometry methods to differentiate between the dihydroxylated vitamin D_3 isomers using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS).

Sample preparation, data acquisition and analysis results presented in this chapter were carried out by the thesis author. Vitamin D_3 metabolites were provided by Professor Dietrich A. Volmer and Pascal Schorr from the Volmer group at the Humboldt University of Berlin.

One manuscript entitled “Differentiation of dihydroxylated vitamin D_3 isomers using tandem mass spectrometry” by Anisha Haris, Yuko P. Y. Lam, Christopher A. Wootton, Alina Theisen, Bryan P. Marzullo, Pascal Schorr, Dietrich Volmer, and Peter B. O'Connor, was accepted for publication in The Journal of the American Society for Mass Spectrometry based on the results presented in this chapter.

4.1 Abstract

Recent studies have shown that the dihydroxylated vitamin D_3 compounds: 1,25-dihydroxyvitamin D_3 (active form) and 24,25-dihydroxyvitamin D_3 (inactive form) have significant biological effects, playing a role in diseases such as osteoporosis. Differentiation and characterization of the isomers by mass spectrometry can be challenging due to the zero-mass difference and minor structural differences between them.

Herein, we investigated the use of various fragmentation methods such as collisional activated dissociation (CAD), infrared multiphoton dissociation (IRMPD), electron induced dissociation (EID), and ultraviolet photodissociation (UVPD), available on a Bruker 12 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) to generate characteristic fragments for the dihydroxylated vitamin D_3 isomers that can be used to distinguish between them.

Isomer-specific fragments were observed and confirmed for the highly active 1,25-dihydroxyvitamin D_3, which were clearly absent in the 24,25-dihydroxyvitamin D_3 MS/MS spectra using all the mentioned fragmentation methods. Thus, enabling clear differentiation between the two dihydroxylated vitamin D_3 isomers, without the need for prior chromatographic separation or derivatization of the metabolites.

4.2 Introduction

Vitamin D refers to a group of fat-soluble organic compounds, often exhibiting certain biological activities. The two major forms of vitamin D are vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol) (Figure 4.1). Vitamin D_2 is found and formed naturally in UV-irradiated plants and fungi.[1] Humans are unable to generate vitamin D_2; instead it is extracted from the plants, which do make it and can be used, for example, to fortify foods such as cereal.[2] Vitamin D_3, on the other hand, is synthesised in the skin of mammals when exposed to UV radiation from the sun. Compared to plant-sourced vitamin D_2, dietary vitamin D_3 is generally only found in animal-sourced foods such as fatty fish and fish liver oil.[3] Although both vitamin D_2 and D_3 enter the bloodstream, supplementation with vitamin D_3 has been shown to be more effective at raising the circulating vitamin D concentrations in the serum compared to vitamin D_2.[4,5]

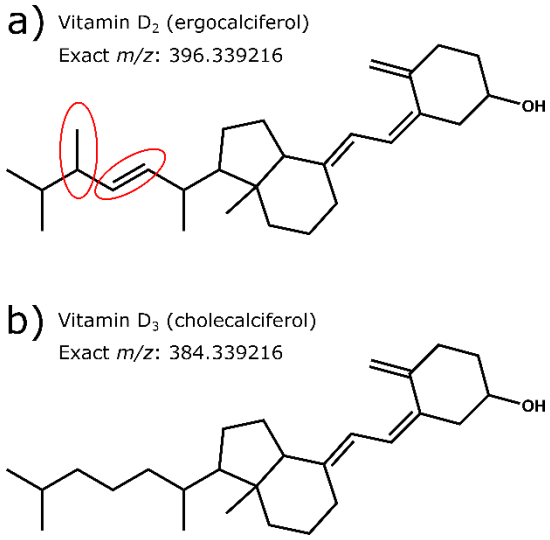


Figure 4. 1 Chemical structures of a) vitamin D_2 and b) vitamin D_3 with the differences between the structures encircled in red on vitamin D_2.

For over a century, it is well known that vitamin D_3 plays an essential role in the absorption and regulation of important minerals such as calcium and phosphate in the body.[6] These nutrients are needed to promote normal bone formation and mineralisation. Vitamin D_3 also provides support for normal muscle function and prevents calcium loss from the kidneys.[7-9] A severe lack of vitamin D_3, also known as vitamin D deficiency/insufficiency can lead to bone deformities, such as rickets in young children and bone pain in adults caused by osteomalacia.[10-13] The physical symptoms of both conditions include bowed legs, muscle weakness as well as soft and fragile bones, which often result in an increased tendency for bone fractures.[14] Vitamin D deficiency has also

been linked to various diseases of concern such as diabetes, heart disease, and neurological disorders, such as Alzheimer's disease and schizophrenia.[15-18]

A simplified diagram of the metabolic pathway for vitamin D_3, resulting in the formation of the dihydroxylated vitamin D_3 isomers of interest is provided in Figure 4.2. Vitamin D_3 (cholecalciferol) is made in the skin from the cholesterol precursor molecule, 7-dehydrocholesterol (7-DHC) under the influence of UV light (290-315 nm, UV_B) from the sun.[1,19] 7-DHC is found in the epidermis, the outer layer of the skin and is crucial for the synthesis of vitamin D_3 in humans and animals. Vitamin D_3 metabolises first to calcidiol, otherwise known as 25-hydroxyvitamin D_3 ($25OHD_3$) in the liver and is then further oxidized to the highly biologically active compound, calcitriol also called 1,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$).[19] During catabolism of vitamin D_3, 24,25-dihydroxyvitamin D_3 ($24,25(OH)_2D_3$) is also formed, which is an isomer of calcitriol and considered to be inactive as a hormone.[19] However, some studies showed that 24,25-dihydroxyvitamin D_3 may have some biological properties of its own.[20-23] For example, in 1982, Sömjen *et. al.*[24] found that 24,25-dihydroxyvitamin D_3 may play a role in the metabolism of developing skeletal tissues of new-born mice; while Seo *et. al.*[25] showed that increased levels of 24,25-dihydroxyvitamin D_3 levels in the serum may correlate to the healing of tibial fractures in chicks.

The biological actions of vitamin D are carried out by the active metabolite, calcitriol because it is a major controlling hormone, known for promoting and stimulating intestinal calcium and phosphate absorption.[8,9,19,26-30] It binds to and directly interacts with the vitamin D receptor (VDR), which is present in the nucleus of cells. Calcitriol is well known to play a critical role in the transcription and expression of genes such as the gene encoding the calcium binding protein, which is involved in the transport of intestinal calcium and other important minerals.[31,32]

The most abundant metabolite, calcidiol is clinically used as a marker compound for determining vitamin D status because of its high concentration levels and its direct link to the vitamin D substrate. However, like 24,25-dihydroxyvitamin D_3, it is biologically inactive. It is difficult to use the active compound, calcitriol, as a biomarker to determine vitamin D_3 sufficiency levels as its half-life is only a few hours and it is only present at very low concentration levels e.g. picomolar range in the blood and serum.[33]

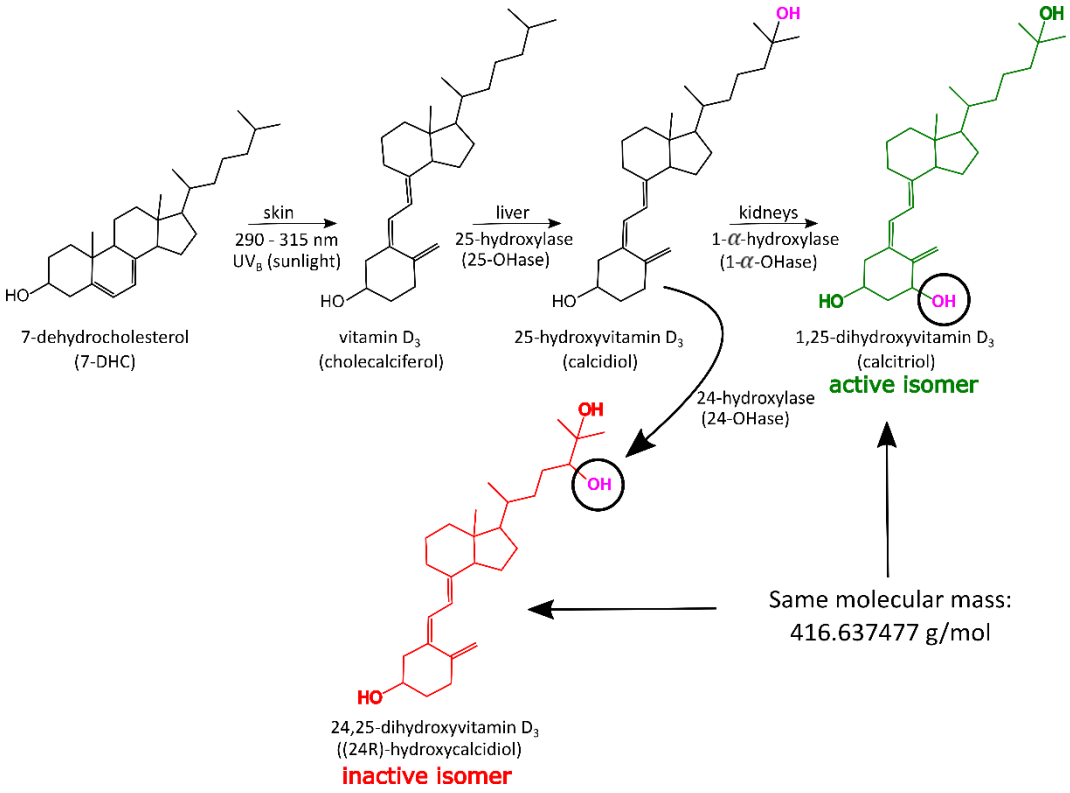


Figure 4. 2 Pathway for vitamin D metabolism with the highlighted OH groups to emphasise the difference in structures of the dihydroxylated isomers (adapted and redrawn from Müller *et. al.* clinical chemistry).[34]

Currently, the methods routinely used for determining vitamin D status in humans include immunoassays[35,36] and liquid chromatography-tandem mass spectrometry (LC-MS/MS)[37,38] via measurement of serum calcidiol levels. Early immunoassays such as competitive protein binding assays with the vitamin D binding protein (DBP) as the binder were used as DBP was able to recognise and bind to calcidiol in the serum. However, the common problem with many of the different types of immunoassays is the difficulty in separating out the intermediate products such as 24,25-dihydroxyvitamin D_3 and other vitamin D metabolites in the serum, which can result in significant variation of the total calcidiol levels. Immunoassays can also take time as only one metabolite is often measured per assay so the selectivity, accuracy and reproducibility may suffer as a result.

LC-MS/MS assays however, provide better selectivity, sensitivity and reproducibility and are widely considered the "gold-standard" method for the measurement of calcidiol and other vitamin D metabolites.[39-43] However, due to the low abundance of certain metabolites such as 1,25-dihydroxyvitamin D_3 and the complex matrices they are detected in such as human serum and blood, qualitative and quantitative

analysis can be difficult due to isomeric interferences that can arise from biological fluids.[44]

Qualitative analysis of the vitamin D metabolites has also been performed using gas chromatography-mass spectrometry (GC-MS) but the metabolites tend to require modification or derivatization using agents such as trimethylsilyl (TMS)[45], 4-phenyl-1,2,4-triazoline-3,5-dione (PTAD)[46-48] or Amplifex[49,50] prior to GC-MS analysis (Figure 4.3).[51] An example of the reaction scheme for calcidiol derivatization with Amplifex is also shown in Figure 4.4.

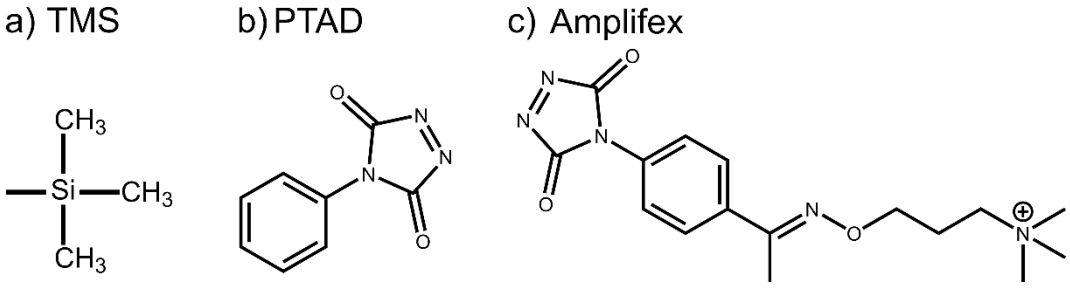


Figure 4. 3 Common derivatization agents used for vitamin D metabolites.

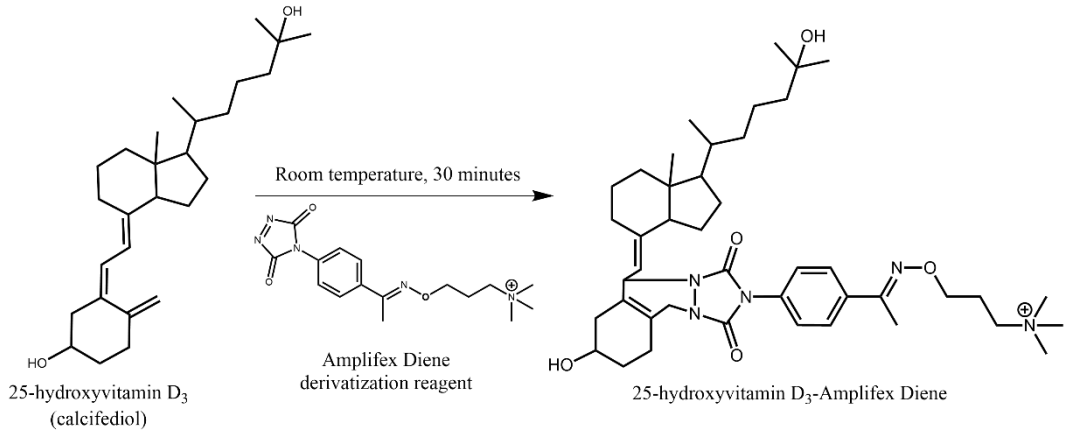


Figure 4. 4 Reaction scheme for calcidiol derivatization with the Amplifex (adapted and redrawn from Yang *et. al.*)[45]

Derivatization agents can also be used with other hyphenated chromatography mass spectrometry technique such as LC-MS/MS. Derivatization can improve the ionisation efficiency of the sample and also decreases any interference from metabolites close in *m/z* to the target species present in the biological media e.g. blood or serum by shifting the *m/z* range of the vitamin D metabolites to higher *m/z* values.[52] However an additional step is required during sample preparation (as shown by Figure 4.4) and may

require the data to be interpreted more carefully. Problems of steric hindrance might lead to mixtures of partially and fully derivatized analytes, hence the under-reaction and over-reaction when carrying out derivatization would need to be carefully controlled and optimised.

Recently, tandem MS methods have been further developed to differentiate between isomeric and epimeric vitamin D_3 metabolites. For example, Qi *et. al.*[53] developed a matrix assisted laser desorption ionisation (MALDI) combined with CAD-MS/MS method based on reactive analyte/matrix adducts using hydrogen-donor matrices such as 1,5-diaminonaphthalene (1,5-DAN), which enabled hydrogen transfers between matrix and analyte prior to desorption of ions. They demonstrated that it was possible to easily differentiate between 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 using the reactive adducts generated via MALDI, resulting in characteristic CAD MS/MS fragments for the isomers post ionisation by MALDI.[53] Chouinard *et. al.*[54] tested the separation capabilities of ion mobility-mass spectrometry (IMS-MS) to distinguish between the gas-phase conformations of 25OHD_3 epimers (Figure 4.5) with the aid of theoretical modelling of the epimers.[54,55]

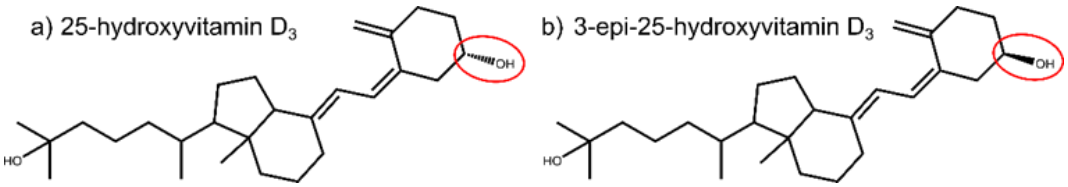


Figure 4. 5 Chemical structures of a) 25-hydroxyvitamin D_3 (also known as calcidiol) and b) 3-epi-25-hydroxyvitamin D_3 (the epimeric form of calcidiol).

These recent developments have encouraged use of different mass spectrometry techniques to further characterise and elucidate the structures of vitamin D metabolites. In this work, the use of a 12 T FT-ICR MS, equipped with various fragmentation methods as investigated, to enable differentiation of the two dihydroxylated vitamin D_3 isomers, without prior chromatographic separation or derivatization of the samples. Slow heating fragmentation methods such as CAD MS/MS were also revisited as a higher resolution MS technique was applied herein compared to previous experiments.

Photodissociation methods including IRMPD and UVPD MS/MS, as well as electron mediated fragmentation techniques such as EID MS/MS were also explored. Multiple diagnostic fragments were observed from the highly active metabolite, 1,25-dihydroxyvitamin D_3, in every tandem MS fragmentation experiment, demonstrating the

capability of using MS/MS on the 12 T FT-ICR MS instrument solely for clear and accurate differentiation of the dihydroxylated vitamin D_3 isomers.

4.3 Experimental Section

Sample preparation

Solvent evaporated standards of 1,25-dihydroxyvitamin D_3 (15 µg) and 24,25-dihydroxyvitamin D_3 (10 µg) were provided by the Volmer group from Humboldt University of Berlin, Germany and prepared as previously described.[53] Ultra-pure water was obtained from a Millipore (Merck Millipore, MA, USA) Direct-Q Milli-Q UV III purification system (18.2 Ω). LC-MS grade methanol (⩾99.9 %) was purchased from VWR Chemicals (Germany) and formic acid was purchased from Honeywell Fluka (Germany). The samples were prepared to stock solutions of 36 µM for the dihydroxylated vitamin D_3 isomers in methanol, which were then stored in a -80 °C freezer. Final samples were diluted with water: methanol (50:50, v/v) with 1 % v/v formic acid into concentrations of 10 µM for MS, CAD, IRMPD, EID and UVPD MS/MS experiments. For the direct infusion relative quantification of 1,25-dihydroxyvitamin D_3 by IRMPD MS/MS, the standard samples of 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 were mixed at 20 %, 40 %, 60 %, 80 %, and 100 % with the final concentration at 10 µM.

FT-ICR MS analysis

A 12 tesla (T) SolariX Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS; Bruker Daltonik GmbH, Bremen, Germany), equipped with an actively shielded superconducting magnet was used for the experiments.

The samples were loaded into borosilicate glass capillary tips (purchased from World Precision Instruments, Inc., Sarasota, FL, USA), which were pulled into nESI emitter tips using a Sutter P-97 capillary Flaming/Brown micropipette puller instrument (Sutter instruments Co., Novato, CA, USA) for nanoelectrospray ionisation (nESI). The samples were sprayed using a homemade nano-electrospray ion source in positive ionisation mode. Ions were externally accumulated in a hexapole collision cell for 0.5 s before they were transferred to an Infinity ICR analyser cell for MS detection.

Mass spectra were acquired with a 4 mega-word (M) data-points (32 bits) over a mass range of *m/z* 98.2 – 1,000 to produce a 1.12 s transient and ~300,000 resolving power at *m/z* 400. Spectra were calibrated to less than 100 parts-per billion (ppb) standard deviation, then assigned to less than 1 ppm.

FT-ICR MS/MS analysis

For all MS/MS experiments, the 1+ protonated precursor ions were isolated at *m/z* 417.3 using the quadrupole mass filter with an isolation window of 5 *m/z*. For CAD MS/MS after mass isolation of the precursor ions, argon was used as the collision gas and the resulting fragments were accumulated in the collision cell. The collision energy was optimized to 10 V. Fragments were then transferred to the ICR cell for detection.

For IRMPD MS/MS, precursor ions were first isolated with the quadrupole and then transmitted to the ICR cell. The ions were then fragmented using a continuous-wave CO_2 laser (Synrad, Washington, USA) with an output wavelength of 10.6 μm. The optimised pulse length of 0.1 s and 50 % laser power from the 25 W laser was used for the fragmentation. The CO_2 laser was introduced from the back of the ICR cell through a BaF_2 window and precursor ions were fragmented inside the ICR cell before excitation and detection.

For the EID MS/MS experiments, the quadrupole isolated ions were accumulated in the hexapole for 1 s. Isolated ions of interest were transferred and trapped in the ICR cell. The trapped ions were then irradiated with medium energy electrons from an indirectly heated hollow dispenser cathode via 1.5 A continuous current for ion fragmentation. The optimised EID MS/MS parameters for the fragmentation of the dihydroxylated vitamin D_3 isomers were a pulse length of 0.4 s, cathode bias/electron energy of 19 V, and extraction lens voltage of 3 V.

Along with a pre-existing IRMPD setup, a 193 nm ArF excimer laser beam (10 Hz; Coherent, UK) was co-introduced from the back of the ICR cell through a BaF_2 window for the UVPD MS/MS experiments. Ions were first quadrupole isolated and then transmitted to the ICR cell for ion fragmentation. The trapped ions were irradiated with 5 laser shots (~5 mJ/pulse measured at the laser head).

A stable telescopic compact high-energy Q-switched pulsed Nd:YAG laser with an output wavelength of 213 nm (5th harmonic of the Nd:YAG laser) (10 Hz; Litron

Lasers, UK) was also used for the UVPD MS/MS experiments and ions were irradiated with 10 laser shots (~1.5 mJ/pulse measured at the laser head).

Data analysis

All spectra were internally calibrated, analysed and manually assigned via DataAnalysis 4.3 software (Bruker Daltonik, GmbH, Bremen, Germany) to achieve sub-ppm accuracy for all assigned fragment peaks. For the relative quantification of 1,25-dihydroxyvitamin D_3 using IRMPD MS/MS data, the significant parameters are mentioned below.

All spectra were calibrated using the fragmented peaks with a minimum threshold of S/N >3 and intensities higher than $1x10^6$ determined by the Bruker FTMS peak picking algorithm. Each calibration point was averaged from three IRMPD MS/MS spectra for each isomer mixture and the R-square of the calibration curve using two different characteristic fragments of 1,25-dihydroxyvitamin D_3 was calculated to be higher than 0.99.

To generate the calibration points using the IRMPD MS/MS data, the following equation [Eqn. 4.1] was used:

$$Relative\ percentage\ ratio\ of\ 1{,}25(OH)_2D_3\ (\%) = \frac{peak\ area\ of\ characteristic\ 1{,}25(OH)_2D_3\ \ fragment}{sum\ of\ all\ fragment\ peak\ areas} \times 100 \quad \text{[Eqn. 4.1]}$$

The percentage ratios calculated using the equation above were then used as points to build a linear calibration curve. The equations of the calibration curves were used to determine the percentage of $1{,}25(OH)_2D_3$ generated in biological samples. However, application of the in-silico calibration curves to human samples in serum or blood is not explored within this body of work as the focus was on the implementation of various fragmentation methods available of the SolariX 12 T FT-ICR MS platform to distinguish between the isomeric dihydroxylated vitamin D_3 metabolites.

All calibration curves were plotted and a linear fit was applied using Origin 2019 (OriginLab Corporation, USA).

4.4 Results and Discussion

The mass spectra of the dihydroxylated vitamin D_3 isomers showed that the singly charged, protonated species was observed for both isomers of interest, demonstrating that it was not possible to differentiate between them solely based on single MS screening (Figure 4.6). In Figure 4.6, the mass spectra showed a significant intensity difference between 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 protonated and sodiated peaks. However, the exact purities, synthetic processes, and storage conditions of both molecules were not completely understood, therefore it was not possible to differentiate the molecules using purely the intensity differences between the mass spectra of the isomers either.

Although the molecular ions of the target species could be readily observed and assigned, there is no mass difference between the isomeric dihydroxylated vitamin D_3 metabolites as both isomers have the same exact mass with the same elemental composition ($C_{27}H_{44}O_3$), hence no further information regarding the structural characterisation of both isomers can be obtained from only the mass spectra.

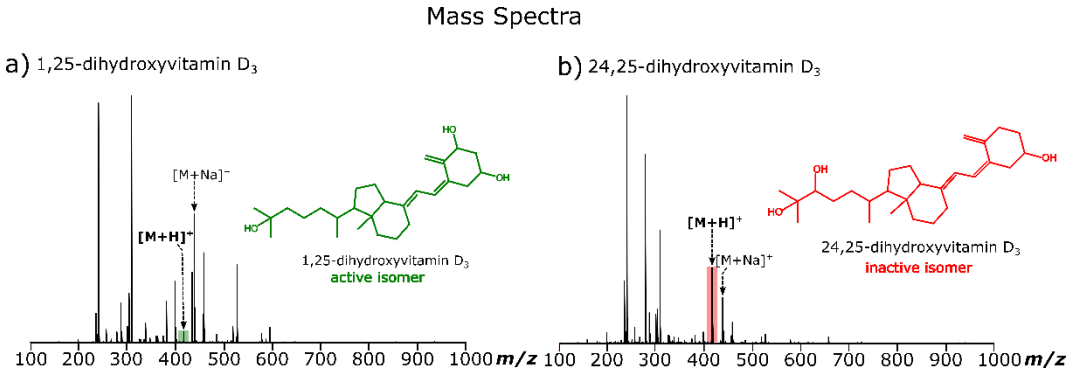


Figure 4. 6 a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 mass spectra.

As illustrated in Figure 4.7, three major peaks ($[M+H-H_2O]^+$, $[M+H-2H_2O]^+$, and $[M+H-3H_2O]^+$) were identified in the CAD MS/MS spectra, corresponding to the loss of the hydroxyl groups from both metabolites. This was also observed in the mass spectra, confirming the fragile nature of the hydroxyl groups on the ring of both isomeric species. A homologous series of hydrocarbon losses ($-CH_2$) resulting from direct carbon-carbon (C-C) cleavages were also observed in all the fragmentation spectra. These fragments contribute to the complex spectra, providing limited structural and no isomer-specific information as these fragments were observed for both 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3.

In previous studies, application of CAD MS/MS to vitamin D_3 compounds resulted in dense spectra, which was also observed in this work due to the typical "picket fence" ion series of hydrocarbon chain losses.[44,56] Other problems noted in the studies include the lack of more easily ionizable groups on the molecules as well as the difficult separation and analysis of the vitamin D metabolites when in complex matrices such as serum or urine. This is because in human samples, the compounds are present in low levels and interference from other species in the matrices can contribute to the ion suppression of the vitamin D_3 compounds, particularly for the active metabolite 1,25-dihydroxyvitamin D_3. Therefore, the experiments conducted in this work were tested on the provided vitamin D_3 standards as a basis for method development.

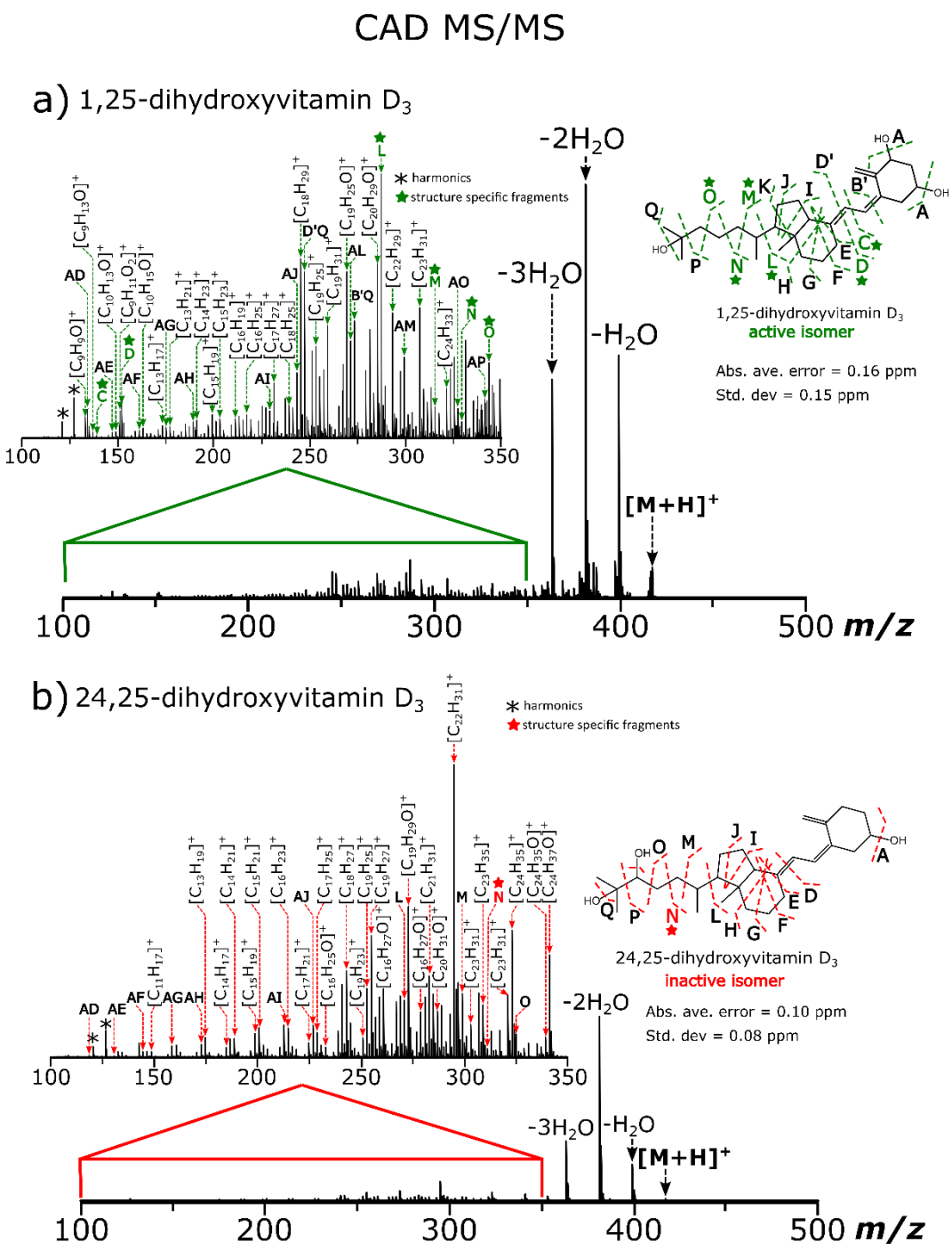


Figure 4. 7 CAD MS/MS spectra with inserts of *m/z* 100 – 350 regions with fragment peaks labelled for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with structure specific fragments denoted by the star symbol. The peak assignment tables a) S4.1 and b) S4.2 are displayed in the supplementary information for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3, respectively.

After collision energy optimisation and detailed analysis of the CAD MS/MS spectra, diagnostic fragments were detected for 1,25-dihydroxyvitamin D_3, which were absent in the 24,25-dihydroxyvitamin D_3 MS/MS spectra.

For each MS/MS method studied (CAD, EID, IRMPD, and UVPD), the parameters required for fragmentation optimisation were individually optimised and up to 100 scans were accumulated to ensure that the characteristic fragments observed for 1,25-dihydroxyvitamin D_3 spectra, were absent for 24,25-dihydroxyvitamin D_3 spectra. This included optimisation of the collision energy for CAD MS/MS experiments, the pulse length for ion interaction with IR or UV photons for both IRMPD and UVPD MS/MS as well as the cathode bias voltage, which is responsible for the energy of the electrons for the EID MS/MS experiments. Figure 4.8 demonstrates the importance of collision energy optimisation for the detection of one of the multiple characteristic fragments detected for 1,25-dihydroxyvitamin D_3, which was absent when a collision energy of 5 V was applied but present with a collision energy of 8 V. The S/N of the characteristic fragment D was further improved (approximately 2-fold increase in S/N) when the optimised collision energy of 10 V was used.

Although fragment D was detected at low *m/z* with a resolving power of 825,058, the benefits of high mass resolution is observed and required to separate the critical isomer-differentiating fragments from other detected peaks in the spectrum. High resolution was also useful for resolving the lower intensity diagnostic fragment ions from the series of CH_2 losses, which dominated all MS/MS spectra of these species.

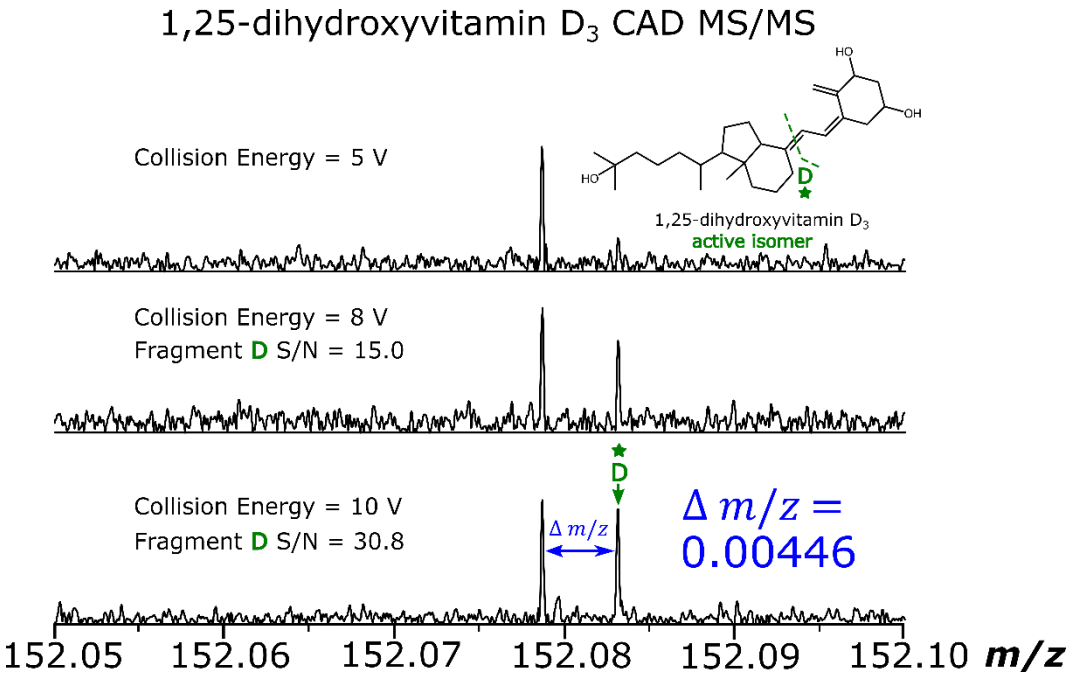


Figure 4. 8 *m/z* scale expansions of the region *m/z* 152.05 – 152.10, highlighting the presence of a structure specific fragment (labelled D) in the CAD MS/MS spectra of the active metabolite, 1,25-dihydroxyvitamin D_3 after collision energy optimisation.

As observed with the CAD MS/MS spectra, IRMPD MS/MS also resulted in complex fragmentation spectra for the dihydroxylated vitamin D_3 isomers. An additional cleavage for 1,25-dihydroxyvitamin D_3 was observed using IRMPD resulting in the formation of the diagnostic fragment B as shown in Figure 4.9a). Both CAD MS/MS and IRMPD MS/MS are based on the excitation of vibrational energy modes and energetically weaker bonds often cleave preferentially. However, testing all the available fragmentation methods presents an opportunity for comparison of the suitability of each method for qualitative and quantitative analysis. Depending on the MS/MS method used, the metabolites may undergo a different fragmentation pathway, resulting in secondary fragmentation, improvement in the number of diagnostic fragments detected or an improvement in the relative intensities of those diagnostic fragments.

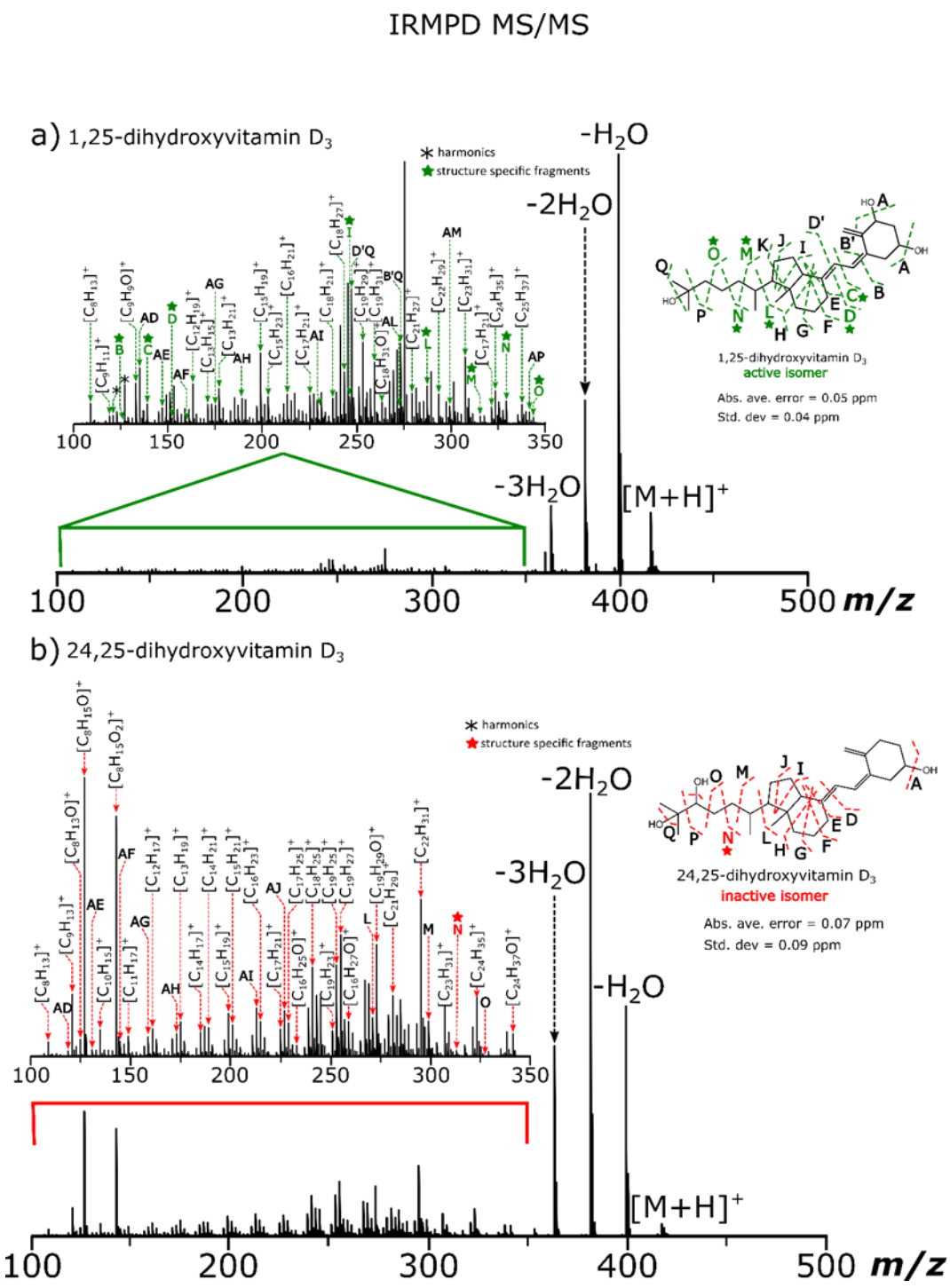


Figure 4. 9 IRMPD MS/MS spectra with inserts of *m/z* 100 – 350 region with fragments labelled for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with structure specific fragments denoted by the star symbol. The corresponding peak assignment tables a) S4.3 and b) S4.4 are displayed in the supplementary information for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3, respectively.

An example of a key diagnostic fragment observed is shown in Figure 4.10, which was observed with CAD, IRMPD MS/MS, and the other MS/MS methods, where

the detected fragment C in the MS/MS spectra for 1,25-dihydroxyvitamin D_3 was present, while absent in the 24,25-dihydroxyvitamin D_3 MS/MS spectra. This indicates that both fragile hydroxyl groups on the ring can be preserved during dissociation of the active 1,25-dihydroxyvitamin D_3 metabolite but not for the inactive 24,25-dihydroxyvitamin D_3 isomer. The sensitivity of the test criteria can also be significantly improved by using IRMPD as a fragmentation technique instead of CAD as an 8-fold improvement in the S/N ratio for the fragment C generated by IRMPD MS/MS is observed compared to the same fragment generated by CAD MS/MS.

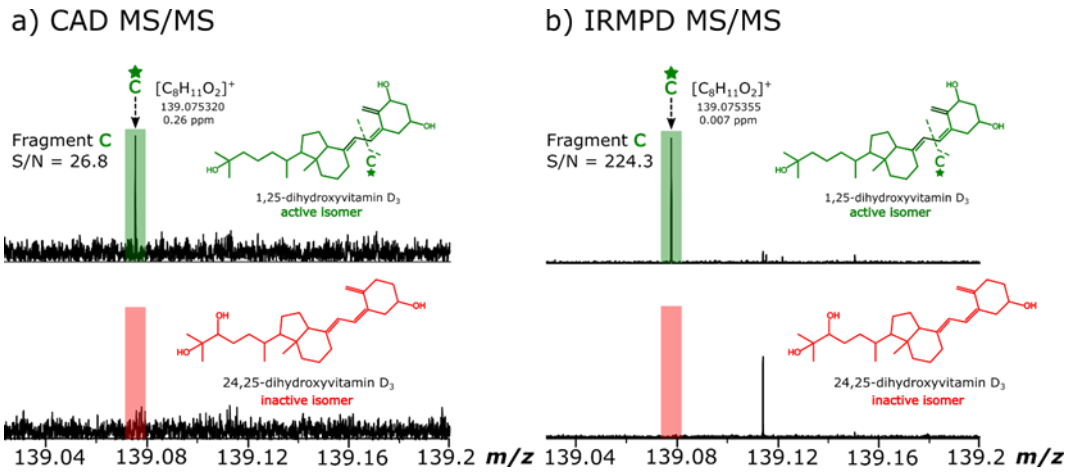


Figure 4. 10 Zoom in of *m/z* 139.02 -139.2 region of a) CAD MS/MS spectra and b) IRMPD MS/MS spectra of 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 showing 1,25-dihydroxyvitamin D_3 specific fragment C. An 8-fold improvement in the S/N is also noted for the diagnostic fragment C using IRMPD MS/MS compared to CAD MS/MS.

The internal calibration of all fragmentation spectra resulted in sub-ppm mass accuracy assignment errors. Peak Assignments were made based on closely following the criteria mentioned here. For example, all product ions and in particular the characteristic fragment ions which enabled differentiation between the isomers, were checked manually and based on low mass errors (<1 ppm). The isotopic simulation of each characteristic fragment ion was checked and matched with the observed fragment ion in the MS/MS spectra obtained. It is important to have sub-ppm mass errors for such dense fragmentation spectra as multiple assignments are possible for each nominal *m/z*, hence it is also necessary to accompany this with detailed exact mass calculation and matching simulation of the assigned elemental formulae as shown in Figure 4.11 a) and b).

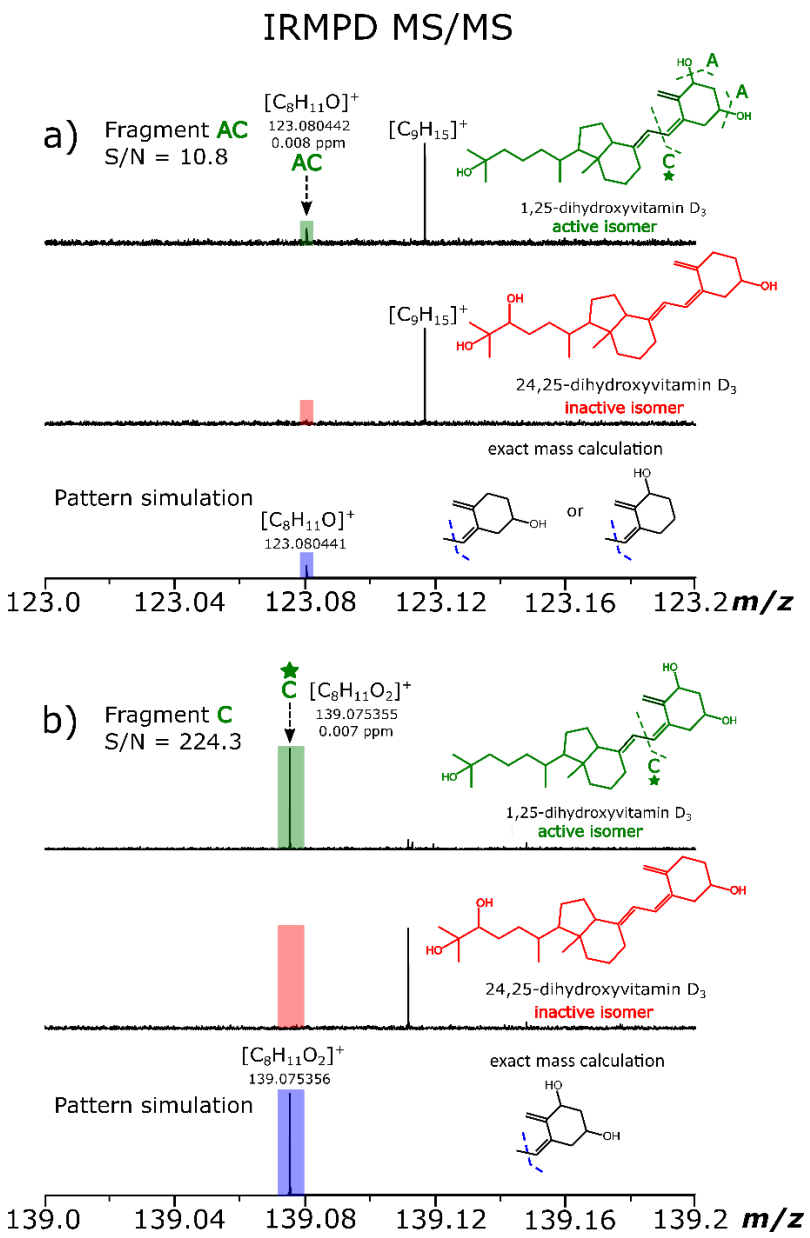


Figure 4. 11 *m/z* scale expansions of IRMPD spectra obtained of a) *m/z* 123.0 – 123.2 and b) *m/z* 139– 139.2 for the IRMPD fragment ions of 1,25$(OH)_2D_3$ (top traces) and 24,25$(OH)_2D_3$ (middle traces). The exact mass calculation and simulation of the assigned elemental formulae with chemical structures are shown in the bottom trace.

As the vitamin D metabolites are singly charged species, ECD MS/MS is not viable due to the charge limitation effect of the ECD process, however EID MS/MS can be applied. Irradiation of singly charged ions with higher energy electrons can be used to perform EID, which results in fragments like those generated with electron ionisation (EI). Both techniques result in extensive fragmentation yet EID produces relatively low intensity fragments and retains precursor ions. This is in contrast with EI, where many are mostly, if not fully fragmented so the molecular ion is low in intensity or absent from the spectra, potentially making a definitive identification of unknown compounds more difficult.

For the vitamin D metabolites studied in this work, with EID MS/MS, the molecular ion can be clearly identified (Figure 4.12). However, this would not be an issue even if the molecular ion was absent from the MS/MS spectra as the MS spectra were obtained and the structures, as well as the identity of the metabolites of interest are known. Both electron-mediated fragmentation methods are radical-based processes and may also result in the formation of secondary and tertiary fragments. This contributes to the complexity of the MS/MS spectra obtained, which often require careful manual interpretation.

After optimisation of the electron energy and pulse length, fragmentation of both dihydroxylated vitamin D_3 isomers resulted in complex EID spectra (Figure 4.12). Although the relative intensity of the diagnostic and non-diagnostic fragments generated by EID were overall lower compared to both CAD and IRMPD (Table 4.1), complementary structural information was obtained with EID.

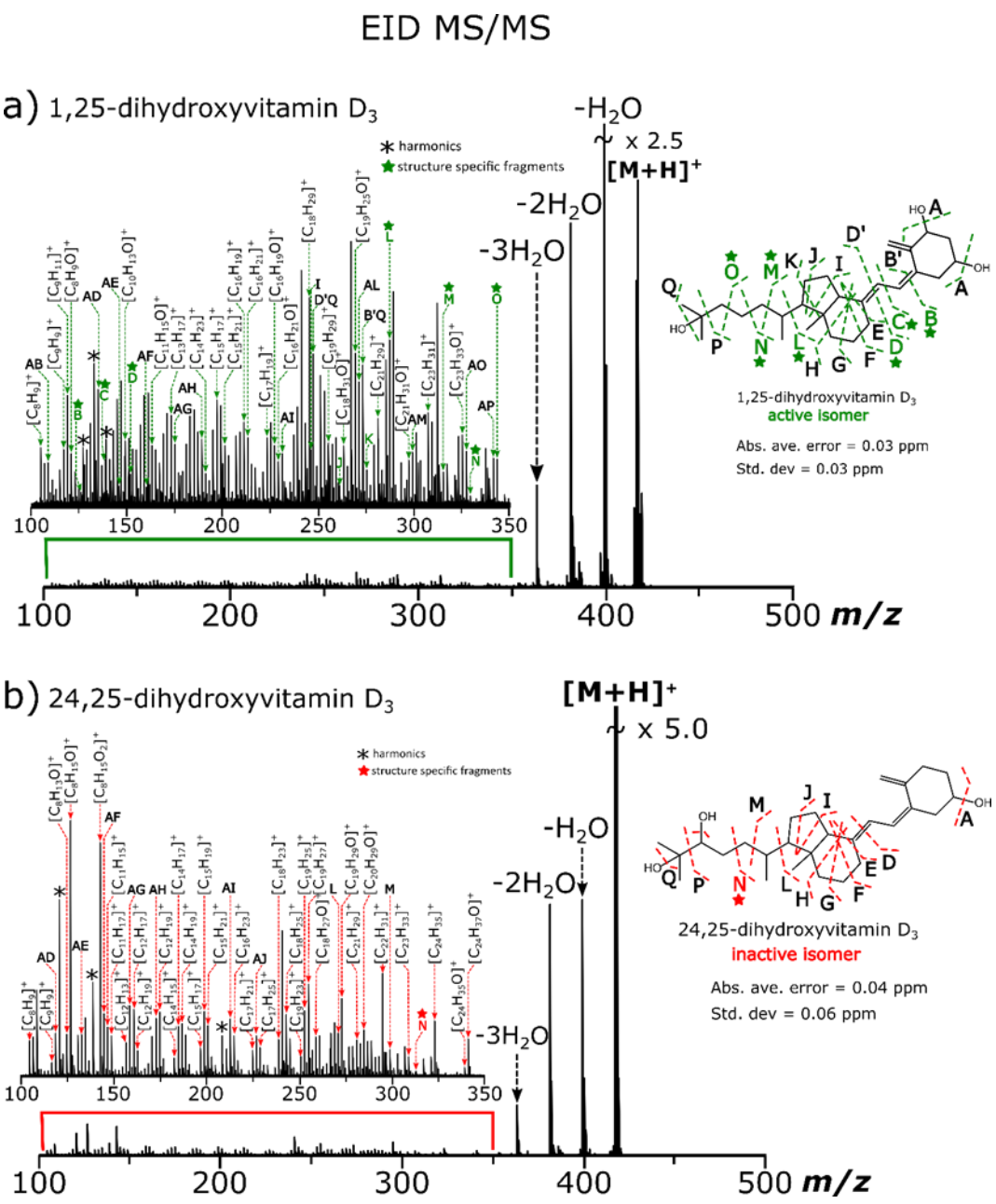


Figure 4. 12 EID MS/MS spectra with inserts of *m/z* 100 – 350 regions with fragment peaks labelled for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with structure specific fragments are denoted by the star symbol. The corresponding peak assignment tables a) S4.5 and b) S4.6 are displayed in the supplementary information for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3, respectively.

Compared to IRMPD, UVPD is a higher energy activation method based on the absorption of UV photons by the analyte ions, which is possible due to the UV chromophore properties of the C=C double bonds present in the 5,6-cistriene system of the vitamin D compounds.

UVPD MS/MS experiments were carried out at 193 nm and 213 nm wavelengths using an ArF Excimer laser and a Nd:YAG laser, respectively. The structural information obtained with 193 nm UVPD for the dihydroxylated vitamin D_3 compounds (Figure 4.13) compared well with the MS/MS data obtained with CAD, IRMPD, and EID MS/MS. This observation may be supported by a combination of the previously proposed UVPD mechanisms, which are known as direct dissociation and internal conversion.[57] Direct dissociation is a result of electronic excitation or relaxation into a dissociative orbital, like that of electron-based fragmentation methods, whereas internal conversion of the photon energy into vibrational modes results in fragmentation in the ground state so the fragments generated will be like those generated by CAD and IRMPD.[57]

With 213 nm UVPD, the fragments obtained were low intensity compared to other MS/MS methods yet structure-specific fragments and cross-ring cleavages across both molecules were observed (Figure 4.14). It is difficult to make a direct comparison between the performance of the 193 nm and 213 nm UVPD on the data obtained as the number of laser shots and the energy output for each laser were different. For the 193 nm UVPD MS/MS experiments, 5 laser shots at 5 mJ/pulse were used and 10 laser shots at 1.5 mJ/pulse were used for 213 nm UVPD MS/MS. However, as shown in Table 4.1, although most fragments were low intensity, many of the isomer-specific fragments (9 out of the 12) listed for 1,25-dihydroxyvitamin D_3 were detected with 213 nm UVPD MS/MS.

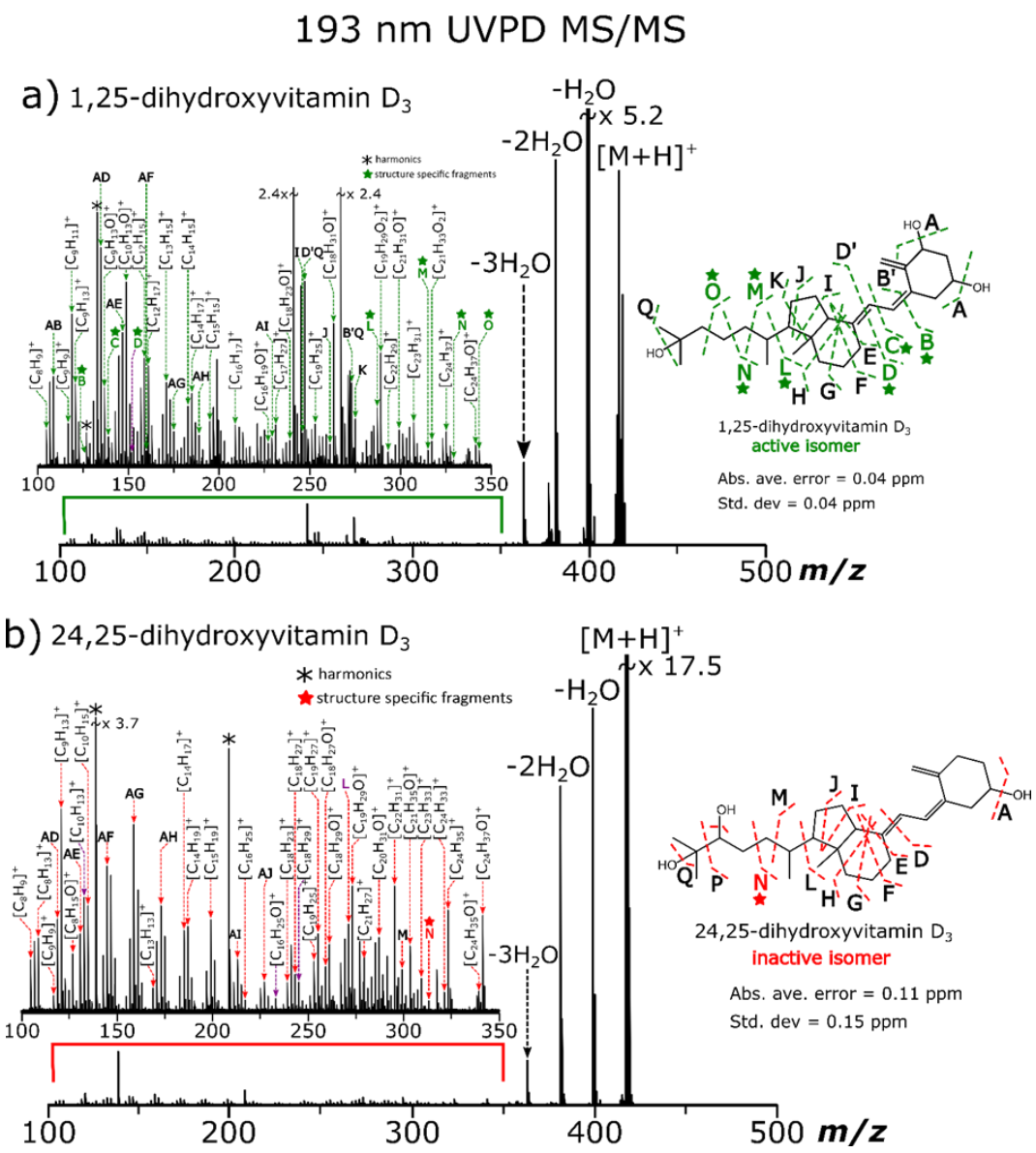


Figure 4. 13 The 193 nm UVPD MS/MS spectra with inserts of *m/z* 100 – 350 regions with the fragment peaks labelled for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with structure specific fragments are denoted by the star symbol. The corresponding peak assignment tables a) S4.7 and b) S4.8 are displayed in the supplementary information for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3, respectively.

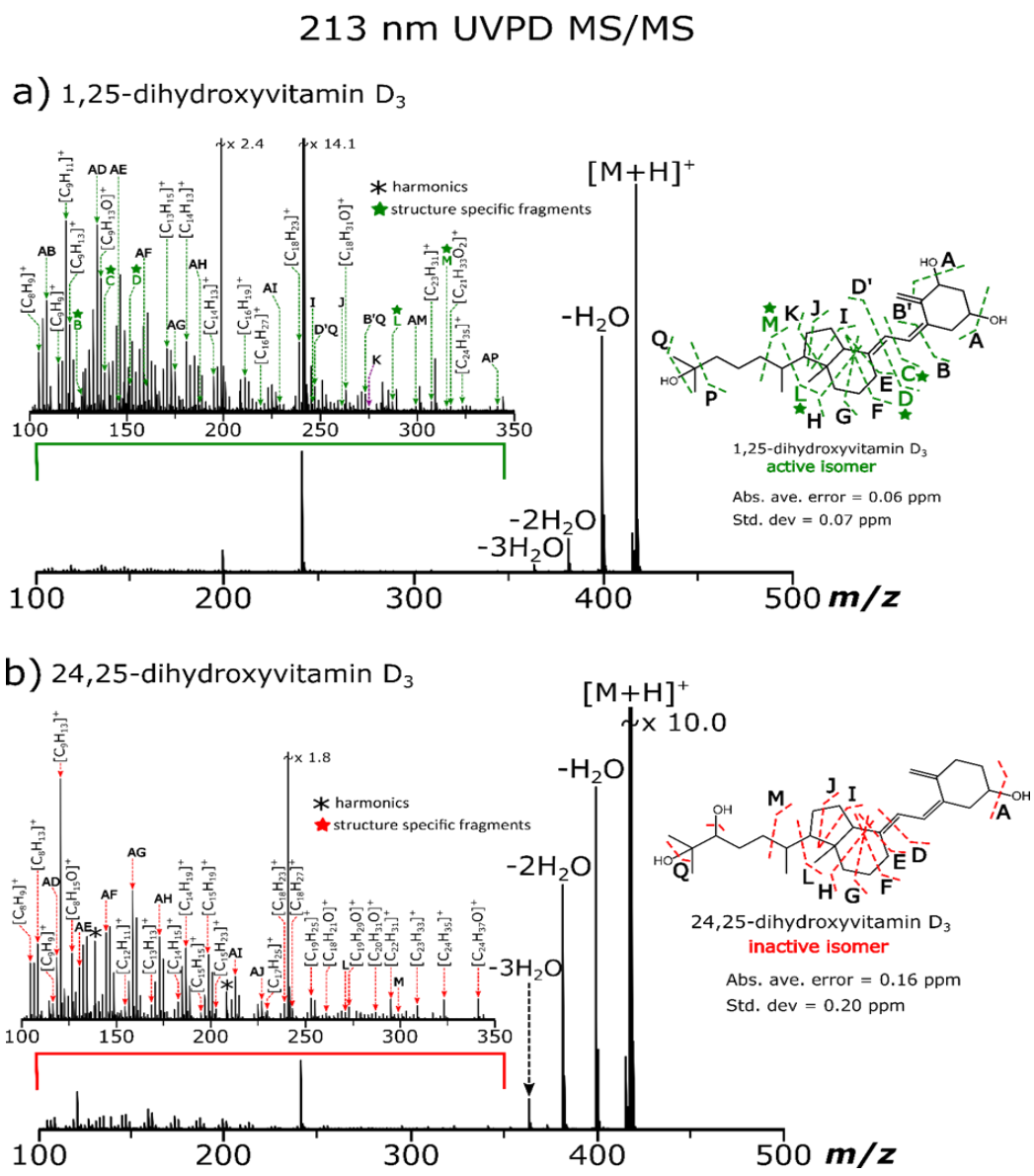


Figure 4. 14 The 213 nm UVPD MS/MS spectra with inserts of *m/z* 100 – 350 regions with the fragment peaks labelled for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3 with structure specific fragments are denoted by the star symbol. The corresponding peak assignment tables a) S4.9 and b) S4.10 are displayed in the supplementary information for a) 1,25-dihydroxyvitamin D_3 and b) 24,25-dihydroxyvitamin D_3, respectively.

Multiple diagnostic fragments were detected as shown in Table 4.1. The table displays the main diagnostic fragments detected in the 1,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$) MS/MS spectra, which were definitively absent in the MS/MS spectra of 24,25-dihydroxyvitamin D_3 ($24,25(OH)_2D_3$), using the various fragmentation methods available. This was shown mainly for the 1,25-dihydroxyvitamin D_3 isomer as this

metabolite had fragments that were also generated for 24,25-dihydroxyvitamin D_3, due to the fragile nature of the hydroxyl groups on the A ring and the aliphatic side chain of the molecule. Extensive fragmentation including cross-ring cleavages of both dihydroxylated vitamin D_3 isomers was observed with all fragmentation methods applied. The assignment of the diagnostic fragments in Table 4.1 corresponds to the assigned cleavages of the $1,25(OH)_2D_3$ compound e.g., fragment "AD" refers to bonds "A" and "D" broken in $1,25(OH)_2D_3$, as shown by the cleavage diagram in Figure 4.14.

Table 4. 1 Fragmentation table for diagnostic fragments, where one or both OH groups were retained on the ring for 1,25-dihydroxyvitamin D_3, which are absent in the 24,25-dihydroxyvitamin D_3 spectra. In the table, "X" denotes the absence of the fragment in the 1,25-dihydroxyvitamin D_3 MS/MS spectra and further explanation about the fragment intensity level (S4.12) is provided in the supplementary information along with an expanded version of the fragmentation table (S4.11).

$1,25(OH)_2D_3$ characteristic theoretical fragment (*m/z*)	**Assignment**	**Fragmentation Method**									
		CAD		**IRMPD**		**EID**		**193 nm UVPD**		**213 nm UVPD**	
		Intensity	**S/N**	**Intensity**	**S/N**	**Intensity**	**S/N**	**Intensity**	**S/N**	**Intensity**	**S/N**
109.064791	AB	X	X	X	X	medium	18.1	medium	53.6	medium	58
127.075356	B	X	X	high	231.5	medium	18.9	low	23.1	low	42
135.080441	AD	high	346.4	high	639.4	high	106.7	high	585.7	high	507
139.075356	C	low	26.8	high	224.3	medium	12.6	low	29.5	low	27.1
147.080441	AE	low	20	low	27.7	low	21.2	medium	40.9	low	40.6
152.083181	D	low	30.8	medium	63.6	high	42.2	medium	36.9	low	24.7
165.091006	E	low	44.2	medium	70.2	low	12.8	low	14.7	low	12.2
287.200557	L	high	390.1	high	423.9	medium	198.7	high	141.2	low	25.3
315.231857	M	high	27.2	medium	77.6	low	39.6	medium	38.3	low	9.2
329.247507	N	low	128.1	low	26.5	low	12.1	low	15.8	X	X
343.263157	O	high	493.1	medium	67.6	low	53.3	medium	36.9	X	X
357.278807	P	medium	78	low	16.7	X	X	X	X	X	X

The development of vitamin D detection and quantification methods are essential for the diagnosis of patients with vitamin D deficiency and other diseases. Methods for quantification of vitamin D metabolites in human serum and plasma samples are often carried out with immunoassays. However, recently LC-MS/MS assays are used for routine analysis and quantification of vitamin D compounds. The success and use of LC-MS/MS methods for quantification can be contributed to the sensitivity, reproducibility, and the capability for detection of several of the vitamin D_3 metabolites used as biomarkers for vitamin D deficiency.

Although LC provides an extra dimension for the separation of compounds in complex mixtures, identification cannot be confirmed based on the retention time alone as co-elution with the many other metabolites of vitamin D remains to be a significant problem. This is particularly an issue for the separation and detection of very low

abundant metabolites such as 1,25-dihydroxyvitamin D_3. Therefore, chromatographic separation and isolation of the metabolites can be beneficial with the aid of MS/MS methods, which can be used for the structural characterization of the compounds.

With the MS/MS methods applied in this work, the fragmentation patterns for dihydroxylated vitamin D isomers were similar yet isomer-specific fragments were detected, particularly for 1,25-dihydroxyvitamin D_3. An alternative direct infusion MS/MS approach is discussed herein using mixtures of the standard vitamin D isomers, where the specific fragments detected for the active dihydroxylated vitamin D_3. The aim is to demonstrate the potential of this method for the relative quantification of vitamin D metabolites in biological matrices.

The highlighted fragment ions at *m/z* 135.08 and *m/z* 287.20 in Table 4.1 were chosen to test the relative quantitation of the isomers as these fragments had the highest relative intensities and S/N out of the characteristic fragments listed in Table 4.1. The isomers, 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 were mixed to obtain mixtures, in which the 1,25-dihydroxyvitamin D_3 content varied from 0 % to 100 % in 20 % increments. Figure 4.15a) demonstrates clear discrimination between the dihydroxylated vitamin D_3 isomers. As the percentage of 1,25-dihydroxyvitamin D_3 in the standard mixtures of 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3 were increased, the intensity of the fragments at *m/z* 135.08 and *m/z* 287.2 generated by IRMPD MS/MS also increased in intensity.

A ratio was calculated using the peak area of the 1,25-dihydroxyvitamin D_3 specific fragment to the sum of all the fragments present in the IRMPD MS/MS spectrum for each isomer mixture. Fluctuations were observed in the calibration curve when only the peak area or the peak intensities of the diagnostic fragment was plotted against the percentage of 1,25-dihydroxyvitamin D_3 in the dihydroxylated vitamin D_3 isomeric mixture, thus this was avoided using the ratio equation [Eqn. 4.1], which is also reiterated below:

$$\text{Relative ratio of } 1,25(\text{OH})_2\text{D}_3\ (\%) = \frac{\text{peak area of characteristic } 1,25(\text{OH})_2\text{D}_3 \text{ fragment}}{\text{sum of all fragment peak areas}} \times 100$$

A calibration curve was obtained with good linearity (R^2>0.99) with the inclusion of the confidently assigned (mass error < 1 ppm) fragments and using the equation [Eqn. 4.1].

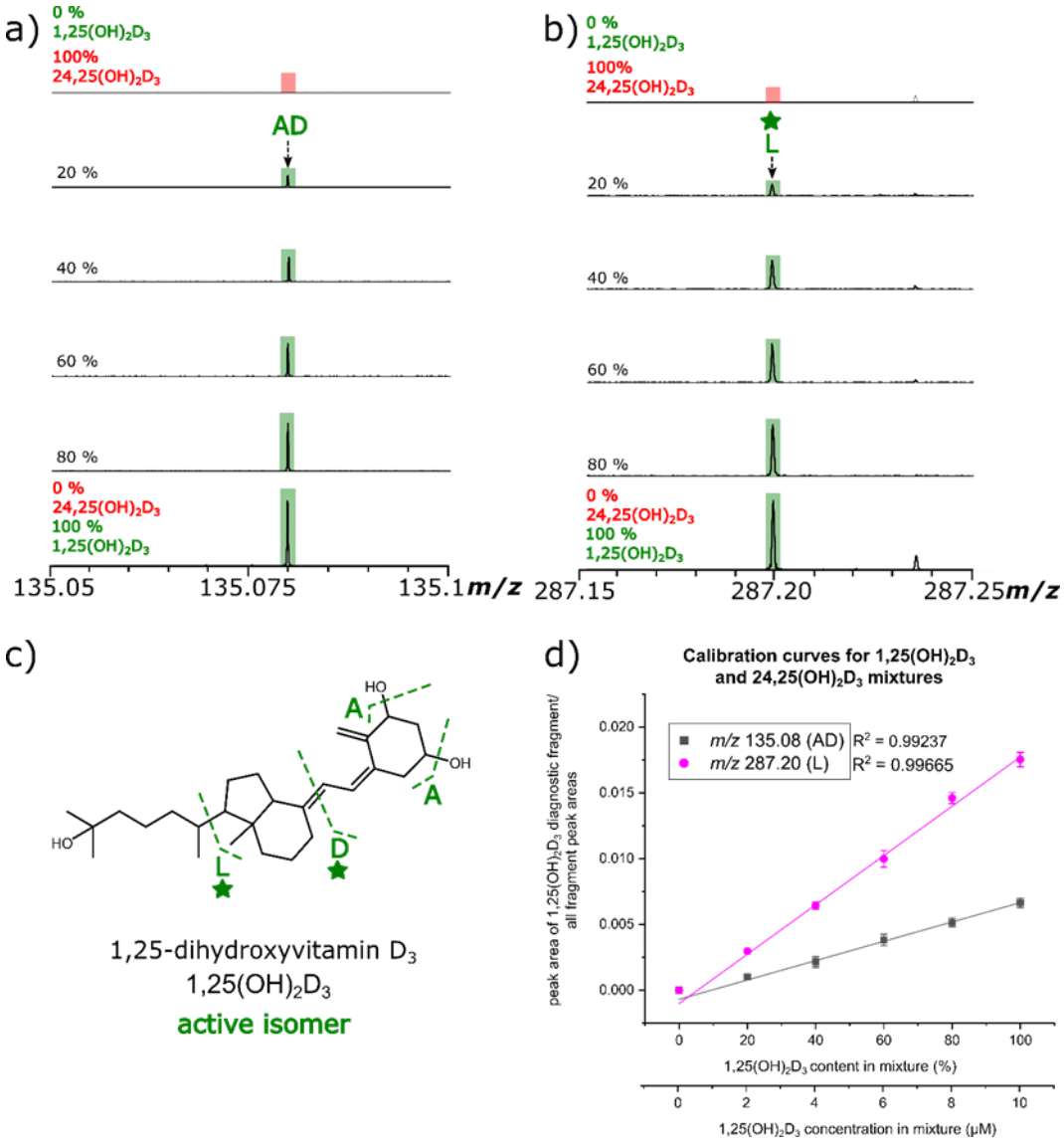


Figure 4. 15 *m/z* scale expansion of a) *m/z* 135.05– 135.10 and b) *m/z* 287.15 – 287.25 from the IRMPD spectra for the characteristic $1,25(OH)_2D_3$ IRMPD fragment ions with increasing increments of $1,25(OH)_2D_3$ in percentage concentration in the mixture (all spectra scaled to same y-axis intensity), c) structure of $1,25(OH)_2D_3$ with associated cleavages to produce the diagnostic fragments "AD" and "L" and d) calibration curves generated using the peak area ratio of the $1,25(OH)_2D_3$ specific "AD" and "L" fragments.

4.5 Conclusions

In this study, the use of CAD MS/MS was revisited and alternative fragmentation methods such as IRMPD, UVPD implemented at two different wavelengths, and EID MSMS/MS were investigated for the differentiation of the isomeric dihydroxylated vitamin D_3 compounds. Isomer-specific fragments were detected for the active metabolite, 1,25-dihydroxyvitamin D_3, which were absent for the inactive metabolite, 24,25-dihydroxyvitamin D_3 after optimisation of the parameters for each MS/MS method and accumulation of scans. The structure-specific fragments generated due to cleavage of the C-6/C-7 bond in the 1,25-dihydroxyvitamin D_3 compound successfully demonstrate the retention of the fragile hydroxyl groups during dissociation using all the available fragmentation methods.

It should be noted that the loss of the hydroxyl groups and series of hydrocarbon chain decompositions for both vitamin D_3 metabolites dominated all the MS/MS spectra obtained. However, after detailed analysis, multiple diagnostic fragments were detected and assigned with high confidence aided by the high resolving power and high mass accuracy capabilities provided by FT-ICR MS.

In summary, diagnostic fragments generated via all MS/MS methods were observed for 1,25-dihydroxyvitamin D_3, enabling quick and easy differentiation between the two dihydroxylated vitamin D_3, without the need for prior chromatographic separation or derivatization of the molecules.

Accurate quantification of the vitamin D metabolites is essential and continues to have a profound impact on the clinical diagnosis of vitamin D deficiency and other related diseases. The issues surrounding quantification include the difficulty in analysis of the samples as the vitamin D metabolites are structurally very similar and they are often directly measured in complex biological matrices such as human serum or blood. Therefore, adequate separation of the metabolites is necessary and definitive product confirmation is required depending on which vitamin D metabolites are targeted.

Preliminary experiments for the quantitative analysis of 1,25-dihydroxyvitamin D_3 were also carried out. Herein, we developed a quantification method that uses the peak area of a selected diagnostic fragment of 1,25-dihydroxyvitamin D_3 divided by the sum of the fragments peak areas to reduce the fluctuation caused by a single fragmented peak. A calibration curve using the diagnostic fragments of 1,25-dihydroxyvitamin D_3 was established with good linearity ($R^2 > 0.99$).

This direct infusion quantification method using MS/MS has the potential to be applied to the vitamin D_3 metabolites detected in matrices such as serum or urine, which are routinely found in low concentrations and often masked by other endogenous material. Hence, chromatographic separation prior to MS/MS analysis may be beneficial whilst the characteristic fragments listed in this work can be used to identify and quantify the biologically active 1,25-dihydroxyvitamin D_3 compound.

4.6 References

(1) Holick, M. F. Vitamin D Deficiency. N. Engl. J. Med. 2007, 357 (3), 266-281.

(2) Garg, M.; Sharma, A.; Vats, S.; Tiwari, V.; Kumari, A.; Mishra, V.; Krishania, M. Vitamins in Cereals: A Critical Review of Content, Health Effects, Processing Losses, Bioaccessibility, Fortification, and Biofortification Strategies for Their Improvement. Front. Nutr. 2021, 8 (254).

(3) Lu, Z.; Chen, T. C.; Zhang, A.; Persons, K. S.; Kohn, N.; Berkowitz, R.; Martinello, S.; Holick, M. F. An evaluation of the vitamin D3 content in fish: Is the vitamin D content adequate to satisfy the dietary requirement for vitamin D? J. Steroid Biochem. Mol. Biol. 2007, 103 (3), 642-644.

(4) Tripkovic, L.; Lambert, H.; Hart, K.; Smith, C. P.; Bucca, G.; Penson, S.; Chope, G.; Hyppönen, E.; Berry, J.; Vieth, R.; Lanham-New, S. Comparison of vitamin D2 and vitamin D3 supplementation in raising serum 25-hydroxyvitamin D status: a systematic review and meta-analysis. Am. J. Clin. Nutr. 2012, 95 (6), 1357-1364.

(5) Glendenning, P.; Chew, G. T.; Inderjeeth, C. A.; Taranto, M.; Fraser, W. D. Calculated free and bioavailable vitamin D metabolite concentrations in vitamin D-deficient hip fracture patients after supplementation with cholecalciferol and ergocalciferol. Bone 2013, 56 (2), 271-275.

(6) Wolf, G. The Discovery of Vitamin D: The Contribution of Adolf Windaus. J. Nutr. 2004, 134 (6), 1299-1302.

(7) Sunyecz, J. A. The use of calcium and vitamin D in the management of osteoporosis. Ther. Clin. Risk Manag. 2008, 4 (4), 827-836.

(8) Adami, S.; Frijlink, W. B.; Bijvoet, O. L. M.; O'Riordan, J. L. H.; Clemens, T. L.; Papapoulos, S. E. Regulation of calcium absorption by 1,25,dihydroxy-vitamin D—Studies of the effects of a bisphosphonate treatment. Calcif. Tissue Int. 1982, 34 (1), 317-320.

(9) Wasserman, R. H. Vitamin D and the Dual Processes of Intestinal Calcium Absorption. J. Nutr. 2004, 134 (11), 3137-3139.

(10) Christodoulou, S.; Goula, T.; Ververidis, A.; Drosos, G. Vitamin D and bone disease. BioMed Res. Int. 2013, 2013, 396541-396541.

(11) Soliman, A. T.; El-Dabbagh, M.; Adel, A.; Ali, M. A.; Aziz Bedair, E. M.; ElAlaily, R. K. Clinical Responses to a Mega-dose of Vitamin D3 in Infants and Toddlers With Vitamin D Deficiency Rickets. J. Trop. Pediatr. 2009, 56 (1), 19-26.

(12) Sahay, M.; Sahay, R. Rickets-vitamin D deficiency and dependency. Indian J. Endocrinol. Metab. 2012, 16 (2), 164-176.

(13) Laird, E.; Ward, M.; McSorley, E.; Strain, J. J.; Wallace, J. Vitamin D and bone health: potential mechanisms. Nutrients 2010, 2 (7), 693-724.

(14) Holick, M. F. Resurrection of vitamin D deficiency and rickets. J. Clin. Invest. 2006, 116 (8), 2062-2072.

(15) Holick, M. F. Vitamin D status: measurement, interpretation, and clinical application. Ann. Epidemiol. 2009, 19 (2), 73-78.

(16) Norman, P. E.; Powell, J. T. Vitamin D and Cardiovascular Disease. Circ. Res. 2014, 114 (2), 379-393.

(17) Valipour, G.; Saneei, P.; Esmaillzadeh, A. Serum Vitamin D Levels in Relation to Schizophrenia: A Systematic Review and Meta-Analysis of Observational Studies. J. Clin. Endocrinol. Metab. 2014, 99 (10), 3863-3872.

(18) Littlejohns, T. J.; Henley, W. E.; Lang, I. A.; Annweiler, C.; Beauchet, O.; Chaves, P. H. M.; Fried, L.; Kestenbaum, B. R.; Kuller, L. H.; Langa, K. M.; Lopez, O. L.; Kos, K.; Soni, M.; Llewellyn, D. J. Vitamin D and the risk of dementia and Alzheimer disease. Neurology 2014, 83 (10), 920.

(19) Bikle, D. D. Vitamin D metabolism, mechanism of action, and clinical applications. Chem. Biol. 2014, 21 (3), 319-329.

(20) Lam, H.-Y.; Schnoes, H. K.; DeLuca, H. F.; Chen, T. C. 24,25-Dihydroxyvitamin D3. Synthesis and biological activity. Biochemistry 1973, 12 (24), 4851-4855.

(21) Henry, H.; Norman, A.; Taylor, A.; Hartenbower, D.; Coburn, J. Biological Activity of 24,25-Dihydroxycholecalciferol in Chicks and Rats. J. Nutr. 1976, 106, 724-734.

(22) Endo, H.; Kiyoki, M.; Kawashima, K.; Naruchi, T.; Hashimoto, Y. Vitamin D3 metabolites and PTH synergistically stimulate bone formation of chick embryonic femur in vitro. Nature 1980, 286 (5770), 262-264.

(23) Galus, K.; Szymendera, J.; Zaleski, A.; Schreyer, K. Effects of 1α-hydroxyvitamin D3 and 24R,25-dihydroxyvitamin D3 on bone remodeling. Calcif. Tissue Int. 1980, 31 (1), 209-213.

(24) Sömjen, D.; Sömjen, G. J.; Weisman, Y.; Binderman, I. Evidence for 24,25-dihydroxycholecalciferol receptors in long bones of newborn rats. Biochem. J. 1982, 204 (1), 31-36.

(25) Seo, E.-G.; Einhorn, T. A.; Norman, A. W. 24R,25-Dihydroxyvitamin D3: An Essential Vitamin D3 Metabolite for Both Normal Bone Integrity and Healing of Tibial Fracture in Chicks*. Endocrinology 1997, 138 (9), 3864-3872.

(26) Christakos, S. Mechanism of action of 1,25-dihydroxyvitamin D3 on intestinal calcium absorption. Rev. Endocr. Metab. Disord. 2012, 13 (1), 39-44.

(27) Fleet, J. C.; Bradley, J.; Reddy, G. S.; Ray, R.; Wood, R. J. 1α,25-(OH)2-Vitamin D3Analogs with Minimalin VivoCalcemic Activity Can Stimulate Significant Transepithelial Calcium Transport and mRNA Expressionin Vitro. Arch. Biochem. Biophys. 1996, 329 (2), 228-234.

(28) DeLuca, H. F.; Plum, L. A.; Clagett-Dame, M. Selective analogs of 1α,25-dihydroxyvitamin D3 for the study of specific functions of Vitamin D. J. Steroid Biochem. Mol. Biol. 2007, 103 (3), 263-268.

(29) Tanaka, Y.; Frank, H.; Deluca, H. F. Biological Activity of 1,25-Dihydroxyvitamin D3 in the Rat1. Endocrinology 1973, 92 (2), 417-422.

(30) Kumar, R. The Metabolism of 1,25-Dihydroxyvitamin D3*. Endocr. Rev. 1980, 1 (3), 258-267.

(31) Norman, A. W.; Roth, J.; Orci, L. The Vitamin D Endocrine System: Steroid Metabolism, Hormone Receptors, and Biological Response (Calcium Binding Proteins)*. Endocr. Rev. 1982, 3 (4), 331-366.

(32) Fleet, J. C.; Schoch, R. D. Molecular mechanisms for regulation of intestinal calcium absorption by vitamin D and other factors. Crit. Rev. Clin. Lab. Sci. 2010, 47 (4), 181-195.

(33) Crawford, B. A.; Labio, E. D.; Strasser, S. I.; McCaughan, G. W. Vitamin D replacement for cirrhosis-related bone disease. Nat Clin Pract Gastroenterol Hepatol 2006, 3 (12), 689-699.

(34) Muller, M. J.; Volmer, D. A. Mass spectrometric profiling of vitamin D metabolites beyond 25-hydroxyvitamin D. Clin. Chem. 2015, 61 (8), 1033-1048.

(35) Heijboer, A. C.; Blankenstein, M. A.; Kema, I. P.; Buijs, M. M. Accuracy of 6 routine 25-hydroxyvitamin D assays: influence of vitamin D binding protein concentration. Clin. Chem. 2012, 58 (3), 543-548.

(36) Cavalier, E.; Lukas, P.; Crine, Y.; Peeters, S.; Carlisi, A.; Le Goff, C.; Gadisseur, R.; Delanaye, P.; Souberbielle, J.-C. Evaluation of automated immunoassays for 25(OH)-vitamin D determination in different critical populations before and after standardization of the assays. Clin. Chim. Acta 2014, 431, 60-65.

(37) van den Ouweland, J. M. W. Analysis of vitamin D metabolites by liquid chromatography-tandem mass spectrometry. TrAC, Trends Anal. Chem. 2016, 84, 117-130.

(38) Wang, Z.; Senn, T.; Kalhorn, T.; Zheng, X. E.; Zheng, S.; Davis, C. L.; Hebert, M. F.; Lin, Y. S.; Thummel, K. E. Simultaneous measurement of plasma vitamin D3 metabolites, including 4β,25-dihydroxyvitamin D3, using liquid chromatography–tandem mass spectrometry. Anal. Biochem. 2011, 418 (1), 126-133.

(39) Shah, I.; James, R.; Barker, J.; Petroczi, A.; Naughton, D. P. Misleading measures in Vitamin D analysis: a novel LC-MS/MS assay to account for epimers and isobars. Nutr. J. 2011, 10, 46-46.

(40) Satoh, M.; Ishige, T.; Ogawa, S.; Nishimura, M.; Matsushita, K.; Higashi, T.; Nomura, F. Development and validation of the simultaneous measurement of four vitamin D metabolites in serum by LC–MS/MS for clinical laboratory applications. Anal. Bioanal. Chem. 2016, 408 (27), 7617-7627.

(41) Liebisch, G.; Matysik, S. Accurate and reliable quantification of 25-hydroxy-vitamin D species by liquid chromatography high-resolution tandem mass spectrometry. J. Lipid Res. 2015, 56 (6), 1234-1239.

(42) Stepman, H. C. M.; Vanderroost, A.; Van Uytfanghe, K.; Thienpont, L. M. Candidate Reference Measurement Procedures for Serum 25-Hydroxyvitamin D3 and 25-Hydroxyvitamin D2 by Using Isotope-Dilution Liquid Chromatography–Tandem Mass Spectrometry. Clin. Chem. 2011, 57 (3), 441-448.

(43) Newman, M. S.; Brandon, T. R.; Groves, M. N.; Gregory, W. L.; Kapur, S.; Zava, D. T. A liquid chromatography/tandem mass spectrometry method for determination of 25-hydroxy vitamin D2 and 25-hydroxy vitamin D3 in dried blood spots: a potential adjunct to diabetes and cardiometabolic risk screening. J. Diabetes Sci. Technol. 2009, 3 (1), 156-162.

(44) Volmer, D. A.; Mendes, L. R. B. C.; Stokes, C. S. Analysis of vitamin D metabolic markers by mass spectrometry: Current techniques, limitations of the "gold standard" method, and anticipated future directions. Mass Spectrom. Rev. 2015, 34 (1), 2-23.

(45) Yang, M.-Y.; Huang, C.-Y.; Chiu, T. H. T.; Chang, K.-C.; Lin, M.-N.; Chen, L.-Y.; Hu, A. Using gas chromatography and mass spectrometry to determine 25-hydroxyvitamin D levels for clinical assessment of vitamin D deficiency. J. Food Drug Anal. 2019, 27 (2), 494-501.

(46) Aronov, P. A.; Hall, L. M.; Dettmer, K.; Stephensen, C. B.; Hammock, B. D. Metabolic profiling of major vitamin D metabolites using Diels–Alder derivatization and ultra-performance liquid chromatography–tandem mass spectrometry. Anal. Bioanal. Chem. 2008, 391 (5), 1917.

(47) Teegarden, M. D.; Riedl, K. M.; Schwartz, S. J. Chromatographic separation of PTAD-derivatized 25-hydroxyvitamin D3 and its C-3 epimer from human serum and murine skin. J. Chromatogr. B 2015, 991, 118-121.

(48) Ogawa, S.; Ooki, S.; Shinoda, K.; Higashi, T. Analysis of urinary vitamin D3 metabolites by liquid chromatography/tandem mass spectrometry with ESI-enhancing and stable isotope-coded derivatization. Anal. Bioanal. Chem. 2014, 406 (26), 6647-6654.

(49) Hedman, C. J.; Wiebe, D. A.; Dey, S.; Plath, J.; Kemnitz, J. W.; Ziegler, T. E. Development of a sensitive LC/MS/MS method for vitamin D metabolites: 1,25 Dihydroxyvitamin D2&3 measurement using a novel derivatization agent. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 2014, 953-954, 62-67.

(50) Bonnet, L.; Margier, M.; Svilar, L.; Couturier, C.; Reboul, E.; Martin, J.-C.; Landrier, J.-F.; Defoort, C. Simple Fast Quantification of Cholecalciferol, 25-Hydroxyvitamin D and 1,25-Dihydroxyvitamin D in Adipose Tissue Using LC-HRMS/MS. Nutrients 2019, 11 (9), 1977.

(51) Coldwell, R. D.; Porteous, C. E.; Trafford, D. J. H.; Makin, H. L. J. Gas chromatography—mass spectrometry and the measurement of vitamin D metabolites in human serum or plasma. Steroids 1987, 49 (1), 155-196.

(52) Ding, S.; Schoenmakers, I.; Jones, K.; Koulman, A.; Prentice, A.; Volmer, D. A. Quantitative determination of vitamin D metabolites in plasma using UHPLC-MS/MS. Anal. Bioanal. Chem. 2010, 398 (2), 779-789.

(53) Qi, Y.; Müller, M. J.; Volmer, D. A. Activation of Reactive MALDI Adduct Ions Enables Differentiation of Dihydroxylated Vitamin D Isomers. J. Am. Soc. Mass Spectrom. 2017, 28 (12), 2532-2537.

(54) Chouinard, C. D.; Cruzeiro, V. W. D.; Beekman, C. R.; Roitberg, A. E.; Yost, R. A. Investigating Differences in Gas-Phase Conformations of 25-Hydroxyvitamin D3 Sodiated Epimers using Ion Mobility-Mass Spectrometry and Theoretical Modeling. J. Am. Soc. Mass Spectrom. 2017, 28 (8), 1497-1505.

(55) Chouinard, C. D.; Cruzeiro, V. W. D.; Kemperman, R. H. J.; Oranzi, N. R.; Roitberg, A. E.; Yost, R. A. Cation-dependent conformations in 25-hydroxyvitamin D3-cation adducts measured by ion mobility-mass spectrometry and theoretical modeling. Int. J. Mass spectrom. 2018, 432, 1-8.

(56) Young, D. C.; Vouros, P.; Holick, M. F.; Higuchi, T. Collisionally induced dissociation in the study of A-ring hydroxylated vitamin D type compounds. Anal. Chem. 1992, 64 (8), 837-842.

(57) R Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. J. Am. Soc. Mass Spectrom. 2017, 28 (9), 1823-1826.

4.7 Supplementary Information

Table S4. 1 Peak assignment table for the CAD MS/MS of the protonated 1,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_9H_{11}	2061866	1077261	119.085527	119.085441	-0.72
	C_9H_9O	5472289	947214	133.064791	133.064727	-0.48
AD	$C_9H_{11}O$	23938552	920511	135.080441	135.080393	-0.36
	$C_9H_{13}O$	15070955	912936	137.096091	137.096044	-0.34
C	$C_8H_{11}O_2$	1983347	950057	139.075356	139.07532	-0.26
	$C_{11}H_{13}$	5323683	868950	145.101177	145.10113	-0.32
AE	$C_{10}H_{11}O$	1526401	949151	147.080441	147.080399	-0.29
	$C_{10}H_{13}O$	9244278	824438	149.096091	149.096066	-0.17
	$C_9H_{11}O_2$	25602036	833988	151.075356	151.075332	-0.16
	$C_{10}H_{15}O$	2383356	878831	151.111742	151.111715	-0.18
D	$C_9H_{12}O_2$	2293053	825058	152.083181	152.083144	-0.24
	$C_9H_{13}O_2$	22569352	820295	153.091006	153.090978	-0.18
	$C_{12}H_{15}$	9021058	803599	159.116827	159.116812	-0.09
AF	$C_{11}H_{13}O$	3294005	801681	161.096091	161.096086	-0.03
	$C_{11}H_{15}O$	2908662	829572	163.111742	163.111731	-0.07
E	$C_{10}H_{12}O_2$	3280050	774403	165.091006	165.090998	-0.04
	$C_{13}H_{15}$	10664191	749222	171.116827	171.116816	-0.06
	$C_{13}H_{17}$	12257985	716753	173.132477	173.132466	-0.06
AG	$C_{12}H_{15}O$	4327030	729898	175.111742	175.111746	0.02
	$C_{13}H_{19}$	11398270	722067	175.148127	175.148116	-0.06
	$C_{13}H_{21}$	13291060	715052	177.163777	177.163786	0.05
	$C_{13}H_{23}$	9741283	695438	179.179427	179.179421	-0.03
	$C_{14}H_{17}$	15953049	681392	185.132477	185.132492	0.08
	$C_{14}H_{19}$	13559317	672381	187.148127	187.148127	0.00
AH	$C_{13}H_{17}O$	9007994	670989	189.127392	189.127389	-0.02
	$C_{14}H_{21}$	20266880	670093	189.163777	189.163776	-0.01
	$C_{14}H_{23}$	22655194	652927	191.179427	191.179424	-0.02
	$C_{12}H_{17}O_2$	2021912	643289	193.122306	193.122326	0.10
	$C_{14}H_{25}$	6070819	667615	193.195077	193.195111	0.18
	$C_{15}H_{17}$	13907062	653713	197.132477	197.132507	0.15
	$C_{15}H_{19}$	26645904	632977	199.148127	199.148148	0.11
	$C_{15}H_{21}$	16617116	623389	201.163777	201.163789	0.06
	$C_{15}H_{23}$	21587866	621190	203.179427	203.179429	0.01
CO	$C_{15}H_{25}$	9506955	611478	205.195077	205.195091	0.07
	$C_{16}H_{17}$	11917886	609393	209.132477	209.132477	0.00
	$C_{16}H_{19}$	19204876	594058	211.148127	211.148169	0.20

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{16}H_{21}$	25816014	592531	213.163777	213.163809	0.15
	$C_{16}H_{23}$	19586184	584901	215.179427	215.17945	0.11
	$C_{16}H_{25}$	27176248	579544	217.195077	217.195081	0.02
CP	$C_{16}H_{27}$	23883234	574684	219.210727	219.210734	0.03
	$C_{16}H_{19}O$	33476658	552356	227.143042	227.143073	0.14
	$C_{17}H_{23}$	34156596	552632	227.179427	227.179442	0.07
AI	$C_{16}H_{21}O$	17477818	544501	229.158692	229.158702	0.04
	$C_{17}H_{25}$	24039612	542390	229.195077	229.195124	0.21
	$C_{16}H_{23}O$	15388990	537950	231.174342	231.174342	0.00
AJ	$C_{17}H_{23}O$	26235928	518740	243.174342	243.174335	-0.03
	$C_{18}H_{27}$	81663000	515083	243.210727	243.210792	0.27
	$C_{18}H_{29}$	201938064	509890	245.226377	245.226437	0.24
I	$C_{16}H_{23}O_2$	19768582	504654	247.169256	247.169277	0.08
D'Q	$C_{18}H_{31}$	205870352	507123	247.242027	247.242063	0.15
	$C_{19}H_{27}$	55498488	493964	255.210727	255.210728	0.00
	$C_{19}H_{29}$	75192184	488940	257.226377	257.226408	0.12
	$C_{19}H_{31}$	147085312	484034	259.242027	259.242058	0.12
J	$C_{17}H_{25}O_2$	12384383	480345	261.184906	261.184876	-0.11
	$C_{18}H_{29}O$	6879361	470525	261.221292	261.221314	0.08
	$C_{18}H_{31}O$	9625863	469322	263.236942	263.236988	0.17
	$C_{19}H_{23}O$	5939731	420864	267.174342	267.17431	-0.12
	$C_{19}H_{25}O$	166222496	466510	269.189992	269.19006	0.25
AL	$C_{19}H_{27}O$	97180456	463452	271.205642	271.205654	0.04
B'Q	$C_{20}H_{33}$	84089784	458038	273.257677	273.257724	0.17
	$C_{18}H_{27}O_2$	11697216	450225	275.200557	275.200592	0.13
	$C_{21}H_{29}$	115039728	447404	281.226377	281.226436	0.21
	$C_{21}H_{31}$	46259840	443816	283.242027	283.24206	0.12
	$C_{20}H_{29}O$	48783116	442013	285.221292	285.221289	-0.01
L	$C_{19}H_{27}O_2$	332488608	436153	287.200557	287.20064	0.29
	$C_{19}H_{29}O_2$	36806704	434029	289.216207	289.216236	0.10
	$C_{22}H_{29}$	52945236	429591	293.226377	293.226361	-0.05
	$C_{22}H_{31}$	16298472	425753	295.242027	295.242021	-0.02
	$C_{22}H_{33}$	18879100	418298	297.257677	297.257696	0.06
AM	$C_{21}H_{31}O$	84763408	420850	299.236942	299.236906	-0.12
	$C_{22}H_{35}$	14971663	406694	299.273328	299.273275	-0.18
	$C_{23}H_{31}$	144401744	408611	307.242027	307.242037	0.03
	$C_{23}H_{33}$	42034664	407296	309.257677	309.257699	0.07
	$C_{23}H_{35}$	13117050	389140	311.273328	311.273278	-0.16
AN	$C_{22}H_{33}O$	4190985	394200	313.252592	313.252551	-0.13
M	$C_{21}H_{31}O_2$	26888088	399910	315.231857	315.231733	-0.39
	$C_{24}H_{35}$	69429728	390699	323.273328	323.273317	-0.03

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{24}H_{37}$	21056110	378065	325.288978	325.288971	-0.02
AO	$C_{23}H_{35}O$	26847998	382999	327.268242	327.268124	-0.36
N	$C_{22}H_{33}O_2$	10656654	365106	329.247507	329.247367	-0.43
	$C_{24}H_{35}O$	29355636	365382	339.268242	339.268127	-0.34
AP	$C_{24}H_{37}O$	18161420	358362	341.283892	341.283732	-0.47
O	$C_{23}H_{35}O_2$	41643940	369087	343.263157	343.263035	-0.36
P	$C_{24}H_{37}O_2$	6802424	332705	357.278807	357.2786	-0.58
$[M+H]^+$ - $3H_2O$	$C_{27}H_{39}$	1894413824	343871	363.304628	363.304715	0.24
$[M+H]^+$ - $2H_2O$	$C_{27}H_{41}O$	3714219008	327414	381.315192	381.315269	0.20
$[M+H]^+$ - H_2O	$C_{27}H_{43}O_2$	4040327936	313373	399.325757	399.325759	0.01
$[M+H]^+$	$C_{27}H_{44}O_3$	66586640	303294	417.336322	417.336034	-0.69
Average error						0.04
Absolute average error						0.16
Standard deviation						0.15

Table S4. 2 Peak assignment table for the CAD MS/MS of the protonated 24,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
AB	C_9H_{11}	1869928	1071496	119.085527	119.085556	0.24
	C_9H_{13}	12396664	1061040	121.101177	121.101208	0.26
	$C_8H_{13}O$	1181563	1089584	125.096091	125.096129	0.30
AC	$C_{10}H_{11}$	1677959	937313	131.085527	131.085554	0.21
	$C_{10}H_{13}$	2038017	889944	133.101177	133.101211	0.26
AD	$C_{11}H_{13}$	6366844	893308	145.101177	145.101194	0.12
	$C_{11}H_{15}$	6961119	883805	147.116827	147.116847	0.14
	$C_{11}H_{17}$	6275924	882670	149.132477	149.132509	0.21
AE	$C_{12}H_{15}$	12733265	829922	159.116827	159.11685	0.14
	$C_{13}H_{15}$	5637053	780588	171.116827	171.116848	0.12
AF	$C_{13}H_{17}$	14657403	741502	173.132477	173.132494	0.10
	$C_{13}H_{19}$	22800692	729735	175.148127	175.148134	0.04
	$C_{14}H_{15}$	2614706	648983	183.116827	183.116852	0.14
	$C_{14}H_{17}$	10953016	670327	185.132477	185.132503	0.14
	$C_{14}H_{19}$	21418670	664199	187.148127	187.148127	0.00
	$C_{14}H_{21}$	21595156	662008	189.163777	189.163778	0.01
	$C_{15}H_{19}$	27630082	618707	199.148127	199.148133	0.03
	$C_{15}H_{21}$	34540292	609773	201.163777	201.163772	-0.02
AG	$C_{16}H_{21}$	37326840	573471	213.163777	213.163774	-0.01

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{16}H_{23}$	34091172	561218	215.179427	215.179413	-0.07
	$C_{17}H_{21}$	16966516	541964	225.163777	225.163773	-0.02
AH	$C_{17}H_{23}$	37188592	540647	227.179427	227.179404	-0.10
	$C_{17}H_{25}$	28857444	532460	229.195077	229.195077	0.00
	$C_{16}H_{25}O$	11507003	525096	233.189992	233.189972	-0.09
	$C_{18}H_{23}$	30439008	514971	239.179427	239.179425	-0.01
	$C_{18}H_{25}$	84186816	512230	241.195077	241.19504	-0.15
	$C_{18}H_{27}$	100554520	509859	243.210727	243.210708	-0.08
	$C_{18}H_{29}$	39622516	501056	245.226377	245.226365	-0.05
	$C_{17}H_{27}O$	10186701	488821	247.205642	247.205644	0.01
	$C_{19}H_{23}$	22505600	492517	251.179427	251.179396	-0.12
	$C_{19}H_{25}$	80828640	490115	253.195077	253.19503	-0.19
	$C_{19}H_{27}$	142137664	489376	255.210727	255.210663	-0.25
	$C_{18}H_{27}O$	70392304	478417	259.205642	259.205622	-0.08
	$C_{18}H_{29}O$	79927888	473736	261.221292	261.221258	-0.13
I	$C_{19}H_{27}O$	65494068	458462	271.205642	271.205596	-0.17
	$C_{19}H_{29}O$	174578832	457433	273.221292	273.221247	-0.16
	$C_{21}H_{27}$	52501960	446115	279.210727	279.210707	-0.07
	$C_{21}H_{29}$	70016560	443005	281.226377	281.226357	-0.07
	$C_{20}H_{27}O$	2186903	449643	283.205642	283.20561	-0.11
	$C_{21}H_{31}$	95444632	443805	283.242027	283.242002	-0.09
	$C_{20}H_{29}O$	73539328	439120	285.221292	285.221254	-0.13
	$C_{20}H_{31}O$	53714796	436218	287.236942	287.236927	-0.05
	$C_{20}H_{33}O$	5771226	430141	289.252592	289.252551	-0.14
	$C_{22}H_{31}$	339894560	424902	295.242027	295.241989	-0.13
J	$C_{21}H_{31}O$	73474568	418451	299.236942	299.236895	-0.16
	$C_{21}H_{33}O$	24789626	414105	301.252592	301.252594	0.01
	$C_{21}H_{35}O$	38273776	410898	303.268242	303.268233	-0.03
	$C_{23}H_{31}$	75506648	409239	307.242027	307.241997	-0.10
	$C_{23}H_{33}$	60367956	403800	309.257677	309.257678	0.00
	$C_{23}H_{35}$	8760529	404676	311.273328	311.273337	0.03
K	$C_{22}H_{33}O$	30573900	398151	313.252592	313.252589	-0.01
	$C_{24}H_{33}$	71788360	389017	321.257677	321.257667	-0.03
	$C_{24}H_{35}$	147974096	388923	323.273328	323.273313	-0.05
	$C_{24}H_{37}$	1444946	425976	325.288978	325.288984	0.02
L	$C_{23}H_{35}O$	4925655	389572	327.268242	327.26822	-0.07
	$C_{24}H_{35}O$	23751706	369922	339.268242	339.268238	-0.01
	$C_{24}H_{37}O$	119270576	368247	341.283892	341.283874	-0.05
M	$C_{24}H_{37}O_2$	2125161	366485	357.278807	357.278835	0.08
$[M+H]^+$ - $3H_2O$	$C_{27}H_{39}$	1050571584	346006	363.304628	363.304579	-0.13

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
$[M+H]^+$ - $2H_2O$	$C_{27}H_{41}O$	3261233408	329163	381.315192	381.315092	-0.26
$[M+H]^+$ - H_2O	$C_{27}H_{43}O_2$	640581440	313847	399.325757	399.325756	0.00
$[M+H]^+$	$C_{27}H_{44}O_3$	43943192	301238	417.336322	417.336392	0.17
Average error						0.01
Absolute average error						0.10
Standard deviation						0.08

Table S4. 3 Peak assignment table for the IRMPD MS/MS of the protonated 1,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_{15}	4141042	1223797	111.116827	111.116828	0.01
	C_9H_9	3162382	1207595	117.069877	117.069883	0.05
AB	C_9H_{11}	9180172	1110064	119.085527	119.085527	0.00
	C_8H_9O	4794784	1067511	121.064791	121.064792	0.01
	C_9H_{13}	9434145	1027320	121.101177	121.101177	0.00
	$C_8H_{11}O$	1227280	1198025	123.080441	123.080442	0.01
	C_9H_{15}	13603153	1043694	123.116827	123.116833	0.05
	$C_7H_{10}O_2$	22377998	988679	127.075356	127.075359	0.02
	C_9H_9O	3407312	893098	133.064791	133.064785	-0.05
AD	$C_9H_{11}O$	61992800	912006	135.080441	135.080441	0.00
	$C_9H_{13}O$	14729875	886280	137.096091	137.09609	-0.01
C	$C_8H_{11}O_2$	21955042	880931	139.075356	139.075355	-0.01
	$C_{11}H_{13}$	13500047	849301	145.101177	145.101166	-0.08
AE	$C_{10}H_{11}O$	2899916	841614	147.080441	147.080426	-0.10
	$C_{10}H_{13}O$	7497256	821052	149.096091	149.09609	-0.01
	$C_9H_{11}O_2$	35243024	814615	151.075356	151.075353	-0.02
D	$C_9H_{12}O_2$	6456592	830834	152.083181	152.083173	-0.05
	$C_9H_{13}O_2$	42139156	808019	153.091006	153.090996	-0.07
	$C_{12}H_{11}$	3405859	799996	155.085527	155.085512	-0.10
	$C_{12}H_{13}$	10882631	812535	157.101177	157.101162	-0.10
	$C_{12}H_{15}$	17157238	782252	159.116827	159.116821	-0.04
AF	$C_{11}H_{13}O$	3969191	760310	161.096091	161.096092	0.01
	$C_{12}H_{17}$	16109234	771371	161.132477	161.132468	-0.06
	$C_{11}H_{15}O$	5644011	764213	163.111742	163.111733	-0.06
	$C_{12}H_{19}$	44210420	765603	163.148127	163.148117	-0.06
E	$C_{10}H_{12}O_2$	7188778	770119	165.091006	165.091003	-0.02

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{13}H_{15}$	22198802	734421	171.116827	171.116815	-0.07
	$C_{13}H_{17}$	21697632	716744	173.132477	173.132466	-0.06
AG	$C_{12}H_{15}O$	4840097	721043	175.111742	175.111733	-0.05
	$C_{13}H_{21}$	39102704	708411	177.163777	177.163773	-0.02
	$C_{14}H_{13}$	2501446	757234	181.101177	181.101175	-0.01
	$C_{14}H_{15}$	14540652	688350	183.116827	183.116816	-0.06
	$C_{14}H_{17}$	29294980	675596	185.132477	185.132474	-0.02
	$C_{14}H_{19}$	20597644	668850	187.148127	187.148114	-0.07
AH	$C_{13}H_{17}O$	11246969	677449	189.127392	189.12738	-0.06
	$C_{15}H_{15}$	5799643	655948	195.116827	195.116805	-0.11
	$C_{15}H_{17}$	19758214	633614	197.132477	197.132476	-0.01
	$C_{15}H_{19}$	35409440	626339	199.148127	199.148122	-0.03
	$C_{15}H_{21}$	21475238	622075	201.163777	201.163768	-0.04
	$C_{15}H_{23}$	30118174	614343	203.179427	203.179413	-0.07
CO	$C_{15}H_{25}$	8950406	609094	205.195077	205.195067	-0.05
	$C_{16}H_{17}$	10781976	602120	209.132477	209.132468	-0.04
	$C_{16}H_{19}$	22222418	594429	211.148127	211.148132	0.02
	$C_{16}H_{21}$	33290112	583929	213.163777	213.163777	0.00
	$C_{16}H_{23}$	25520288	578839	215.179427	215.179424	-0.01
	$C_{16}H_{25}$	34314676	574544	217.195077	217.195065	-0.06
CP	$C_{16}H_{27}$	15055544	564836	219.210727	219.210712	-0.07
	$C_{17}H_{19}$	10728598	560044	223.148127	223.148113	-0.06
	$C_{17}H_{21}$	31931770	549860	225.163777	225.163775	-0.01
	$C_{16}H_{19}O$	11527761	547202	227.143042	227.143041	0.00
AI	$C_{16}H_{21}O$	14839585	543657	229.158692	229.158679	-0.06
	$C_{17}H_{25}$	26799908	539293	229.195077	229.195084	0.03
	$C_{16}H_{23}O$	20008936	539591	231.174342	231.174326	-0.07
	$C_{17}H_{27}$	34766828	537008	231.210727	231.210734	0.03
	$C_{17}H_{29}$	13113514	527816	233.226377	233.226389	0.05
	$C_{18}H_{21}$	28041734	522105	237.163777	237.163785	0.03
	$C_{18}H_{23}$	35480236	518591	239.179427	239.179433	0.03
	$C_{18}H_{25}$	31261516	518563	241.195077	241.195069	-0.03
AJ	$C_{17}H_{23}O$	21914594	513123	243.174342	243.174322	-0.08
	$C_{18}H_{27}$	58065892	510607	243.210727	243.210737	0.04
	$C_{18}H_{29}$	156288128	507276	245.226377	245.226381	0.02
I	$C_{16}H_{22}O_2$	24454404	499315	247.169256	247.169252	-0.02
D'Q	$C_{18}H_{31}$	149341456	502566	247.242027	247.242018	-0.04
	$C_{19}H_{25}$	90932888	491510	253.195077	253.195064	-0.05
	$C_{19}H_{29}$	38602640	480314	257.226377	257.226373	-0.02
	$C_{19}H_{31}$	68454408	478619	259.242027	259.242016	-0.04
J	$C_{17}H_{24}O_2$	10025082	478826	261.184906	261.184886	-0.08

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{18}H_{29}O$	10991740	476765	261.221292	261.221295	0.01
	$C_{18}H_{31}O$	24353010	468928	263.236942	263.236958	0.06
AL	$C_{19}H_{27}O$	82430648	457319	271.205642	271.20563	-0.04
B'Q	$C_{20}H_{33}$	89410344	452369	273.257677	273.25768	0.01
	$C_{18}H_{26}O_2$	8926107	443503	275.200557	275.20057	0.05
	$C_{21}H_{27}$	33991812	436608	279.210727	279.210731	0.01
	$C_{21}H_{29}$	39603452	434783	281.226377	281.226386	0.03
	$C_{21}H_{31}$	23899508	427873	283.242027	283.242028	0.00
L	$C_{19}H_{27}O_2$	42282604	428683	287.200557	287.200563	0.02
	$C_{19}H_{29}O_2$	28569328	420913	289.216207	289.21622	0.04
	$C_{22}H_{31}$	6949008	429052	295.242027	295.242035	0.03
AM	$C_{21}H_{31}O$	20066750	404205	299.236942	299.236911	-0.10
	$C_{23}H_{31}$	73815136	397661	307.242027	307.242011	-0.05
	$C_{23}H_{33}$	28136736	393719	309.257677	309.257688	0.04
AN	$C_{22}H_{33}O$	1278894	331561	313.252592	313.252607	0.05
M	$C_{21}H_{31}O_2$	8755069	374332	315.231857	315.231813	-0.14
	$C_{21}H_{33}O_2$	9345112	390307	317.247507	317.247525	0.06
	$C_{24}H_{33}$	19081760	369762	321.257677	321.257666	-0.03
	$C_{24}H_{35}$	36993808	371881	323.273328	323.273328	0.00
	$C_{24}H_{37}$	17722374	368194	325.288978	325.288993	0.05
AO	$C_{23}H_{35}O$	15183102	363428	327.268242	327.268218	-0.07
N	$C_{22}H_{33}O_2$	3162620	388905	329.247507	329.247431	-0.23
	$C_{25}H_{37}$	21881404	353321	337.288978	337.288963	-0.04
	$C_{24}H_{35}O$	13845332	350171	339.268242	339.268217	-0.07
AP	$C_{24}H_{37}O$	13186168	342876	341.283892	341.283849	-0.13
O	$C_{23}H_{35}O_2$	7857562	348589	343.263157	343.263133	-0.07
P	$C_{24}H_{37}O_2$	2128920	356065	357.278807	357.278835	0.08
$[M+H]^+ - 3H_2O$	$C_{27}H_{39}$	827027968	341771	363.304628	363.30462	-0.02
$[M+H]^+ - 2H_2O$	$C_{27}H_{41}O$	2130935296	326424	381.315192	381.315202	0.03
$[M+H]^+ - H_2O$	$C_{27}H_{43}O_2$	5138254336	314268	399.325757	399.325692	-0.16
$[M+H]^+$	$C_{27}H_{44}O_3$	263033312	294122	417.336322	417.336321	0.00
Average error						0.03
Absolute average error						0.05
Standard deviation						0.04

Table S4. 4 Peak assignment table for the IRMPD MS/MS of the protonated 24,25 dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_{13}	32794890	1196925	109.101177	109.101222	0.41
AB	C_9H_{11}	12499045	1054654	119.085527	119.085564	0.31
	C_9H_{13}	144448384	1052295	121.101177	121.101209	0.26
	$C_8H_{13}O$	39061508	994911	125.096091	125.096125	0.27
	$C_8H_{15}O$	649833600	988496	127.111742	127.111758	0.13
AC	$C_{10}H_{11}$	14868465	922309	131.085527	131.085566	0.30
	$C_{10}H_{13}$	13019374	945840	133.101177	133.101205	0.21
	$C_{10}H_{15}$	61624340	917892	135.116827	135.116853	0.19
	$C_{11}H_{11}$	3722737	850705	143.085527	143.085542	0.10
AD	$C_{11}H_{13}$	34826168	853145	145.101177	145.101195	0.12
	$C_{11}H_{15}$	28377496	841503	147.116827	147.116845	0.12
	$C_{11}H_{17}$	46254992	821634	149.132477	149.132502	0.17
	$C_{12}H_{11}$	3924427	833354	155.085527	155.085545	0.12
	$C_{12}H_{13}$	23462916	796546	157.101177	157.101189	0.08
AE	$C_{12}H_{15}$	44738120	776744	159.116827	159.116842	0.09
	$C_{12}H_{17}$	64866452	775570	161.132477	161.132488	0.07
	$C_{12}H_{19}$	38963944	766225	163.148127	163.14814	0.08
	$C_{13}H_{13}$	9758154	744709	169.101177	169.101186	0.05
	$C_{13}H_{15}$	40191004	725762	171.116827	171.116834	0.04
AF	$C_{13}H_{17}$	52295268	721671	173.132477	173.132484	0.04
	$C_{13}H_{19}$	80894112	713751	175.148127	175.14813	0.02
	$C_{14}H_{17}$	55722736	671054	185.132477	185.132485	0.04
	$C_{14}H_{19}$	69469384	666583	187.148127	187.148127	0.00
	$C_{14}H_{21}$	66283140	662398	189.163777	189.163776	-0.01
	$C_{15}H_{19}$	99904936	626481	199.148127	199.148128	0.01
	$C_{15}H_{21}$	72232656	621101	201.163777	201.163775	-0.01
AG	$C_{16}H_{21}$	111981648	585253	213.163777	213.163777	0.00
	$C_{16}H_{25}$	43210168	573491	217.195077	217.195068	-0.04
	$C_{17}H_{21}$	62961672	553554	225.163777	225.163775	-0.01
AH	$C_{17}H_{23}$	106985616	550762	227.179427	227.179418	-0.04
	$C_{17}H_{25}$	77232400	542255	229.195077	229.195079	0.01
	$C_{16}H_{25}O$	24463878	537023	233.189992	233.189983	-0.04
	$C_{18}H_{23}$	111393656	520174	239.179427	239.179428	0.00
	$C_{18}H_{25}$	209632240	517834	241.195077	241.195065	-0.05
	$C_{18}H_{27}$	141704320	512010	243.210727	243.21073	0.01
	$C_{18}H_{29}$	135478528	507500	245.226377	245.226376	0.00
	$C_{19}H_{25}$	211566448	491770	253.195077	253.195061	-0.06
	$C_{19}H_{27}$	282148896	487686	255.210727	255.210708	-0.07

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{19}H_{29}$	84312272	482227	257.226377	257.226366	-0.04
	$C_{18}H_{27}O$	81723792	481220	259.205642	259.205644	0.01
	$C_{18}H_{29}O$	54116944	475520	261.221292	261.22129	-0.01
	$C_{20}H_{27}$	176374960	465600	267.210727	267.21071	-0.06
	$C_{20}H_{29}$	169002384	462490	269.226377	269.226357	-0.07
I	$C_{19}H_{27}O$	88802920	458558	271.205642	271.205628	-0.05
	$C_{19}H_{29}O$	259344192	455485	273.221292	273.221287	-0.02
	$C_{21}H_{27}$	95635952	442722	279.210727	279.210725	-0.01
	$C_{21}H_{29}$	141118176	441552	281.226377	281.22638	0.01
	$C_{20}H_{29}O$	131976368	436108	285.221292	285.221282	-0.04
	$C_{20}H_{31}O$	61017504	430867	287.236942	287.236945	0.01
	$C_{20}H_{33}O$	6254222	436577	289.252592	289.252576	-0.06
	$C_{22}H_{31}$	365646144	420677	295.242027	295.242024	-0.01
J	$C_{21}H_{31}O$	81548064	413280	299.236942	299.236919	-0.08
	$C_{21}H_{33}O$	18985990	405109	301.252592	301.252601	0.03
	$C_{21}H_{35}O$	18855146	403445	303.268242	303.268234	-0.03
	$C_{23}H_{31}$	116414112	402309	307.242027	307.242013	-0.05
	$C_{23}H_{33}$	56039292	397078	309.257677	309.257688	0.04
K	$C_{22}H_{33}O$	13269278	396106	313.252592	313.252599	0.02
	$C_{24}H_{33}$	67525688	379972	321.257677	321.257679	0.01
	$C_{24}H_{35}$	139746560	379937	323.273328	323.273331	0.01
L	$C_{23}H_{35}O$	1291381	435373	327.268242	327.268255	0.04
	$C_{24}H_{35}O$	14357176	350233	339.268242	339.268244	0.01
	$C_{24}H_{37}O$	51864428	353125	341.283892	341.28389	-0.01
$[M+H]^+ - 3H_2O$	$C_{27}H_{39}$	996371264	339751	363.304628	363.304616	-0.03
$[M+H]^+ - 2H_2O$	$C_{27}H_{41}O$	2313384448	325828	381.315192	381.315177	-0.04
$[M+H]^+ - H_2O$	$C_{27}H_{43}O_2$	1206090240	308590	399.325757	399.325759	0.01
$[M+H]^+$	$C_{27}H_{44}O_3$	62603664	278062	417.336322	417.33637	0.12
Average error						0.04
Absolute average error						0.07
Standard deviation						0.09

Table S4. 5 Peak assignment table for the EID MS/MS of the protonated 1,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	21745554	1151374	105.069877	105.069877	0.00
AB	C_7H_9O	5964417	1105781	109.064791	109.064792	0.01
	C_9H_9	11830329	1049674	117.069877	117.06988	0.03
	C_9H_{11}	15875639	1043152	119.085527	119.085529	0.02
	C_8H_9O	3233805	1072666	121.064791	121.064791	0.00
B	$C_7H_{10}O_2$	4893909	1030521	127.075356	127.075358	0.02
	C_9H_9O	23503828	941431	133.064791	133.064793	0.02
AD	$C_9H_{11}O$	23789826	921376	135.080441	135.080445	0.03
	$C_9H_{13}O$	7391829	931763	137.096091	137.096097	0.04
C	$C_8H_{11}O_2$	6098119	899829	139.075356	139.075361	0.04
	$C_{11}H_{13}$	15155438	865594	145.101177	145.101182	0.03
AE	$C_{10}H_{11}O$	2339543	861627	147.080441	147.080436	-0.03
	$C_{10}H_{13}O$	3729634	866417	149.096091	149.0961	0.06
	$C_9H_{11}O_2$	34396920	831224	151.075356	151.075357	0.01
D	$C_9H_{12}O_2$	102794256	822435	152.083181	152.083181	0.00
	$C_9H_{13}O_2$	18203950	821412	153.091006	153.091007	0.01
	$C_{12}H_{15}$	14861349	785008	159.116827	159.116829	0.01
AF	$C_{11}H_{13}O$	3069574	776242	161.096091	161.096087	-0.02
	$C_{11}H_{15}O$	2252282	763080	163.111742	163.11173	-0.07
E	$C_{10}H_{12}O_2$	2562152	793103	165.091006	165.091009	0.02
	$C_{13}H_{15}$	12708566	738795	171.116827	171.116827	0.00
	$C_{13}H_{17}$	11684174	727635	173.132477	173.13248	0.02
AG	$C_{12}H_{15}O$	1432756	795431	175.111742	175.111754	0.07
	$C_{13}H_{19}$	8393663	722879	175.148127	175.148132	0.03
	$C_{13}H_{21}$	9359911	697506	177.163777	177.163776	-0.01
	$C_{13}H_{23}$	5438596	702218	179.179427	179.179425	-0.01
	$C_{14}H_{17}$	11805476	682507	185.132477	185.132479	0.01
	$C_{14}H_{19}$	7968060	679133	187.148127	187.148134	0.04
	$C_{13}H_{17}O$	3351859	707028	189.127392	189.127388	-0.02
	$C_{14}H_{21}$	8966972	660049	189.163777	189.163777	0.00
	$C_{14}H_{23}$	7812389	669199	191.179427	191.179423	-0.02
	$C_{14}H_{25}$	1819381	692395	193.195077	193.195078	0.01
	$C_{15}H_{17}$	8963647	649317	197.132477	197.132482	0.03
	$C_{15}H_{19}$	10753998	624941	199.148127	199.14813	0.02
	$C_{15}H_{21}$	6139208	647897	201.163777	201.163775	-0.01
	$C_{15}H_{23}$	6394542	625112	203.179427	203.179428	0.00
CO	$C_{15}H_{25}$	2119169	606449	205.195077	205.195085	0.04
	$C_{16}H_{17}$	6609508	597781	209.132477	209.13248	0.01

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{16}H_{19}$	6869892	615474	211.148127	211.148129	0.01
	$C_{16}H_{21}$	7071892	595982	213.163777	213.16378	0.01
	$C_{16}H_{23}$	4808084	589301	215.179427	215.179422	-0.02
	$C_{16}H_{25}$	6851207	591871	217.195077	217.195076	0.00
CP	$C_{16}H_{27}$	3515756	608398	219.210727	219.210726	0.00
	$C_{16}H_{19}O$	2532989	565614	227.143042	227.143049	0.03
	$C_{17}H_{23}$	6315648	549340	227.179427	227.179433	0.03
AI	$C_{16}H_{21}O$	2878762	536560	229.158692	229.158691	0.00
	$C_{17}H_{25}$	4610862	572738	229.195077	229.195075	-0.01
	$C_{16}H_{23}O$	3845072	575035	231.174342	231.174344	0.01
	$C_{18}H_{25}$	5171871	530008	241.195077	241.195059	-0.07
AJ	$C_{17}H_{23}O$	3675934	528060	243.174342	243.174345	0.01
	$C_{18}H_{27}$	11402016	520715	243.210727	243.210729	0.01
	$C_{18}H_{29}$	27061150	520987	245.226377	245.226381	0.02
I	$C_{16}H_{22}O_2$	4609046	519625	247.169256	247.169249	-0.03
D'Q	$C_{18}H_{31}$	32082970	504844	247.242027	247.242027	0.00
	$C_{19}H_{29}$	5105265	493262	257.226377	257.226377	0.00
	$C_{19}H_{31}$	6321896	483961	259.242027	259.24203	0.01
J	$C_{17}H_{24}O_2$	2180443	494168	261.184906	261.184897	-0.03
	$C_{18}H_{29}O$	1826141	541555	261.221292	261.22132	0.11
	$C_{18}H_{31}O$	8505301	484603	263.236942	263.236941	0.00
	$C_{19}H_{25}O$	10422584	472959	269.189992	269.189992	0.00
AL	$C_{19}H_{27}O$	12378546	464222	271.205642	271.205649	0.03
B'Q	$C_{20}H_{33}$	11559471	463605	273.257677	273.257673	-0.01
	$C_{21}H_{29}$	4465699	471764	281.226377	281.226371	-0.02
	$C_{21}H_{31}$	2718886	455237	283.242027	283.242009	-0.06
L	$C_{19}H_{27}O_2$	6839726	445370	287.200557	287.200554	-0.01
	$C_{19}H_{29}O_2$	13039160	434836	289.216207	289.216206	0.00
	$C_{22}H_{29}$	4446035	440516	293.226377	293.226373	-0.01
AM	$C_{21}H_{31}O$	3420436	448597	299.236942	299.236934	-0.03
	$C_{23}H_{31}$	8723357	415981	307.242027	307.242026	0.00
	$C_{23}H_{33}$	2063947	428409	309.257677	309.257723	0.15
M	$C_{21}H_{31}O_2$	2054762	401182	315.231857	315.231882	0.08
	$C_{24}H_{35}$	3794304	423870	323.273328	323.273321	-0.02
	$C_{23}H_{33}O$	3541580	430028	325.252592	325.252612	0.06
	$C_{24}H_{37}$	2174544	408656	325.288978	325.289004	0.08
AO	$C_{23}H_{35}O$	2199328	411714	327.268242	327.268246	0.01
N	$C_{22}H_{33}O_2$	2251236	384755	329.247507	329.247513	0.02
	$C_{24}H_{35}O$	1669210	402380	339.268242	339.268286	0.13
AP	$C_{24}H_{37}O$	1953095	399336	341.283892	341.283854	-0.11
O	$C_{23}H_{35}O_2$	1248819	413526	343.263157	343.263223	0.19

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
$[M+H]^{1+}$ - $3H_2O$	$C_{27}H_{39}$	102396960	346378	363.304628	363.304631	0.01
$[M+H]^{1+}$ - $2H_2O$	$C_{27}H_{41}O$	366968256	330378	381.315192	381.315188	-0.01
$[M+H]^{1+}$ - H_2O	$C_{27}H_{43}O_2$	1173587840	313586	399.325757	399.325738	-0.05
$[M+H]^{1+}$	$C_{27}H_{44}O_3$	350588896	299779	417.336322	417.336317	-0.01
Average error						0.01
Absolute average error						0.03
Standard deviation						0.03

Table S4. 6 Peak assignment table for the EID MS/MS of the protonated 24,25 dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	18844638	1145651	105.069877	105.069888	0.10
	C_9H_9	7289019	1045181	117.069877	117.069884	0.06
AB	C_9H_{11}	22662456	1035656	119.085527	119.085536	0.08
	$C_8H_{13}O$	22515680	979822	125.096091	125.096103	0.10
	$C_8H_{15}O$	134854240	955111	127.111742	127.11175	0.06
AC	$C_{10}H_{11}$	21878770	940167	131.085527	131.085534	0.05
	$C_{10}H_{13}$	21992704	915657	133.101177	133.101181	0.03
	$C_{10}H_{15}$	31024698	900174	135.116827	135.116834	0.05
	$C_9H_{12}O$	1864736	984297	136.088266	136.088268	0.01
	$C_{11}H_{11}$	10245764	853042	143.085527	143.085531	0.03
	$C_8H_{15}O_2$	123945096	856517	143.106656	143.106661	0.03
AD	$C_{11}H_{13}$	33323628	843867	145.101177	145.101181	0.03
	$C_{11}H_{15}$	26104944	824577	147.116827	147.116831	0.03
	$C_{11}H_{17}$	21757580	825650	149.132477	149.132481	0.03
	$C_{12}H_{13}$	17960866	788574	157.101177	157.101179	0.01
AE	$C_{12}H_{15}$	37380624	769875	159.116827	159.116828	0.01
	$C_{12}H_{17}$	35484812	768739	161.132477	161.132477	0.00
	$C_{12}H_{19}$	13694730	761122	163.148127	163.148129	0.01
	$C_{13}H_{15}$	21231864	713370	171.116827	171.116827	0.00
AF	$C_{13}H_{17}$	30528370	707578	173.132477	173.132477	0.00
	$C_{13}H_{19}$	34203104	707453	175.148127	175.148127	0.00
	$C_{14}H_{15}$	9874170	687268	183.116827	183.116826	-0.01
	$C_{14}H_{17}$	26177814	660014	185.132477	185.132476	-0.01
	$C_{14}H_{19}$	27771674	655968	187.148127	187.148127	0.00

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{14}H_{21}$	22107396	655496	189.163777	189.163779	0.01
	$C_{15}H_{19}$	34209000	622183	199.148127	199.148122	-0.03
	$C_{15}H_{21}$	26725944	621027	201.163777	201.163774	-0.01
	$C_{15}H_{23}$	10372977	605300	203.179427	203.179424	-0.01
AG	$C_{16}H_{21}$	30389346	578700	213.163777	213.163772	-0.02
	$C_{16}H_{23}$	21472560	575191	215.179427	215.179422	-0.02
	$C_{16}H_{25}$	8388079	579923	217.195077	217.195076	0.00
	$C_{17}H_{21}$	13865070	551834	225.163777	225.163784	0.03
AH	$C_{17}H_{23}$	21498104	549497	227.179427	227.179424	-0.01
	$C_{17}H_{25}$	15128954	540581	229.195077	229.19507	-0.03
	$C_{16}H_{25}O$	5815914	527677	233.189992	233.189989	-0.01
	$C_{18}H_{23}$	19471284	518914	239.179427	239.179429	0.01
	$C_{18}H_{25}$	35908640	512497	241.195077	241.195072	-0.02
	$C_{19}H_{27}$	48688972	484809	255.210727	255.210725	-0.01
	$C_{18}H_{27}O$	20942936	482884	259.205642	259.205635	-0.03
	$C_{18}H_{29}O$	21471460	481776	261.221292	261.221285	-0.03
I	$C_{19}H_{27}O$	23004112	464644	271.205642	271.205643	0.00
	$C_{19}H_{29}O$	41301104	457314	273.221292	273.221289	-0.01
	$C_{21}H_{29}$	19285780	441858	281.226377	281.226382	0.02
	$C_{20}H_{29}O$	24590454	438562	285.221292	285.221291	0.00
	$C_{20}H_{33}O$	3975654	444339	289.252592	289.252564	-0.10
	$C_{22}H_{31}$	54287388	425696	295.242027	295.24203	0.01
J	$C_{21}H_{31}O$	14200228	421688	299.236942	299.236951	0.03
	$C_{21}H_{33}O$	4520557	430555	301.252592	301.252611	0.06
	$C_{23}H_{33}$	10482078	403587	309.257677	309.25769	0.04
K	$C_{22}H_{33}O$	2728770	417195	313.252592	313.252566	-0.08
	$C_{24}H_{35}$	29329280	387861	323.273328	323.273336	0.02
L	$C_{24}H_{37}O_2$	4873906	379840	357.278807	357.278775	-0.09
	$C_{24}H_{35}O$	1655922	353743	339.268242	339.268249	0.02
$[M+H]^{1+}$ - $3H_2O$	$C_{27}H_{39}$	215382672	342927	363.304628	363.304639	0.03
$[M+H]^{1+}$ - $2H_2O$	$C_{27}H_{41}O$	1077642368	323776	381.315192	381.315187	-0.01
$[M+H]^{1+}$ - H_2O	$C_{27}H_{43}O_2$	1103859200	309031	399.325757	399.325755	-0.01
$[M+H]^{1+}$	$C_{27}H_{44}O_3$	8841828352	295556	417.336322	417.33614	-0.44
Average error						0.00
Absolute average error						0.04
Standard deviation						0.06

Table S4. 7 Peak assignment table for the 193 nm UVPD MS/MS of the protonated 1,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	13742154	1084022	105.069877	105.069877	0.00
AB	C_7H_9O	7533579	1062180	109.064791	109.064789	-0.02
	C_8H_{15}	5477098	1012224	111.116827	111.116824	-0.03
	C_9H_9	15969965	1046541	117.069877	117.069875	-0.02
	C_9H_{11}	55903916	1014701	119.085527	119.085522	-0.04
	C_8H_9O	5620035	1018142	121.064791	121.064787	-0.03
	C_9H_{13}	30283972	1002986	121.101177	121.101172	-0.04
	$C_8H_{11}O$	3166527	1049248	123.080441	123.080442	0.01
	C_9H_{15}	16803584	1010279	123.116827	123.116824	-0.02
B	$C_7H_{10}O_2$	3403886	1039053	127.075356	127.075349	-0.06
	C_9H_9O	11441856	950805	133.064791	133.064783	-0.06
AD	$C_9H_{11}O$	80283672	914365	135.080441	135.080437	-0.03
	$C_9H_{13}O$	29013398	905520	137.096091	137.096088	-0.02
C	$C_8H_{11}O_2$	4315962	968182	139.075356	139.075354	-0.01
	$C_{11}H_{13}$	40700184	865289	145.101177	145.101172	-0.03
AE	$C_{10}H_{11}O$	5923146	871874	147.080441	147.080431	-0.07
	$C_{10}H_{13}O$	8735650	848776	149.096091	149.096091	0.00
	$C_9H_{11}O_2$	23089162	830598	151.075356	151.075354	-0.01
D	$C_9H_{12}O_2$	5398861	820900	152.083181	152.08318	-0.01
	$C_9H_{13}O_2$	14555797	817125	153.091006	153.091001	-0.03
	$C_{12}H_{11}$	13179184	821729	155.085527	155.085523	-0.03
	$C_{12}H_{13}$	28868582	801816	157.101177	157.101173	-0.03
	$C_{12}H_{15}$	40034988	791246	159.116827	159.116825	-0.01
AF	$C_{11}H_{13}O$	12022132	789447	161.096091	161.09609	-0.01
	$C_{12}H_{17}$	37172608	781012	161.132477	161.132475	-0.01
	$C_{11}H_{15}O$	5187670	791056	163.111742	163.111744	0.01
	$C_{12}H_{19}$	15282275	775113	163.148127	163.148126	-0.01
E	$C_{10}H_{12}O_2$	2351923	930356	165.091006	165.090986	-0.12
	$C_{13}H_{15}$	31032332	734899	171.116827	171.116827	0.00
	$C_{13}H_{17}$	24809220	731732	173.132477	173.132476	-0.01
AG	$C_{12}H_{15}O$	11332077	732250	175.111742	175.11174	-0.01
	$C_{14}H_{11}$	2253488	766817	179.085527	179.085508	-0.11
	$C_{14}H_{13}$	10928782	702038	181.101177	181.101176	-0.01
	$C_{14}H_{15}$	22586718	693794	183.116827	183.116827	0.00
	$C_{14}H_{17}$	26925088	684101	185.132477	185.132478	0.01
	$C_{14}H_{19}$	16456400	672931	187.148127	187.148128	0.01
AH	$C_{13}H_{17}O$	5110624	683570	189.127392	189.127396	0.02
	$C_{15}H_{13}$	1843524	756056	193.101177	193.101171	-0.03

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{15}H_{15}$	16852128	649942	195.116827	195.116829	0.01
	$C_{15}H_{17}$	21890790	631269	197.132477	197.132479	0.01
	$C_{15}H_{19}$	22853914	633091	199.148127	199.148128	0.01
	$C_{15}H_{21}$	13421368	635914	201.163777	201.163781	0.02
	$C_{15}H_{23}$	6365509	635533	203.179427	203.179428	0.00
CO	$C_{15}H_{25}$	1789758	613906	205.195077	205.195107	0.15
	$C_{16}H_{17}$	15434469	596630	209.132477	209.132476	0.00
	$C_{16}H_{19}$	14826666	608144	211.148127	211.148133	0.03
	$C_{16}H_{21}$	12764766	596733	213.163777	213.163782	0.02
	$C_{16}H_{23}$	7185410	600084	215.179427	215.179431	0.02
	$C_{16}H_{25}$	7505302	585992	217.195077	217.195082	0.02
CP	$C_{16}H_{27}$	4112154	605066	219.210727	219.210723	-0.02
	$C_{16}H_{19}O$	4104325	586179	227.143042	227.143041	0.00
AI	$C_{16}H_{21}O$	3130308	574432	229.158692	229.158664	-0.12
	$C_{17}H_{25}$	5494218	544845	229.195077	229.195076	0.00
	$C_{16}H_{23}O$	5563127	530303	231.174342	231.174343	0.00
	$C_{17}H_{27}$	15667964	539172	231.210727	231.210733	0.03
	$C_{18}H_{23}$	9920324	529900	239.179427	239.179439	0.05
	$C_{18}H_{25}$	7086143	518414	241.195077	241.195084	0.03
AJ	$C_{17}H_{23}O$	6746412	520511	243.174342	243.174346	0.02
	$C_{18}H_{27}$	19552560	510063	243.210727	243.210741	0.06
	$C_{18}H_{29}$	66312728	509273	245.226377	245.226386	0.04
I	$C_{16}H_{22}O_2$	10736366	507724	247.169256	247.169264	0.03
D'Q	$C_{18}H_{31}$	68590328	502623	247.242027	247.242037	0.04
	$C_{19}H_{25}$	15773016	497725	253.195077	253.195084	0.03
	$C_{19}H_{29}$	11185875	491779	257.226377	257.226392	0.06
	$C_{19}H_{31}$	12033928	480198	259.242027	259.242035	0.03
J	$C_{17}H_{24}O_2$	7000114	493796	261.184906	261.184916	0.04
	$C_{18}H_{29}O$	8517685	485959	261.221292	261.221297	0.02
	$C_{18}H_{31}O$	52945120	472063	263.236942	263.236954	0.05
AL	$C_{19}H_{27}O$	32801638	456819	271.205642	271.205653	0.04
B'Q	$C_{20}H_{33}$	31007750	453624	273.257677	273.2577	0.08
	$C_{21}H_{29}$	9762418	432742	281.226377	281.226396	0.07
	$C_{21}H_{31}$	4469521	429681	283.242027	283.24204	0.05
L	$C_{19}H_{27}O_2$	21755982	428818	287.200557	287.200571	0.05
	$C_{19}H_{29}O_2$	41707764	425396	289.216207	289.216223	0.06
	$C_{22}H_{31}$	1740028	432588	295.242027	295.24201	-0.06
AM	$C_{21}H_{31}O$	13532269	413723	299.236942	299.236951	0.03
	$C_{23}H_{31}$	16314276	406376	307.242027	307.242042	0.05
	$C_{23}H_{33}$	4263047	421120	309.257677	309.257707	0.10
M	$C_{21}H_{31}O_2$	6241120	391991	315.231857	315.23185	-0.02

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{21}H_{33}O_2$	9477224	388439	317.247507	317.247535	0.09
	$C_{24}H_{35}$	10581947	383473	323.273328	323.273353	0.08
	$C_{24}H_{37}$	5118188	391997	325.288978	325.288994	0.05
AO	$C_{23}H_{35}O$	7850529	388796	327.268242	327.268244	0.01
N	$C_{22}H_{33}O_2$	2713440	368619	329.247507	329.247471	-0.11
	$C_{24}H_{35}O$	2443857	387737	339.268242	339.268281	0.11
AP	$C_{24}H_{37}O$	8188888	347207	341.283892	341.283882	-0.03
O	$C_{23}H_{35}O_2$	6172003	371085	343.263157	343.263171	0.04
$[M+H]^+$ - $3H_2O$	$C_{27}H_{39}$	453302144	333248	363.304628	363.30462	-0.02
$[M+H]^+$ - $2H_2O$	$C_{27}H_{41}O$	2120543232	318552	381.315192	381.315147	-0.12
$[M+H]^+$ - H_2O	$C_{27}H_{43}O_2$	1.1182E+10	307498	399.325757	399.325664	-0.23
$[M+H]^+$	$C_{27}H_{44}O_3$	2048201216	288484	417.336322	417.336247	-0.18
Average error						0.00
Absolute average error						0.04
Standard deviation						0.04

Table S4. 8 Peak assignment table for the 193 nm UVPD MS/MS of the protonated 24,25- dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	17712800	1059487	105.069877	105.069944	0.64
	C_8H_{11}	23954304	963545	107.085527	107.085587	0.56
	C_8H_{13}	24980924	967380	109.101177	109.101235	0.53
	C_9H_9	5715644	893719	117.069877	117.069921	0.38
AB	C_9H_{11}	31213176	919667	119.085527	119.085572	0.38
	C_9H_{13}	68264080	925553	121.101177	121.101219	0.35
	$C_8H_{13}O$	7556916	902875	125.096091	125.096129	0.30
	$C_8H_{15}O$	19691976	871688	127.111742	127.111777	0.28
AC	$C_{10}H_{11}$	26288002	864176	131.085527	131.085561	0.26
	$C_{10}H_{13}$	38538416	866276	133.101177	133.101208	0.23
	$C_{10}H_{15}$	35796492	856483	135.116827	135.116855	0.21
	$C_{11}H_{11}$	14143775	829719	143.085527	143.08555	0.16
AD	$C_{11}H_{13}$	49407816	819100	145.101177	145.101197	0.14
	$C_{11}H_{15}$	45961620	803419	147.116827	147.116846	0.13
	$C_{11}H_{17}$	18154492	793348	149.132477	149.132496	0.13
	$C_{12}H_{11}$	5500855	821106	155.085527	155.085539	0.08
	$C_{12}H_{13}$	24680066	759477	157.101177	157.10119	0.08

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
AE	$C_{12}H_{15}$	63619420	750320	159.116827	159.116839	0.08
	$C_{12}H_{17}$	41079884	744070	161.132477	161.132488	0.07
	$C_{12}H_{19}$	9879364	731004	163.148127	163.14814	0.08
	$C_{13}H_{13}$	8218688	710857	169.101177	169.101184	0.04
	$C_{13}H_{15}$	23843668	706342	171.116827	171.116832	0.03
AF	$C_{13}H_{17}$	36017768	699663	173.132477	173.132482	0.03
	$C_{13}H_{19}$	25557372	695992	175.148127	175.148131	0.02
	$C_{14}H_{17}$	27445308	652846	185.132477	185.132478	0.01
	$C_{14}H_{19}$	28957452	653374	187.148127	187.148126	-0.01
	$C_{14}H_{21}$	11576778	659773	189.163777	189.163774	-0.02
	$C_{15}H_{19}$	31545894	614969	199.148127	199.148125	-0.01
	$C_{15}H_{21}$	17646698	608502	201.163777	201.163772	-0.02
AG	$C_{16}H_{21}$	17848420	578889	213.163777	213.163767	-0.05
	$C_{16}H_{25}$	3725319	550113	217.195077	217.195077	0.00
	$C_{17}H_{21}$	8090119	546034	225.163777	225.163759	-0.08
AH	$C_{17}H_{23}$	10290032	538181	227.179427	227.179413	-0.06
	$C_{17}H_{25}$	5512396	531586	229.195077	229.195081	0.02
	$C_{16}H_{25}O$	4806484	554746	233.189992	233.18998	-0.05
	$C_{18}H_{23}$	10001087	520395	239.179427	239.179417	-0.04
	$C_{18}H_{25}$	16190413	508650	241.195077	241.195068	-0.04
	$C_{18}H_{27}$	12840144	501549	243.210727	243.21072	-0.03
	$C_{18}H_{29}$	10198472	503112	245.226377	245.226375	-0.01
	$C_{19}H_{25}$	17187124	492462	253.195077	253.195063	-0.06
	$C_{19}H_{27}$	26493434	478186	255.210727	255.210712	-0.06
	$C_{19}H_{29}$	9011897	474156	257.226377	257.226371	-0.02
	$C_{18}H_{27}O$	15508332	475697	259.205642	259.205634	-0.03
	$C_{18}H_{29}O$	22930460	456872	261.221292	261.221288	-0.02
	$C_{20}H_{27}$	9348607	470450	267.210727	267.210708	-0.07
	$C_{20}H_{29}$	24659606	458857	269.226377	269.226365	-0.04
I	$C_{19}H_{27}O$	29632296	453861	271.205642	271.20563	-0.04
	$C_{19}H_{29}O$	42338232	443223	273.221292	273.221287	-0.02
	$C_{21}H_{27}$	10482003	438130	279.210727	279.210727	0.00
	$C_{21}H_{29}$	8143322	450387	281.226377	281.226392	0.05
	$C_{20}H_{29}O$	23407330	425577	285.221292	285.22129	-0.01
	$C_{20}H_{31}O$	25021286	419411	287.236942	287.236942	0.00
	$C_{20}H_{33}O$	7351276	421495	289.252592	289.252588	-0.01
	$C_{22}H_{31}$	42511748	413578	295.242027	295.242027	0.00
J	$C_{21}H_{31}O$	14384799	401733	299.236942	299.23693	-0.04
	$C_{21}H_{33}O$	5789493	419512	301.252592	301.252609	0.06
	$C_{21}H_{35}O$	10422222	405905	303.268242	303.268239	-0.01
	$C_{23}H_{31}$	7698702	391351	307.242027	307.24204	0.04

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{23}H_{33}$	11489721	400825	309.257677	309.257698	0.07
K	$C_{22}H_{33}O$	3940607	388112	313.252592	313.252593	0.00
	$C_{24}H_{33}$	6899268	389133	321.257677	321.257699	0.07
	$C_{24}H_{35}$	34338592	372583	323.273328	323.273335	0.02
	$C_{24}H_{35}O$	7132963	366082	339.268242	339.268259	0.05
	$C_{24}H_{37}O$	32473164	350800	341.283892	341.283905	0.04
L	$C_{24}H_{37}O_2$	17230564	343238	357.278807	357.27882	0.04
$[M+H]^+ - 3H_2O$	$C_{27}H_{39}$	271241088	333614	363.304628	363.30465	0.06
$[M+H]^+ - 2H_2O$	$C_{27}H_{41}O$	1960107904	315056	381.315192	381.315208	0.04
$[M+H]^+ - H_2O$	$C_{27}H_{43}O_2$	2450845952	300639	399.325757	399.325759	0.01
$[M+H]^+$	$C_{27}H_{44}O_3$	4.3386E+10	295929	417.336322	417.336044	-0.67
Average error						0.06
Absolute average error						0.11
Standard deviation						0.15

Table S4. 9 Peak assignment table for the 213 nm UVPD MS/MS of the protonated 1,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	15672403	1151259	105.069877	105.069877	0.00
AB	C_7H_9O	5671770	1056906	109.064791	109.064791	0.00
	C_8H_{15}	6888588	1090361	111.116827	111.116824	-0.03
	C_9H_9	13430582	1011544	117.069877	117.069875	-0.02
	C_9H_{11}	49778484	990567	119.085527	119.085524	-0.03
	C_8H_9O	5599569	999581	121.064791	121.064791	0.00
	C_9H_{13}	22951762	977536	121.101177	121.101176	-0.01
	$C_8H_{11}O$	2289688	938768	123.080441	123.080442	0.01
	C_9H_{15}	13836954	966363	123.116827	123.116825	-0.02
B	$C_7H_{10}O_2$	4159928	962405	127.075356	127.075358	0.02
	C_9H_9O	9463323	894363	133.064791	133.064794	0.02
AD	$C_9H_{11}O$	48670132	889956	135.080441	135.080442	0.01
	$C_9H_{13}O$	35000188	876406	137.096091	137.096094	0.02
C	$C_8H_{11}O_2$	2791150	823341	139.075356	139.075361	0.04
	$C_{11}H_{13}$	22749264	829030	145.101177	145.101181	0.03
AE	$C_{10}H_{11}O$	4133081	884212	147.080441	147.080446	0.03
	$C_{10}H_{13}O$	5150603	857740	149.096091	149.0961	0.06
	$C_9H_{11}O_2$	11786313	811135	151.075356	151.075359	0.02

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
D	$C_9H_{12}O_2$	2610871	849275	152.083181	152.08318	-0.01
	$C_9H_{13}O_2$	18639658	788350	153.091006	153.091011	0.03
	$C_{12}H_{11}$	10544663	790102	155.085527	155.085534	0.05
	$C_{12}H_{13}$	16558361	770181	157.101177	157.101184	0.04
	$C_{12}H_{15}$	22596648	765381	159.116827	159.116834	0.04
AF	$C_{11}H_{13}O$	7331118	744543	161.096091	161.096102	0.07
	$C_{12}H_{17}$	25896764	752798	161.132477	161.132483	0.04
	$C_{11}H_{15}O$	2998087	744423	163.111742	163.111737	-0.03
	$C_{12}H_{19}$	13368918	745713	163.148127	163.148134	0.04
E	$C_{10}H_{12}O_2$	1414098	761802	165.091006	165.091028	0.13
	$C_{13}H_{15}$	16768625	714157	171.116827	171.116835	0.05
	$C_{13}H_{17}$	16210278	707726	173.132477	173.132486	0.05
AG	$C_{12}H_{15}O$	6639675	729354	175.111742	175.111749	0.04
	$C_{14}H_{11}$	3494311	693657	179.085527	179.085535	0.04
	$C_{14}H_{13}$	7539774	672597	181.101177	181.101187	0.06
	$C_{14}H_{15}$	12363443	672572	183.116827	183.116836	0.05
	$C_{14}H_{17}$	14874883	658403	185.132477	185.132482	0.03
	$C_{14}H_{19}$	12227371	654713	187.148127	187.148139	0.06
AH	$C_{13}H_{17}O$	3160862	625464	189.127392	189.12741	0.10
	$C_{15}H_{13}$	1939845	666314	193.101177	193.101185	0.04
	$C_{15}H_{15}$	9479680	636826	195.116827	195.116836	0.05
	$C_{15}H_{17}$	11967316	623179	197.132477	197.132486	0.05
	$C_{15}H_{19}$	12997762	613691	199.148127	199.148138	0.06
	$C_{15}H_{21}$	7969165	592310	201.163777	201.163788	0.05
	$C_{15}H_{23}$	4253304	602797	203.179427	203.179441	0.07
	$C_{16}H_{17}$	8211585	586414	209.132477	209.132492	0.07
	$C_{16}H_{19}$	9263358	574695	211.148127	211.14814	0.06
	$C_{16}H_{21}$	8128870	584929	213.163777	213.163793	0.08
	$C_{16}H_{23}$	3685306	617326	215.179427	215.179433	0.03
	$C_{16}H_{25}$	2767356	557563	217.195077	217.195104	0.12
CP	$C_{16}H_{27}$	1415982	569506	219.210727	219.210722	-0.02
	$C_{16}H_{19}O$	4412040	546302	227.143042	227.143056	0.06
AI	$C_{16}H_{21}O$	2011789	593723	229.158692	229.158665	-0.12
	$C_{17}H_{25}$	2292877	568617	229.195077	229.195095	0.08
	$C_{16}H_{23}O$	1421712	564422	231.174342	231.174373	0.13
	$C_{17}H_{27}$	2171024	583507	231.210727	231.210728	0.00
	$C_{18}H_{23}$	4945581	518261	239.179427	239.179442	0.06
	$C_{18}H_{25}$	4529859	485225	241.195077	241.195099	0.09
AJ	$C_{17}H_{23}O$	1498339	484873	243.174342	243.174346	0.02
	$C_{18}H_{27}$	5268196	490698	243.210727	243.210743	0.07
	$C_{18}H_{29}$	12161809	494207	245.226377	245.226387	0.04

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
I	$C_{16}H_{22}O_2$	2013515	507795	247.169256	247.16925	-0.02
D'Q	$C_{18}H_{31}$	6907725	488724	247.242027	247.242034	0.03
	$C_{19}H_{25}$	4780132	488605	253.195077	253.195092	0.06
	$C_{19}H_{29}$	2540403	442551	257.226377	257.226368	-0.03
	$C_{19}H_{31}$	3244561	495967	259.242027	259.242022	-0.02
J	$C_{17}H_{24}O_2$	1328823	483799	261.184906	261.1849	-0.02
	$C_{18}H_{31}O$	5919603	446807	263.236942	263.236967	0.09
AL	$C_{19}H_{27}O$	4820673	425219	271.205642	271.205643	0.00
B'Q	$C_{20}H_{33}$	4935604	457158	273.257677	273.257695	0.07
K	$C_{18}H_{26}O_2$	1314466	449077	275.200557	275.20055	-0.03
	$C_{21}H_{29}$	2633134	422861	281.226377	281.226386	0.03
L	$C_{19}H_{27}O_2$	2932959	405613	287.200557	287.200568	0.04
	$C_{19}H_{29}O_2$	6208506	414032	289.216207	289.216208	0.00
AM	$C_{21}H_{31}O$	1698177	477998	299.236942	299.236952	0.03
	$C_{23}H_{31}$	2934216	414648	307.242027	307.242035	0.03
M	$C_{21}H_{31}O_2$	1232320	422446	315.231857	315.23174	-0.37
	$C_{21}H_{33}O_2$	1666228	454891	317.247507	317.247464	-0.14
	$C_{24}H_{35}$	1740154	392020	323.273328	323.273315	-0.04
AP	$C_{24}H_{37}O$	1661324	401187	341.283892	341.2839	0.02
$[M+H]^+ - 3H_2O$	$C_{27}H_{39}$	57241780	310990	363.304628	363.304615	-0.04
$[M+H]^+ - 2H_2O$	$C_{27}H_{41}O$	2.58E+08	290981	381.315192	381.315144	-0.13
$[M+H]^+ - H_2O$	$C_{27}H_{43}O_2$	1.82E+09	274191	399.325757	399.325579	-0.45
$[M+H]^+$	$C_{27}H_{44}O_3$	3.58E+08	257406	417.336322	417.336218	-0.25
Average error						0.01
Absolute average error						0.06
Standard deviation						0.07

Table S4. 10 Peak assignment table for the 213 nm UVPD MS/MS of the protonated 24,25-dihydroxylated vitamin D_3 isomer.

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	C_8H_9	32881288	993093	105.069877	105.069924	0.45
	C_8H_{11}	33115198	956632	107.085527	107.085572	0.42
	C_8H_{13}	43788144	942111	109.101177	109.101221	0.40
	C_9H_9	9479829	797628	117.069877	117.069922	0.38
AD	C_9H_{11}	36245648	851076	119.085527	119.085567	0.34
	C_9H_{13}	1.39E+08	886775	121.101177	121.101203	0.21
	$C_8H_{11}O$	1322552	998741	123.080441	123.080476	0.28
	$C_8H_{13}O$	13515112	780069	125.096091	125.096127	0.29
	$C_8H_{15}O$	38782480	794507	127.111742	127.111779	0.29
AE	$C_{10}H_{11}$	30194694	745352	131.085527	131.085563	0.27
	$C_{10}H_{13}$	43497820	678928	133.101177	133.101206	0.22
	$C_{10}H_{15}$	48274152	675628	135.116827	135.11686	0.24
	$C_{11}H_{11}$	14663380	678338	143.085527	143.085561	0.24
AF	$C_{11}H_{13}$	50556216	644653	145.101177	145.101206	0.20
	$C_{11}H_{15}$	54138816	661344	147.116827	147.116855	0.19
	$C_{11}H_{17}$	27239012	630114	149.132477	149.132508	0.21
	$C_{12}H_{11}$	7587517	625584	155.085527	155.085556	0.19
	$C_{12}H_{13}$	22403510	586650	157.101177	157.101202	0.16
AF	$C_{12}H_{15}$	74735800	571460	159.116827	159.116844	0.11
	$C_{12}H_{17}$	59220936	595874	161.132477	161.132497	0.12
	$C_{12}H_{19}$	13951708	581309	163.148127	163.148153	0.16
	$C_{13}H_{13}$	8712179	465072	169.101177	169.101198	0.12
	$C_{13}H_{15}$	21958396	443574	171.116827	171.116844	0.10
AH	$C_{13}H_{17}$	47961592	480775	173.132477	173.132481	0.02
	$C_{13}H_{19}$	35442920	455054	175.148127	175.148136	0.05
	$C_{14}H_{17}$	31116826	453005	185.132477	185.132478	0.01
	$C_{14}H_{19}$	41693288	440129	187.148127	187.148127	0.00
	$C_{14}H_{21}$	19845776	442528	189.163777	189.163771	-0.03
	$C_{15}H_{15}$	4023556	471765	195.116827	195.116818	-0.05
	$C_{15}H_{17}$	14429588	418279	197.132477	197.132469	-0.04
	$C_{15}H_{19}$	37749508	427533	199.148127	199.148124	-0.02
	$C_{15}H_{21}$	27561042	436697	201.163777	201.163777	0.00
	$C_{15}H_{23}$	6179202	402756	203.179427	203.179439	0.06
AI	$C_{16}H_{21}$	25207266	413414	213.163777	213.163764	-0.06
	$C_{17}H_{25}$	4449089	412414	229.195077	229.195061	-0.07
	$C_{18}H_{23}$	9479774	376462	239.179427	239.179408	-0.08
	$C_{19}H_{27}$	11292916	372238	255.210727	255.210742	0.06
	$C_{18}H_{27}O$	2620914	389771	259.205642	259.205649	0.03

Assignment	Elemental composition	Intensity	Resolution	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
	$C_{18}H_{29}O$	3801215	335258	261.221292	261.221264	-0.11
L	$C_{19}H_{27}O$	4223906	353120	271.205642	271.205602	-0.15
	$C_{19}H_{29}O$	7267927	370413	273.221292	273.221293	0.00
	$C_{21}H_{29}$	3108691	332520	281.226377	281.226385	0.03
	$C_{20}H_{29}O$	3475175	334525	285.221292	285.221285	-0.02
	$C_{20}H_{33}O$	1850505	424272	289.252592	289.252642	0.17
	$C_{22}H_{31}$	12050177	331542	295.242027	295.242051	0.08
M	$C_{21}H_{31}O$	3499177	403852	299.236942	299.236933	-0.03
	$C_{21}H_{33}O$	1760126	423229	301.252592	301.252547	-0.15
	$C_{23}H_{33}$	2759874	357867	309.257677	309.257675	-0.01
	$C_{24}H_{35}$	11633214	314116	323.273328	323.273397	0.21
$[M+H]^+ - 3H_2O$	$C_{27}H_{39}$	1.14E+08	289404	363.304628	363.304735	0.29
$[M+H]^+ - 2H_2O$	$C_{27}H_{41}O$	9.16E+08	277034	381.315192	381.315195	0.01
$[M+H]^+ - H_2O$	$C_{27}H_{43}O_2$	1.3E+09	264409	399.325757	399.325758	0.00
$[M+H]^+$	$C_{27}H_{44}O_3$	1.43E+10	272661	417.336322	417.335761	-1.34
Average error						0.08
Absolute average error						0.16
Standard deviation						0.20

Table S4. 11 Fragmentation table for characteristic fragments, where one or both OH groups retained on the ring for $1,25(OH)_2D_3$ which are absent in the $24,25(OH)_2D_3$ MS/MS spectra.

$1,25(OH)_2D_3$ characteristic theoretical fragment (m/z)	Assignment	Fragmentation Method															
		CAD			IRMPD			EID			193 nm UVPD			213 nm UVPD			
		Intensity	S/N	Fragment to precursor intensity ratio (%)	Intensity	S/N	Fragment to precursor intensity ratio (%)	Intensity	S/N	Fragment to precursor intensity ratio (%)	Intensity	S/N	Fragment to precursor intensity ratio (%)	Intensity	S/N	Fragment to precursor intensity ratio (%)
109.064791	AB	X	X	X	X	X	X	medium	18.1	1.7	medium	53.6	0.37	medium	58	1.59
127.075356	B	X	X	X	high	232	8.51	medium	18.9	1.4	low	23.1	0.17	low	42	1.16
135.080441	AD	high	346.4	35.95	high	639	23.57	high	106.7	6.79	high	585.7	3.92	high	507	13.61
139.075356	C	low	26.8	2.98	high	224	8.35	medium	12.6	1.74	low	29.5	0.21	low	27.1	0.78
147.080441	AE	low	20	2.29	low	27.7	1.1	low	21.2	0.67	medium	40.9	0.29	low	40.6	1.16
152.083181	D	low	30.8	3.44	medium	63.6	2.45	high	42.2	29.32	medium	36.9	0.26	low	24.7	0.73
165.091006	E	low	44.2	4.93	medium	70.2	2.73	low	12.8	0.73	low	14.7	0.11	low	12.2	0.4
287.200557	L	high	390.1	499.33	high	423.9	16.07	medium	198.7	1.95	high	141.2	1.06	low	25.3	0.82
315.231857	M	high	27.2	40.38	medium	77.6	3.33	low	39.6	0.59	medium	38.3	0.3	low	9.2	0.34
329.247507	N	low	128.1	16	low	26.5	1.2	low	12.1	0.64	low	15.8	0.13	X	X	X
343.263157	O	high	493.1	27.27	medium	67.6	2.99	low	53.3	0.36	medium	36.9	0.3	X	X	X
357.278807	P	medium	78	10.22	low	16.7	0.81	X	X	X	X	X	X	X	X	X

Table S4. 12 Table showing the relative intensity range used to designate the low-high fragment intensity levels (low-high) for the characteristic fragments of 1,25-dihydroxyvitamin D_3.

Fragment intensity level	Relative intensity range
Low	$1x10^6 – 5x10^6$
Medium	$5x10^6 – 1x10^7$
High	$>1x10^7$

Equation S4. 1 Equation to calculate the percentage fragmentation intensity to precursor intensity ratio for the characteristic fragments for 1,25-dihydroxyvitamin D_3.

$$\text{Fragment to precursor intensity ratio (\%)} = \frac{\text{Fragment intensity}}{\text{Precursor intensity}} \times 100$$

5. Development of matrix-assisted laser desorption ionisation-time of flight mass spectrometry for the detection of SARS-COV-2 proteins

In this chapter, the application of matrix-assisted laser desorption ionisation-time of flight mass spectrometry (MALDI-TOF MS) to the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins is discussed with primary focus placed on sample preparation techniques for the enrichment and extraction of SARS-CoV-2 proteins.

This was a collaborative project, where sample preparation, data acquisition and analysis results presented in this chapter were carried out by the thesis author and Dr. Yuko P. Y. Lam. In-house software for batch processing of the MALDI-TOF data was provided by Bryan P. Marzullo and Hugh Jones. Samples, including standard viral proteins and patient samples were provided by the Arden Tissue Bank from the University Hospital Coventry & Warwickshire (UHCW) NHS Trust as well as the Facilitating Accelerated Clinical Validation of Novel Diagnostics for COVID-19 (FALCON) repository.

MALDI-TOF MS instrumentation and consumables, which were utilised for this project was provided by Bruker Daltonics.

5.1. Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus responsible for causing the coronavirus disease 2019 (COVID-19) pandemic, which has resulted in over 150,000 deaths in the UK alone. During the initial COVID-19 outbreak, the COVID-19 testing rate was low due to the limitations of human resources and the necessary equipment required for diagnostic tests. The aim of this project was to optimise the methodology for the screening of biomarker proteins in COVID-19 patient samples using mass spectrometry within 4 months to cope with the high demand of testing in the early stages of the COVID-19 outbreak. With the limited time, the focus of the project was placed on optimising methods for the benchtop MALDI-TOF MS instrument as it is a common MS instrument found in clinical test centres and hospital microbiology laboratories in the UK. Using the standard SARS-CoV-2 proteins, the most suitable protein enrichment and detection methods for COVID-19 samples were determined via testing of 73 variables from over 20 different optimisation experiments. A further 34 variables were tested and applied for the detection and enrichment of the COVID-19 proteins in the patient swab samples. A tryptic digest of the biomarker N-protein was also analysed by FT-ICR MS, which was compared to the results obtained from the MALDI-TOF instrument. The optimised method was further applied to the patient samples; the result, however, was not conclusive due to the limited replicates provided from the source of the patient samples as well as inconsistent sample preparation methods between various testing centres.

5.2. Introduction

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a large, enveloped ribonucleic acid (RNA) virus responsible for causing the global outbreak of the coronavirus disease 2019 (COVID-19).[1] SARS-CoV-2 belongs to the Coronaviridae family and is categorised into the genus known as β coronavirus, showing similarities to known coronaviruses in the same category such as the severe acute respiratory syndrome coronavirus (SARS-CoV)[2] and the Middle East respiratory syndrome coronavirus (MERS-CoV).[3] SARS-CoV-2 has four main structural proteins, which are the spike (S) glycoprotein, envelope (E) glycoprotein, membrane (M) glycoprotein, and the nucleocapsid (N) protein as well as sixteen non-structural proteins (nsp1-16). The main structural proteins play a crucial role in the infection of host cells,[4] fusion between viral and host cell membranes,[5] assembly of the virus[6] and release of the viral particles[7] whereas the non-structural proteins are responsible for viral replication and transcription (also known as messenger RNA (mRNA) production.[8,9]

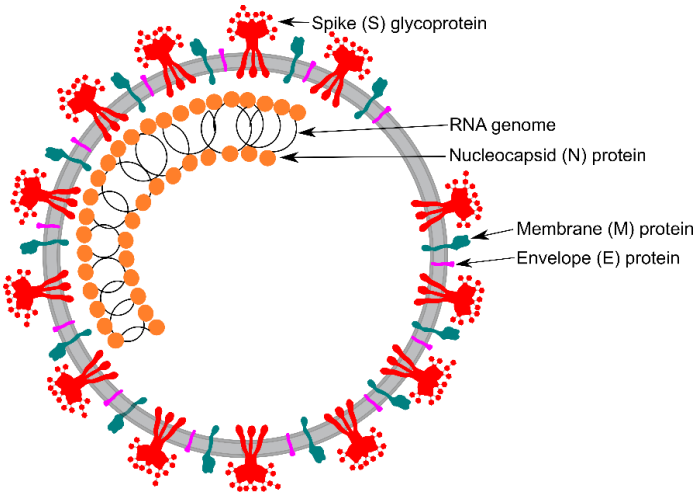


Figure 5. 1 Schematic diagram of the SARS-CoV-2 structure.

The protruding projections on the surface of SARS-CoV-2, known as the S glycoprotein, is of special interest. It is responsible for directly recognising the host cell surface receptors, as well as mediating attachment and cellular entry of the virus by assisting in the fusion of the viral and host cell membranes.[5,10,11]

The mutations in the SARS-CoV-2 genome and more specifically in the S protein sequence has been linked to changes in the properties of the virus. For example, one of the major S protein variants of interest is D614G. This is a single point amino acid mutation in the S protein, where the aspartic acid (D) at position 614 in the amino acid

sequence is replaced by glycine (G). Studies have shown that the variant D614G may be associated with a higher viral load in patients with COVID-19 as well as enhanced transmission of the virus between people.[12-14]

The S protein is also a critical target for drug development and vaccine studies as it ultimately enables viral transmission into the host cells. This takes place via binding of the receptor-binding domain (RBD) in the S protein to angiotensin-converting enzyme 2 (ACE2), which has been identified as the cellular receptor for SARS-CoV-2.[5,11] Targeting this domain of the S protein of SARS-CoV-2 with a vaccine could be effective in blocking the virus from entering the host cells and prevent the viral replication process.[15-18] Therefore, it is also important to assess the many different mutations of the spike protein to understand the impact these changes may cause in the functional properties of the virus, which may also impact vaccine efficacy. This is crucial in the development of effective vaccine therapies so that the general population can be protected against the severe consequences of infection and re-infection of SARS-CoV-2.

Methods for testing COVID-19 and clinical diagnosis

Some of the current diagnostic tests for SARS-CoV-2 infection uses nucleic acid (e.g., RNA),[19] immune-based assays (serological tests),[20] and protein-based (enzyme-linked immunosorbent assay (ELISA)[21] detection methods. However, the main recommended diagnostic test involves detection and amplification of the viral RNA using methods such as reverse transcription polymerase chain reaction (RT-PCR), which is generally carried out on symptomatic patients during the acute phase of SARS-CoV-2 infection.

A nasopharyngeal and an oropharyngeal swab is taken from the patient and tested for the presence of viral RNA using RT-PCR. As SARS-CoV-2 contains only RNA, the first step involves extraction and isolation of all the RNA from the patient's sample. The RNA in the sample is then converted to complementary-deoxyribonucleic acid (cDNA) using the enzyme reverse transcriptase as PCR uses DNA polymerase and therefore only recognises DNA. The PCR enables amplification and detection of a specific DNA molecule, which in this case is the viral cDNA. In real-time PCR, the levels of viral cDNA can be monitored via the addition of a probe during the PCR process, which gives off fluorescence when a new DNA molecule is formed, hence an increase in viral cDNA can be monitored by an increase in the fluorescent signal.

More recently, single-use antigen rapid diagnostic tests (Ag-RDT), which target specific antigens from SARS-CoV-2 have been developed. Samples are obtained from a nasopharyngeal and an oropharyngeal swab, which is mixed with a viral protein extraction buffer. This mixture is then applied to the specimen well of the test device that allows the sample to flow past the specific protein antibody via capillary action to the test line. If the viral proteins are present, the antibodies on the substrate will bind to the proteins forming an antigen-antibody complex and a colour develops to indicate the result, which can be visually read. These tests also have the benefit of speed and ease of use as they can provide a result within 30 minutes and can be carried out by people without the need for special training or expertise in a laboratory environment.

However, the rapid antigen tests lack in sensitivity compared to RT-PCR, resulting in an increased risk of false-negative results, especially when there is a low viral load. Hence the analytical specificity and sensitivity of the methods need to be considered to determine the rate of false negative and false positive results as this may have a significant impact on a positive COVID-19 diagnosis or when ruling out infection.

MS methods for SARS-CoV-2 detection

Matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) has emerged as a promising analytical tool for the rapid and sensitive detection of microorganisms. Over recent years, MALDI-TOF MS has found routine use in clinical microbiology laboratories as it is an easy to use, fast and high throughput technique. It has increasingly been used for microbial identification including detection of harmful bacteria in contaminated water and food, as well as detection of pathogens in blood and urine samples.[22]

Commercial MALDI-TOF platforms such as the MALDI Biotyper (Bruker Daltonics) and Vitek MS (BioMérieux) have successfully been implemented into routine clinical testing in hospital microbiology laboratories. Many studies have demonstrated the potential of MALDI-TOF for the sensitive detection of microbiological species. For example, in 2010, comparisons between MALDI-TOF MS with traditional biochemical tests routinely used for the identification of bacterial species were carried out and a high-confidence correct identification was found in 99.1% of cases.[23] MALDI-TOF MS can also be performed directly for identifying bacterial isolates in blood cultures. In a previous study, 584 positive blood cultures were tested and 562 were found to contain unique bacterial species.[24]

Identification of microbes by MALDI-TOF MS is generally based on matching the peptide mass fingerprint (PMF) of the unknown organism with known PMFs contained in a database.[22] Peptide mass fingerprinting involves digesting the unknown protein sample or organism with a specific enzyme e.g. trypsin to generate peptides and then the experimental peptide molecular masses are matched to theoretical peptide masses from spectral reference databases.[25] Microbes can also be identified via matching the *m/z* of biomarkers in the unknown organism to biomarkers in a proteome database.

The spectral reference databases of known organisms are continuously expanding but there are limitations to the use of databases such as the similarities between organisms and the limited number of spectra for unique species in the database can lead to difficulties in identifications, which may result in a misdiagnosis. These issues can be overcome by additional testing and tandem MS/MS. The use of MALDI-TOF MS has generally shown overall improvements in microbial disease diagnosis, diagnosis time as well as clinical care.

Although RT-PCR is the current "gold standard" method for COVID-19 testing, there is an urgent need to increasing testing capacity to meet the global demands as the different variants of the virus continue to spread across the population. An alternative yet highly sensitive, reliable, and rapid testing method is required to directly detect the virus from patient samples such as nasopharyngeal swabs, oropharyngeal swabs, swab, and gargle solutions. MS platforms in general offer high mass accuracy, sensitivity, and specificity with the ability to detect and identify viral peptides to viral proteins as well as whole viruses. MALDI-TOF MS can be an ideal platform for the analysis of the SARS-CoV-2 virus as it is a fast technique, capable of providing higher specificity and sensitivity compared to RT-PCR and lateral flow antigen tests.

Recent studies have demonstrated the potential for implementing MS and MALDI-TOF MS as rapid detection methods for SARS-CoV-2. Iles et al. used a Shimadzu MALDI-TOF 8020 on gargle solutions spiked with cultures of SARS-CoV-2 to test for the viral proteins.[26] Gargle samples were filtered and acetone precipitated for enrichment of the virus particles. The viral proteins were then extracted and solubilized using a specific extraction formulation buffer called LBSD-X. With the MALDI-TOF, for the gargle solutions, the S1 subunit of the spike protein was detected. Ihling et al. used nano-HPLC/Nano-ESI-Orbitrap-MS/MS also on acetone precipitated gargle solutions for the detection of the tryptic peptides of the N protein.[27]. Nikolaev et al. successfully developed an LC-MS/MS method utilising a nano-HPLC coupled to a tims-TOF Pro (Bruker

Daltonics) for the detection of tryptic peptides of the viral N protein from nasal epithelial swabs.[28] The use of chromatography in combination with MS also provides an extra dimension of separation of the species, particularly for the digested viral proteins.

For SARS-CoV-2, patient samples are generally collected as nasopharyngeal swabs, oropharyngeal swabs and more recently swab samples, which must be fully inactivated using chemicals, heat, or ultraviolet radiation. The extraction of the virus from the specimens are often based on traditional methods of cell lysis and protein extraction prior to MS analysis such as acetone precipitation. Therefore, the optimisation of sample preparation and the development of viral protein extraction methods are crucial for MS and in particular MALDI-TOF MS detection.

5.3. Experimental Section

Sample preparation for recombinant SARS-CoV-2 proteins

The S1 subunit of the recombinant SARS-CoV-2 spike S protein and the S2 subunit, derived from Escherichia Coli (E.Coli), were provided by RayBiotech Inc. (RayBiotech Inc., Peachtree Corners, GA, USA). The full-length SARS-COV-2 Spike trimer, derived from Hek293 cells, was obtained from Peak Proteins Ltd. (Peak Proteins Ltd., Alderley Park, Macclesfield, UK). The His-tagged SARS-COV-2 Nucleocapsid Protein (NCAP or N), produced in E.Coli, was obtained from Sheffield University. All samples were aliquoted into smaller volumes to reduce multiple freeze–thaw cycles and stored at -80 °C until thawed prior to sample preparation and analysis. Details of the various sample preparation methods for the recombinant SARS-CoV-2 proteins are provided in Supplementary Table S5.1.

Standard N and S-protein enrichment by SpeedVac dry down

The standard N-protein and S-protein were mixed and diluted in 3mL 80% ethanol solution to 0.005 μM. The mixture sample was dried in a Savant SPD121P SpeedVac concentrator (Thermo Scientific, Waltham, MA, USA)with heat at 45 °C for 6 hours. After the drying down, the sample was reconstituted in 50:50 H_2O:ACN with 0.1%TFA and sonicated for 30 minutes. Prior to MALDI-TOF MS analysis, the sample was mixed with 40 mg/mL Sinapinic Acid (SA) in a 2:1 volume ratio and spotted on the MALDI plate).

Digestion of recombinant SARS-CoV-2 proteins

The N-protein was dissolved in 100 mM ammonium bicarbonate (ABC) solution to 1 μg/μL. Disulphide bonds were then reduced using 50 mM dithiothreitol (DTT, Sigma Aldrich Company Ltd.) for 30 minutes at 60 °C, followed by alkylation with 100 mM of iodoacetamide (IAA, Sigma Aldrich Company Ltd.) and the samples were stored in the dark at room temperature for 1 hour. The solution was then tryptic digested with 1 mg/mL trypsin (Sigma Aldrich Company Ltd.) in 100 mM ABC solution at 37 °C for 16 hours. After the tryptic digestion, samples were desalted using SOLAμ SPE C18 cartridges (ThermoFisher, Waltham, MA, USA) with the elution buffer 80 % ACN and 0.1 % formic acid. The desalted samples were further diluted with 20 % ACN and 0.1 % formic acid into final concentrations of 0.2 μg/μL for direct infusion MS and CAD MS/MS analysis.

Sample preparation for patient samples

Virus inactivation

Chemically inactivated (concentration of 80 % ethanol and 20 % water) COVID-19 positive and negative patient swab samples were provided by the Arden Tissue Bank at the University Hospital Coventry & Warwickshire (UHCW) NHS Trust (Table 5.1). The sample tubes were disinfected on arrival and all work was carried out in accordance with the guidelines within the designation of a P1 laboratory.

Table 5. 1 Table of the COVID-19 negative and positive tested patient swab samples.

Patient sample number	COVID-19 status	Additional information
2	Negative	
3	Negative	
10	Negative	
17	Positive	Symptoms shown (as stated on the sample vial)
18	Positive	Symptoms shown (as stated on the sample vial)
19	Positive	

Centrifugation method

Each of the patient swab samples were transferred to a 15 mL Falcon conical tube, which were then sonicated for 30 minutes. The solutions in the tubes were transferred to 1.5 mL Eppendorf tube for centrifugation at 14,000 rpm at room temperature for 30 minutes. The supernatant and pellet were separated, and the supernatant was transferred to a new Eppendorf tube. The pellet was then reconstituted in 50:50 H_2O:ACN with 0.1% TFA.

Protein enrichment by SpeedVac dry down

The centrifugation method was followed as mentioned above. After the pellet was reconstituted in 50:50 H_2O:ACN with 0.1% TFA, 400 µL of the pellet solution was transferred to a to 1.5 mL Eppendorf tube. The samples were then dried with a Savant SPD121P SpeedVac concentrator (Thermo Scientific, Waltham, MA, USA). After the drying down, the sample was reconstituted in 50:50 H_2O:ACN with 0.1%TFA and sonicated for 30 minutes. Prior to MALDI-TOF MS analysis, the sample was mixed with 40 mg/mL Sinapinic Acid (SA) in a 2:1 volume ratio and spotted on the MALDI plate.

MALDI-TOF MS analysis

Experiments were carried out using a Bruker Microflex LT MALDI-TOF MS (Bruker Daltonik GmbH, Bremen, Germany), equipped with an accelerating voltage 20 kV and a nitrogen laser (337 nm). The experiments were performed in positive ion, linear-ion mode. The standard settings for MALDI MS profiling of proteins using the Bruker default method for the 66 kDa bovine serum album (BSA) protein was used with changes made to the parameters mentioned in this section and in Supplementary Table S5.3. The random rastering of the sample spot on the MALDI target plate was set to automatic, with 50 laser shots fired per raster spot. For the standard protein optimisation experiments and for the patient swab samples, 2000 scans were acquired over the mass range of *m/z* 5,000-300,000. Further details of the instrument detection parameters used are listed in the Supplementary Table S5.3.

FT-ICR MS analysis

Experiments were also carried out using a 12 tesla (T) SolariX Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS; Bruker Daltonik GmbH, Bremen, Germany), equipped with a shielded superconducting magnet.

Direct infusion experiments

For the direct infusion experiments, the samples were loaded into borosilicate glass capillary tips (purchased from World Precision Instruments, Inc., Sarasota, FL, USA), which were pulled using a Sutter P-97 capillary Flaming/Brown micropipette puller instrument (Sutter instruments Co., Novato, CA, USA). The pulled tips were optimised for a low-flow nano-electrospray ionisation (nESI) experiments.

All samples were sprayed in positive ionisation mode. Mass spectra were acquired with a 4 mega-word (MW) data-points (32 bits) over a mass range of *m/z* 147 – 3,000 to produce a 1.68 s transient and ~460,000 resolving power at *m/z* 400.

Positively charged ions were transmitted through a glass capillary to a quadrupole and then externally accumulated in a hexapole collision cell for 0.5 s before transferred to an infinity cell for MS excitation and detection.

For CAD MS/MS experiments, 2+ precursor ions of the tryptic peptides of the N-protein digest were first quadrupole isolated at *m/z* 443.7 and *m/z* 563.8 with an isolation window of 5 *m/z*. The ions were then subjected to collisions with argon gas in the collision cell. The optimised collision energy (CE) of 6 V and 10 V was applied to the tryptic peptides detected and isolated at *m/z* 443.7 and *m/z* 563.8, respectively. Fragments, together with the precursor ions, were then transferred to the infinity cell for mass detection.

Data analysis

All mass spectra acquired on the Microflex MALDI-TOF were processed and analysed using the flexAnalysis software (Bruker). All FT-ICR mass spectra were analysed using DataAnalysis 4.3 (Bruker), internally calibrated and fragments were assigned manually with a mass error <1 ppm (supplementary table S5.4– S5.5). All spectra were internally calibrated with known *m/z* fragmented peaks that contain minimum threshold of S/N >3 and peaks were picked with relative intensities higher than $1x10^6$ according to the Bruker FTMS peak picking algorithm.

5.4. Results and Discussion

Optimisation of matrices for MALDI-TOF MS analysis of the S1 and S2 subunits of the spike protein

Sample-matrix conditions are known to significantly influence the detection of peptides and proteins using MALDI-TOF MS.[29] In this work, 10 different matrices were tested and the results indicate that out of all the matrices, Ferulic acid (FA) resulted in the detection of the S1 and S2 subunits of the spike protein with the highest signal-to-noise ratio (S/N) compared to other matrices. The best signal for both S1 and S2 units were achieved by mixing 30 mg/mL FA solution with 500 fmol/µL of the protein in 1:1 ratio. The mass spectra of the S1 and S2 subunits are displayed in Figure 5.2, where the S1 monomer is detected around *m/z* 80,000 and the S2 monomer and dimer were detected at *m/z* 60,000 and *m/z* 120,000, respectively.

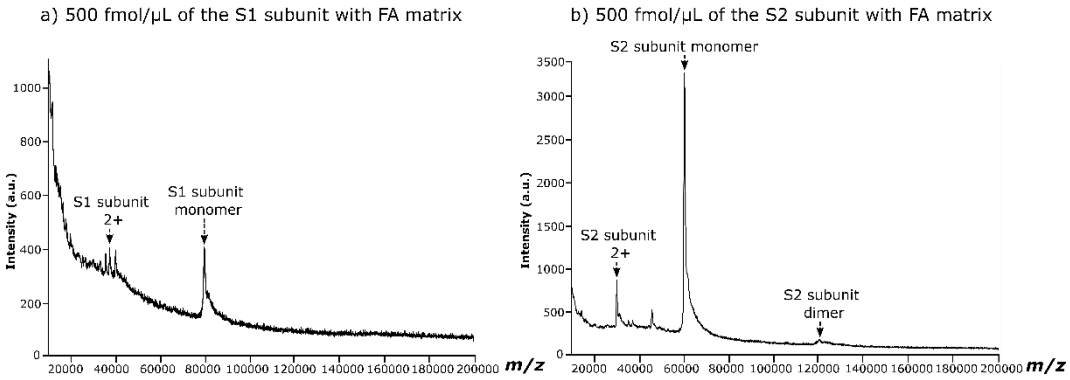


Figure 5. 2 MALDI-TOF MS detection of a) S1 subunit of the spike protein and b) S2 subunit of the spike protein using FA as a matrix.

Although, the use of FA as the matrix resulted in the best signal detection for both the S1 and S2 subunits of the spike protein, manual control to search for regions of "sweet spots" within the MALDI samples was necessary. This can be attributed to the observed uneven co-crystallisation of the protein sample and the FA matrix as depicted by the images of the sample spots on the MALDI target plate in Figure 5.4b. In contrast, when sinapinic acid (SA) was used as the matrix, the co-crystallisation of the SA matrix and the protein was homogenous as shown by Figure 5.4a, indicating SA matrix is more suitable for a robust and high throughput experiment.

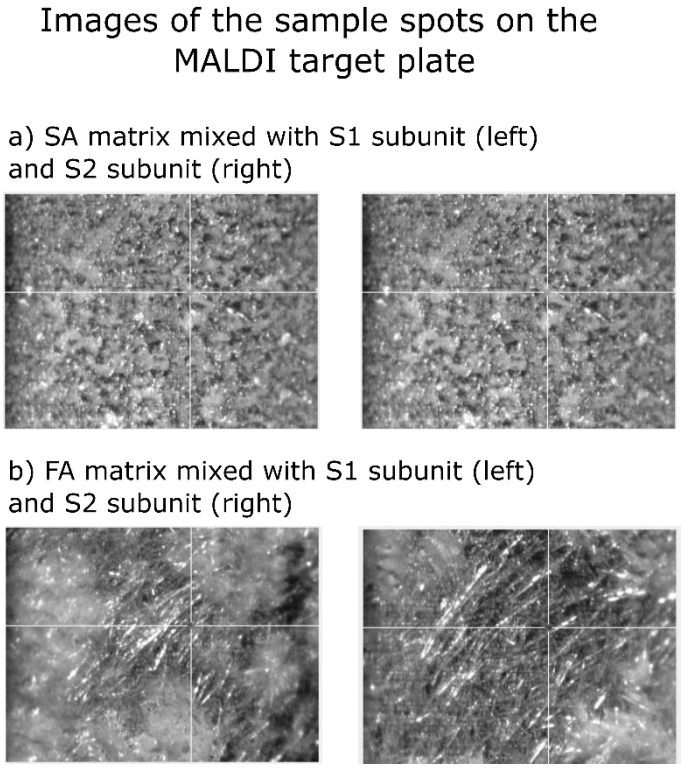


Figure 5. 3 Images of sample spot on MALDI target plate a) SA matrix mixed with S1 subunit and S2 subunit of the S-protein and b) FA matrix mixed with S1 subunit and S2 subunit of the S-protein

Previous studies have shown that SA is highly suitable matrix, known to provide consistent signal detection for large proteins and also viral glycoproteins.[30-33] Although, the S/N for the S2 monomer using FA as the matrix was approximately 3-fold higher than when SA was used, the same peaks were detected with SA and it proved to be the second best matrix out of the 10 matrices tested for the detection of the S1 and S2 subunits (Figure 5.3). The best signal obtained for both S1 and S2 was mixing 500 fmol/µL of the protein solutions with 40 mg/mL of the SA matrix in a 2:1 ratio. When matrices such as α-Cyano-4-hydroxycinnamic acid (CHCA) were used, neither the S1 nor S2 subunit was detected, as CHCA is commonly used for the detection of lower mass ions such as small molecules and peptides. The additional benefit of using SA as the matrix is that due to the homogeneity of co-crystallisation of the sample and matrix, the random walk movement on the MALDI spot can be applied, where the laser irradiation spot moves automatically, instead of manually searching for the sweet spots due to the uneven co-crystallisation when FA is used as a matrix.

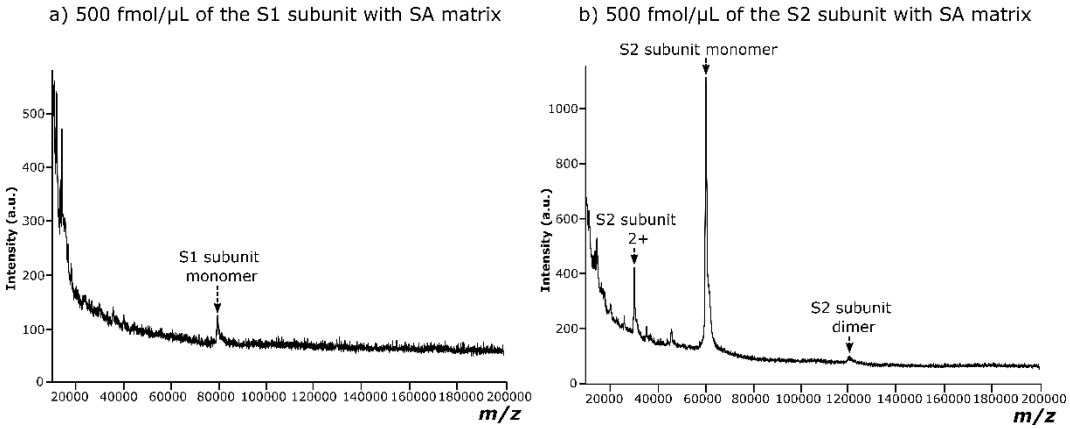


Figure 5. 4 MALDI-TOF MS detection of a) S1 subunit of the spike protein and b) S2 subunit of the spike protein using SA as a matrix.

N-protein and full-length S-protein mixture detection by MALDI-TOF MS

It is expected that the detection of the viral proteins in the patient samples will be affected by the presence of other proteins. To further develop the sample preparation and methods for optimisation, the standard N-protein and S-protein were mixed in a 1:1 concentration ratio into 0.5 µM, with SA (40 mg/mL) as the matrix (2:1 ratio of protein sample mixture to the matrix).

Figure 5.5 clearly demonstrates the presence of N-protein monomer at *m/z* 46,000, the doubly charged (2+) form and the dimer in both the N-protein MS and in the mixture of N and S-protein spectrum. On the other hand, the S-protein monomer peak and doubly charged protein peak were only detected at low intensities in the MS of the S-protein, where the S/N of the S-protein monomer was approximately 50 times lower compared to the N-protein monomer. It was easier to detect of the N-protein in the N and S-protein mixture sample as the S-protein is at a much higher *m/z* ratio (*m/z* 180,000 compared to *m/z* 46,000 for the N-protein) and there is also the ion suppression effect resulting from increased signal detection of the N-protein, which may explain the absence of the S-protein in the N and S-protein mixture spectrum.

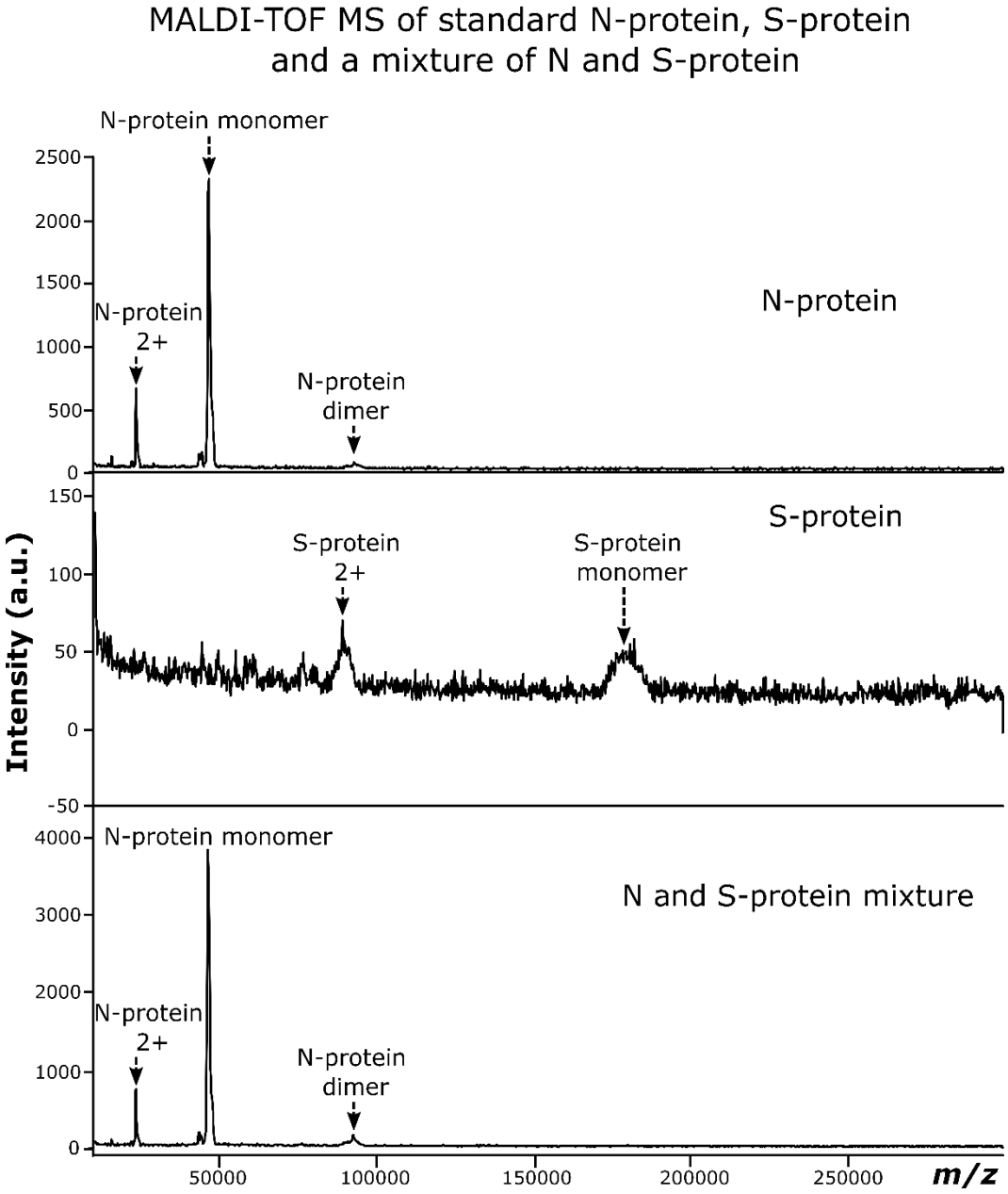


Figure 5. 5 **MALDI-TOF MS detection of N-protein (top), S-protein (middle) and a mixture of the N and S-protein (1:1 concentration ratio) using SA matrix.**

Further mixing method optimisation for the detection of the standard N and S-proteins in the sample mixture

As MALDI is strongly affected by the behaviour of the matrix and sample crystallisation, different mixing methods were investigated for the optimum detection of the N and S-proteins in the sample mixture. Both slow and fast crystallisation of the matrix and sample can result in the formation of different kinds of crystals, therefore the best method for embedding the sample molecules into the matrix crystals is required. MALDI spotting methods tested include the dried droplet technique,[34] three-layered sandwich method,[35] the two-layered method,[36] and mixing the sample and matrix in a 96-well plate prior to spotting on the MALDI target plate.

Although the S-protein was not detected in the MS of the sample mixture, the N-protein monomer and 2+ peak was detected, when FA, SA and a combination of both matrices were used (Figure 5.6). The best N-protein monomer signal detected was using the SA matrix (40 mg/mL) with mixing in 96-well plate before spotting on MALDI target plate. The lowest S/N for the N-protein in the N and S mixture sample was detected with the FA matrix, which was approximately 7 times lower compared to when the SA matrix was used. A mixture of the FA and SA matrix did show an improvement in the S/N compared to when only FA was used as a matrix. The S/N of the N-protein monomer peak increased 6-fold compared to when only the FA matrix was used. However, simply using SA as the matrix and the 96-well plate mixing method proved to be the most effective at improving the protein signal detection.

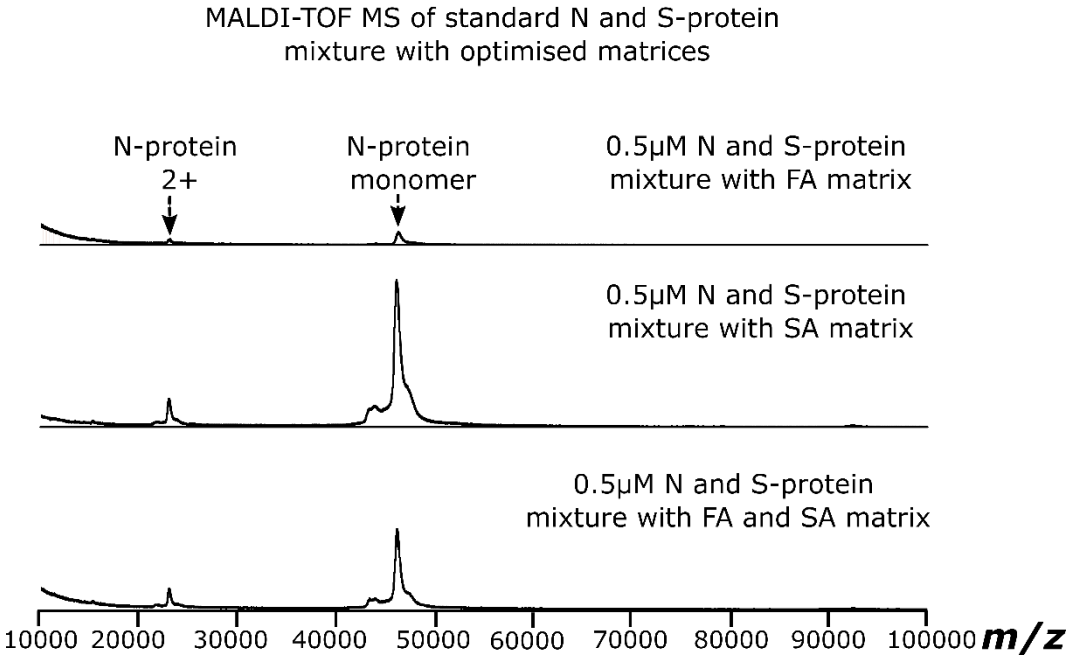


Figure 5. 6 Effect of FA matrix, SA matrix, and a combination of both matrices on the detected N-protein in the N and S-protein mixture when sample and matrix in a 96-well plate mixing method was used (y-axis for each spectrum was set to same relative intensity).

Matrix solvent variation for N and S protein mixture

The solvents used for the matrices and the samples also play an important role. Figure 5.7 depicts the matrix solvent, which resulted in the highest signal of the N-protein as well as the matrix solvent combination, which resulted in the lowest S/N of the N-protein in the N and S protein mixture sample. With acetonitrile/water (50/50) + 0.1 % TFA as the matrix solvent, the N-protein monomer S/N was approximately 14-fold higher than when hexane/ethanol (50/50) + 0.1 % TFA was used as the matrix solvent, suggesting that acetonitrile/water (50/50) + 0.1 % TFA was the optimal matrix solvent. Other matrix solvent variations with the detection percentages, relative intensities, and S/N values are provided in the Supplementary Table S5.1.

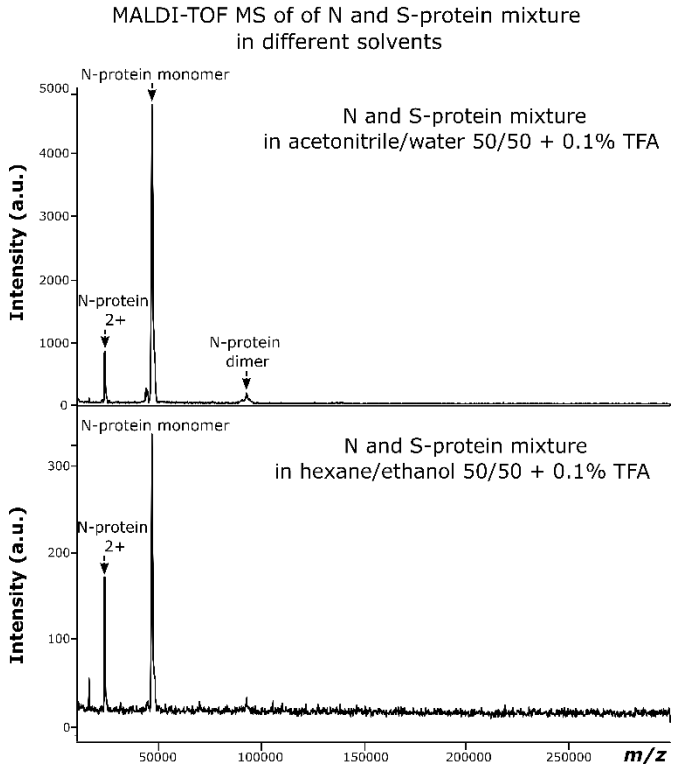


Figure 5. 7 MALDI-TOF MS of the mixture of the N and S-protein (1:1 concentration ratio) sample with different SA matrix solvents.

Enrichment of the standard N and S-proteins in the sample mixture by the SpeedVac dry down method

As the patient samples were provided in a viral deactivation buffer (high percentage concentration of ethanol), the conditions were reproduced for the standard N and S protein mixture. This is necessary to determine an effective way for the viral N and S protein enrichment when the viral biomarker proteins are diluted, which would decrease the concentration in the sample, making detection of the proteins difficult.

The mass spectrum for the diluted N and S-protein mixture sample in 80 % ethanol and the mass spectrum for the enrichment method involving drying down the sample and reconstitution in solvent is provided in Figure 5.8. The data shows, as expected, that at a low concentration (0.005µM), when diluted with ethanol, the N-protein and S-protein were not detected whereas the drying the samples down with the SpeedVac is an effective way to concentrate N-protein from ethanol as both the N-protein monomer and the N-protein 2+ peak are observed.

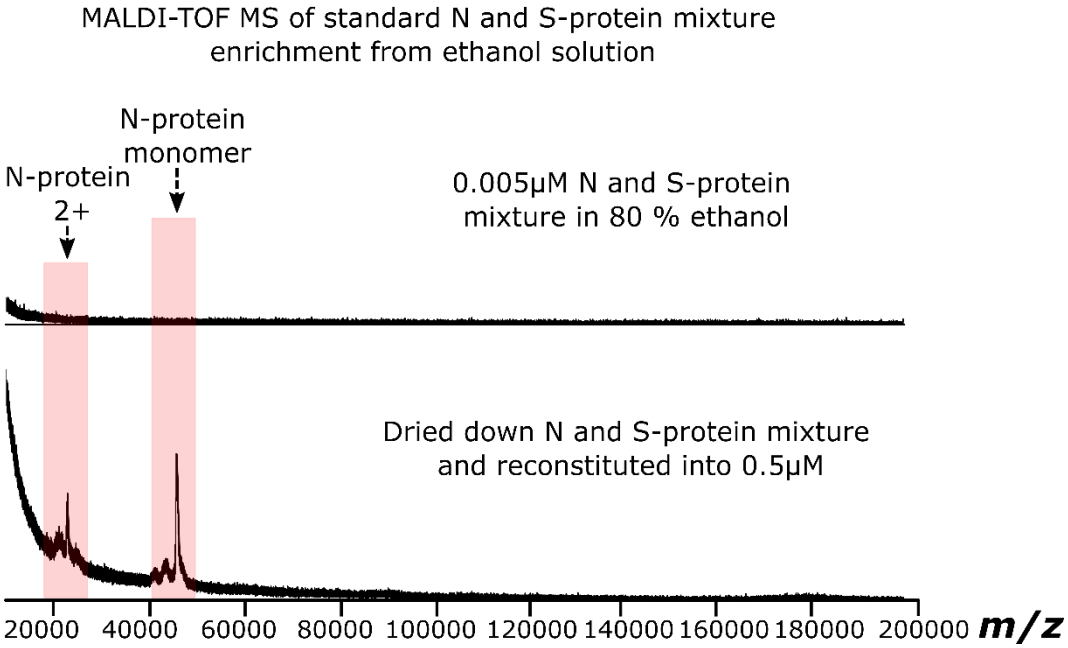


Figure 5. 8 MALDI-TOF MS of standard N and S-protein mixture diluted in ethanol (top spectrum) and dried down to concentration the N and S-protein (bottom spectrum).

Tryptic digestion of the standard N-protein and analysis via FT-ICR MS

It can be challenging to characterise the viral biomarker proteins via top-down methods due to the high mass of the species and because many of the viral proteins are post-translationally modified such as the heavily glycosylated spike protein, which can negatively impact the sensitivity of detection. Therefore, bottom-up methods were also considered. The tryptic digest results for the N-protein using MALDI-TOF MS can be difficult to interpret as there is no isotopic resolution and because of the interference of matrix-related peaks in the low mass region.

Figure 5.9 depicts the mass spectra obtained on the FT-ICR MS for the tryptic digested N-protein with no sample pre-treatment and the tryptic digested N-protein desalted with the SPE C18 cartridges. Significant differences between both spectra can be observed, as without sample pre-treatment, the signal was unstable, the peptides were highly charged, and the peptide signal intensities were approximately 15.6 times lower than the peak intensities of the peptides observed in the desalted N-protein tryptic digest mass spectrum.

Sample pre-treatment has improved the sensitivity of detection of the peptides in the N-protein tryptic digest spectrum. This was also observed for the N-protein tryptic digest using MALDI-TOF MS (Supplementary Figure S5.1). As a result, two peptides

were selected and subjected to CAD MS/MS analysis on the solariX 12 T FT-ICR mass spectrometer to confirm the peptide sequence as shown by Figure 5.10. Fragment peaks were assigned with high confidence as sub-ppm mass errors were obtained. Isotopic resolution was also easily achieved using FT-ICR MS, as for the MS results, a resolving power of approximately 420,000 was obtained at *m/z* 400 and for the CAD MS/MS of both peptides, a resolving power of approximately 460,000 was obtained at *m/z* 400.

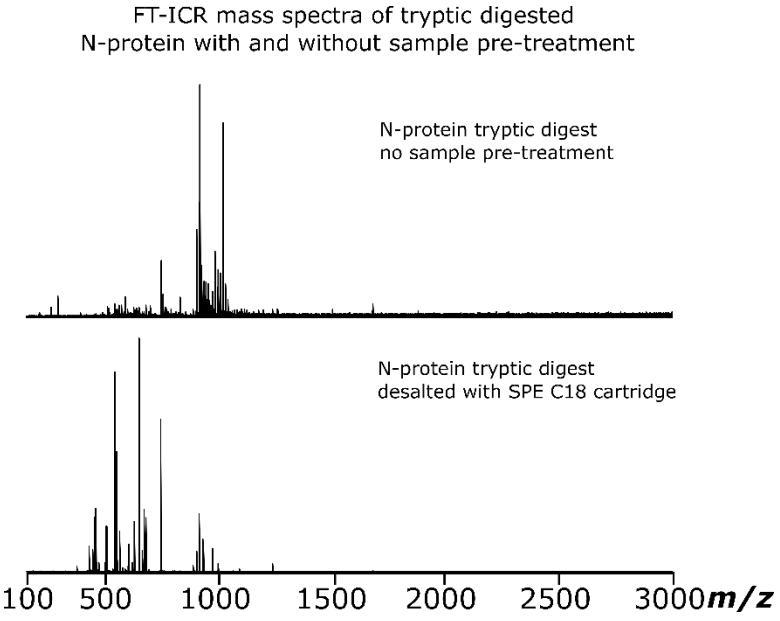


Figure 5. 9 FT-ICR MS of tryptic digested N-protein with no sample pre-treatment (top spectrum) and the tryptic digested N-protein desalted with the SPE C18 cartridges (bottom spectrum).

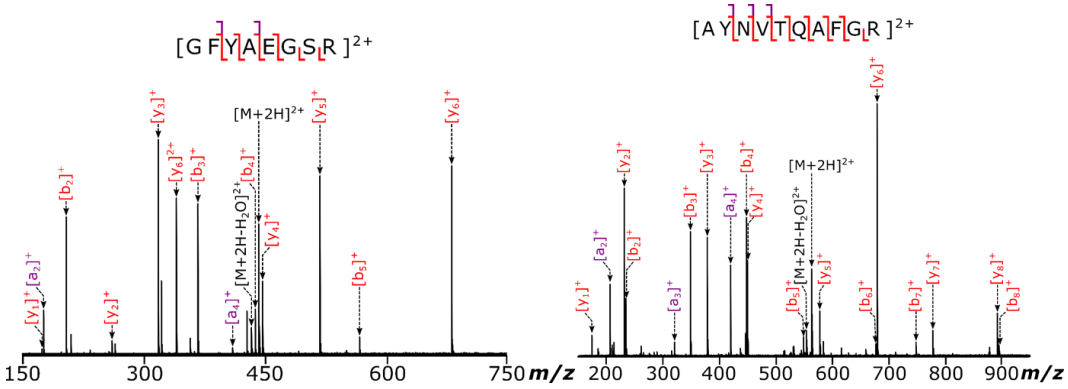


Figure 5. 10 CAD MS/MS spectra of two tryptic peptides of the N-protein. Peak assignment tables for the assigned CAD MS/MS spectra with absolute average mass errors approximately < 0.29 ± 0.27 ppm for both peptides (Table S5.4– S5.5).

MALDI-TOF MS analysis of COVID-19 positive and negative patient swab samples

The provided swabs in solution were sonicated first to aid movement of any proteins/viral particles present on the cotton bud back into the solution. Sonication uses sound energy to agitate particles, which is traditionally applied in cell lysis experiments for disruption of the cell membrane, to release the contents of the cell.[37] Previous studies have shown that sonication of the swab has provided greater recovery of cells, with more consistency, than for the standard swabs without sonication.[38-40]

After sonication of the swabs, the samples were subjected to centrifugation, which is a method to separate molecules based on their densities by spinning them in solution around an axis i.e., in a centrifuge rotor and applying centrifugal force at high speed. This method has been used to collect cells, to precipitate DNA, as well as to purify and concentrate virus particles. Recent studies for the detection of SARS-CoV-2 biomarkers have demonstrated the retention of the pellet for further experiments as the viral particles were detected in the pellet, whilst the centrifugation process resulted in the removal of lipids, particulates, and other unwanted proteins in the supernatant.[26,27,41]

Figure 5.11 depicts the mass spectra obtained using MALDI-TOF MS of the COVID-19 positive and negative patient swab samples. Peaks in the MS, common to all samples were detected. However, due to the variation of samples provided, such as patient sample number 17 and 18, were displaying symptoms, peaks unique to those samples were observed. For example, human serum albumin was detected at approximately *m/z* 66,000 in the MS of COVID-19 positive patient sample number 17.

Although sonication and centrifugation steps were taken, the N-protein, S-protein or the S protein fragments (S1 and S2 subunit) were not detected in the MALDI-TOF mass spectra.

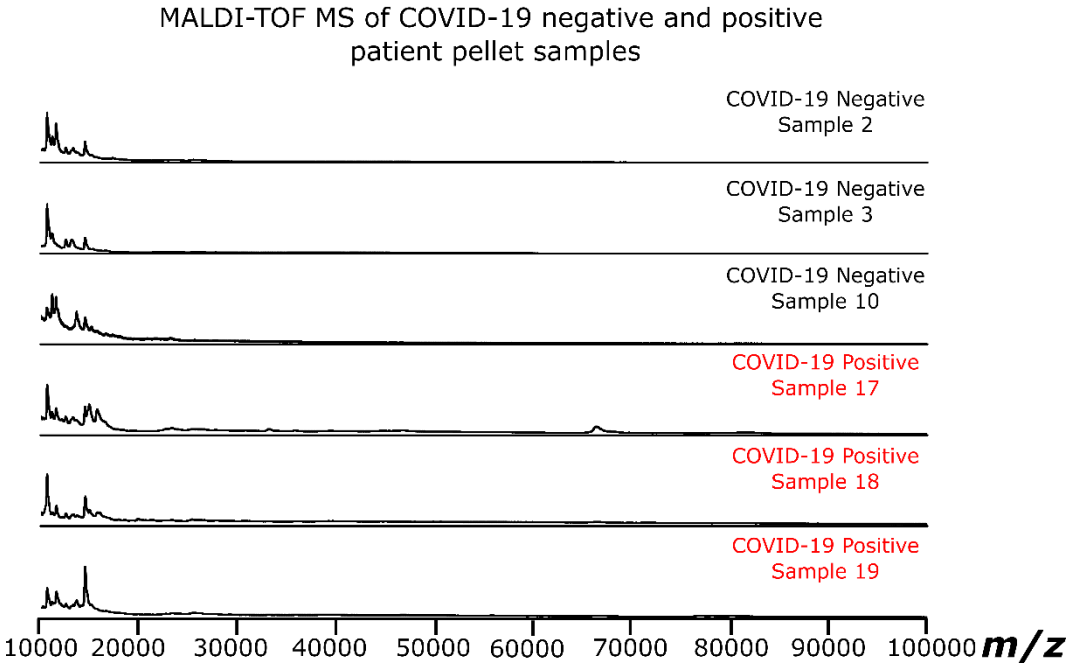


Figure 5. 11 MALDI-TOF MS of COVID-19 negative and positive patient swab pellet samples with no enrichment methods applied.

Figure 5.11 depicts the mass spectra obtained using MALDI-TOF MS of the COVID-19 positive and negative patient swab samples. Peaks in the MS, common to all samples were detected. However, due to the variation of samples provided, such as patient sample number 17 and 18, were displaying symptoms, peaks unique to those samples were observed. For example, human serum albumin was detected at approximately *m/z* 66,000 in the MS of COVID-19 positive patient sample number 17. Although sonication and centrifugation steps were taken, the N-protein, S-protein or the S protein fragments (S1 and S2 subunit) were not detected in the MALDI-TOF mass spectra.

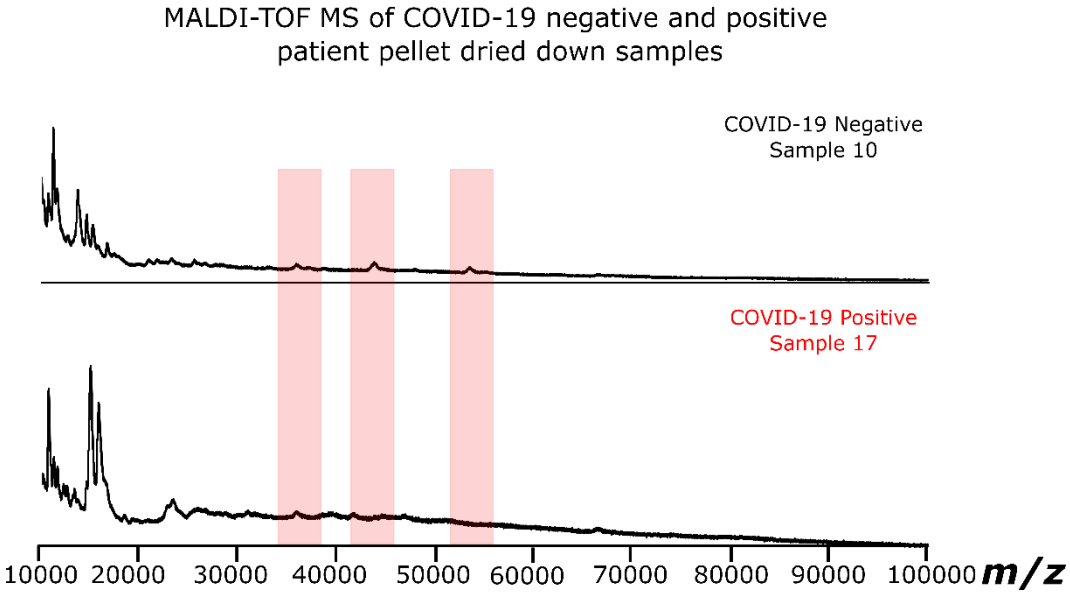


Figure 5. 12 MALDI-TOF MS of COVID-19 negative and positive patient swab pellet samples were dried down to concentrate the viral proteins in the samples.

Other sample pre-treatment methods include the use of centrifugal filters, such as molecular weight cut off (MWCO) filters, which can be used to concentrate, purify, desalt and remove detergents.[42,43] In this case, the 100k MWCO filter was used to enrich the high mass viral proteins e.g. S-protein (180 kDa). Other methods include the use of detergents such as dimethyl sulfoxide (DMSO), which is a polar organic solvent able to dissolve both polar and non-polar compounds. It's commonly used to dissolve neurotoxic agents, used in cell lysis, acts as a chemical penetration enhancer.[44-46]

Figure 5.13 shows that with the 100k MWCO filter more proteins are observed for COVID-19 positive patient number 17 compared to the fresh sample even though the peaks are low in intensity. DMSO has a negative impact on the co-crystallisation process but is a good solubilising agent. The peak at *m/z* 66,000 is suspected to be human serum albumin, which was detected and the improvement in the S/N of an unknown protein at *m/z* 150.00 was also observed. However, although the enrichment and purification steps applied herein, the SARS-CoV-2 viral biomarker N and S-proteins were not detected.

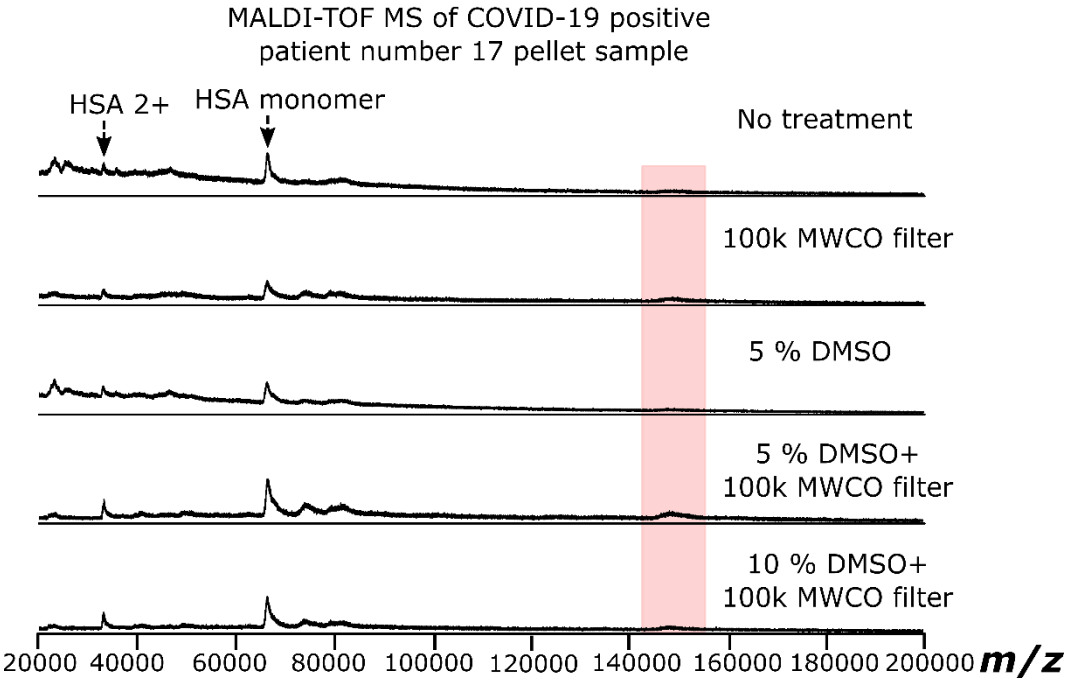


Figure 5. 13 Effect of the individual sample pre-treatment and combination of MWCO filter and detergents for the improvement of protein signal detection for COVID-19 positive patient sample number 17.

5.5. Conclusions

The experimental results demonstrate the different methods for protein enrichment and extraction were investigated, which were subsequently applied to the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) standard proteins, including the SARS-CoV-2 positive and negative swab patient samples.

The different approaches explored herein identified sample pre-treatment methods such as high molecular weight cut-off filters and the use of detergents often used in cell lysis, aided the enrichment and extraction of proteins present in the patient samples. This was characterised by improved signal intensities of proteins such as human serum albumin present in the MALDI-TOF MS of sample number 17. The various sample preparation techniques were firstly applied to the standard viral SARS-COV-2 proteins, which demonstrated enrichment of the proteins and improved S/N of the detected species. However, due to the complexity of the patient swab samples, the same enrichment and extraction approaches did not result in detection of the N and S viral biomarker proteins in the SARS-COV-2 positive patient swab samples.

Overall, the experimental results demonstrate that it is crucial to optimise the both the sample preparation and the MALDI matrix preparation, which includes the type of matrix used, solvents, mixing methods, and detergents to name a few. Optimisation of the sample pre-treatment steps applied to the standard viral SARS-CoV-2 proteins prior to MALDI-TOF analysis of the complex patient swab samples is also necessary. In future, the sample preparation methods discussed herein for viral protein enrichment and extraction can be expanded to include different assays, such as affinity capture beads, which target the critical biomarkers of SARS-CoV-2. Thus, further studies are required to explore the full potential of MALDI-TOF MS as a clinically useful tool for the screening of SARS-CoV-2.

5.6. References

(1) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N. Engl. J. Med. 2020, 382 (8), 727-733.

(2) Peiris, J. S. M.; Lai, S. T.; Poon, L. L. M.; Guan, Y.; Yam, L. Y. C.; Lim, W.; Nicholls, J.; Yee, W. K. S.; Yan, W. W.; Cheung, M. T.; Cheng, V. C. C.; Chan, K. H.; Tsang, D. N. C.; Yung, R. W. H.; Ng, T. K.; Yuen, K. Y. Coronavirus as a possible cause of severe acute respiratory syndrome. The Lancet 2003, 361 (9366), 1319-1325.

(3) Zaki, A. M.; van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D. M. E.; Fouchier, R. A. M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. N. Engl. J. Med. 2012, 367 (19), 1814-1820.

(4) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. Proc. Natl. Acad. Sci. U.S.A. 2020, 117 (21), 11727.

(5) Benton, D. J.; Wrobel, A. G.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. Nature 2020, 588 (7837), 327-330.

(6) McBride, R.; van Zyl, M.; Fielding, B. C. The coronavirus nucleocapsid is a multifunctional protein. Viruses 2014, 6 (8), 2991-3018.

(7) Siu, Y. L.; Teoh, K. T.; Lo, J.; Chan, C. M.; Kien, F.; Escriou, N.; Tsao, S. W.; Nicholls, J. M.; Altmeyer, R.; Peiris, J. S. M.; Bruzzone, R.; Nal, B. The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. J. Virol. 2008, 82 (22), 11318.

(8) Snijder, E. J.; Decroly, E.; Ziebuhr, J. In Adv. Virus Res., Ziebuhr, J., Ed.^Eds.; Academic Press, 2016, p^pp 59-126.

(9) Yadav, R.; Chaudhary, J. K.; Jain, N.; Chaudhary, P. K.; Khanra, S.; Dhamija, P.; Sharma, A.; Kumar, A.; Handu, S. Role of Structural and Non-Structural Proteins and Therapeutic Targets of SARS-CoV-2 for COVID-19. Cells 2021, 10 (4), 821.

(10) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu. Rev. Virol. 2016, 3 (1), 237-261.

(11) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 2020, 581 (7807), 215-220.

(12) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; Angyal, A.; Brown, R. L.; Carrilero, L.; Green, L. R., et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell 2020, 182 (4), 812-827.e819.

(13) Volz, E.; Hill, V.; McCrone, J. T.; Price, A.; Jorgensen, D.; O'Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F. F.; Rey, S. M.; Nicholls, S. M.; Colquhoun, R. M.; da Silva Filipe, A.; Shepherd, J.; Pascall, D. J.; Shah, R.; Jesudason, N.; Li, K.; Jarrett, R., et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell 2021, 184 (1), 64-75.e11.

(14) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J. E.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P. C.; Kyratsous, C. A.; Dudkina, N. V., et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell 2020, 183 (3), 739-751.e738.

(15) Bangaru, S.; Ozorowski, G.; Turner, H. L.; Antanasijevic, A.; Huang, D.; Wang, X.; Torres, J. L.; Diedrich, J. K.; Tian, J.-H.; Portnoff, A. D.; Patel, N.; Massare, M. J.; Yates, J. R.; Nemazee, D.; Paulson, J. C.; Glenn, G.; Smith, G.; Ward, A. B. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. Science 2020, 370 (6520), 1089.

(16) Dai, L.; Gao, G. F. Viral targets for vaccines against COVID-19. Nat. Rev. Immunol. 2021, 21 (2), 73-82.

(17) Du, L.; He, Y.; Zhou, Y.; Liu, S.; Zheng, B.-J.; Jiang, S. The spike protein of SARS-CoV — a target for vaccine and therapeutic development. Nat. Rev. Microbiol. 2009, 7 (3), 226-236.

(18) Salvatori, G.; Luberto, L.; Maffei, M.; Aurisicchio, L.; Roscilli, G.; Palombo, F.; Marra, E. SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. J. Transl. Med. 2020, 18 (1), 222.

(19) Tombuloglu, H.; Sabit, H.; Al-Suhaimi, E.; Al Jindan, R.; Alkharsah, K. R. Development of multiplex real-time RT-PCR assay for the detection of SARS-CoV-2. PLoS One 2021, 16 (4), e0250942.

(20) Lisboa Bastos, M.; Tavaziva, G.; Abidi, S. K.; Campbell, J. R.; Haraoui, L.-P.; Johnston, J. C.; Lan, Z.; Law, S.; MacLean, E.; Trajman, A.; Menzies, D.; Benedetti, A.; Ahmad Khan, F. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ 2020, 370, m2516.

(21) MacMullan, M. A.; Ibrayeva, A.; Trettner, K.; Deming, L.; Das, S.; Tran, F.; Moreno, J. R.; Casian, J. G.; Chellamuthu, P.; Kraft, J.; Kozak, K.; Turner, F. E.; Slepnev, V. I.; Le Page, L. M. ELISA detection of SARS-CoV-2 antibodies in saliva. Sci. Rep. 2020, 10 (1), 20818.

(22) Singhal, N.; Kumar, M.; Kanaujia, P. K.; Virdi, J. S. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front. Microbiol. 2015, 6, 791-791.

(23) Cherkaoui, A.; Hibbs, J.; Emonet, S.; Tangomo, M.; Girard, M.; Francois, P.; Schrenzel, J. Comparison of Two Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry Methods with Conventional Phenotypic Identification for Routine Identification of Bacteria to the Species Level. J. Clin. Microbiol. 2010, 48 (4), 1169.

(24) La Scola, B.; Raoult, D. Direct Identification of Bacteria in Positive Blood Culture Bottles by Matrix-Assisted Laser Desorption Ionisation Time-of-Flight Mass Spectrometry. PLoS One 2009, 4 (11), e8041.

(25) Thiede, B.; Höhenwarter, W.; Krah, A.; Mattow, J.; Schmid, M.; Schmidt, F.; Jungblut, P. R. Peptide mass fingerprinting. Methods 2005, 35 (3), 237-247.

(26) Iles, R. K.; Zmuidinaite, R.; Iles, J. K.; Carnell, G.; Sampson, A.; Heeney, J. L. A clinical MALDI-ToF Mass spectrometry assay for SARS-CoV-2: Rational design and multi-disciplinary team work. medRxiv 2020, 2020.2008.2022.20176669.

(27) Ihling, C.; Tänzler, D.; Hagemann, S.; Kehlen, A.; Hüttelmaier, S.; Sinz, A. Mass Spectrometric Identification of SARS-CoV-2 Proteins from Gargle Solution Samples of COVID-19 Patients. bioRxiv 2020, 2020.2004.2018.047878.

(28) Nikolaev, E. N.; Indeykina, M. I.; Brzhozovskiy, A. G.; Bugrova, A. E.; Kononikhin, A. S.; Starodubtseva, N. L.; Petrotchenko, E. V.; Kovalev, G. I.; Borchers, C. H.; Sukhikh, G. T. Mass-Spectrometric Detection of SARS-CoV-2 Virus in Scrapings of the Epithelium of the Nasopharynx of Infected Patients via Nucleocapsid N Protein. J. Proteome Res. 2020, 19 (11), 4393-4397.

(29) Smolira, A.; Wessely-Szponder, J. Importance of the Matrix and the Matrix/Sample Ratio in MALDI-TOF-MS Analysis of Cathelicidins Obtained from Porcine Neutrophils. Applied Biochemistry and Biotechnology 2015, 175 (4), 2050-2065.

(30) Kussmann, M.; Roepstorff, P. Characterisation of the covalent structure of proteins from biological material by MALDI mass spectrometry ‣ possibilities and limitations. Spectroscopy 1998, 14, 710163.

(31) Beavis, R. C.; Chait, B. T.; Fales, H. M. Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. Rapid Commun. Mass Spectrom. 1989, 3 (12), 432-435.

(32) Kim, Y. J.; Freas, A.; Fenselau, C. Analysis of Viral Glycoproteins by MALDI-TOF Mass Spectrometry. Anal. Chem. 2001, 73 (7), 1544-1548.

(33) Wortmann, A.; Pimenova, T.; Alves, S.; Zenobi, R. Investigation of the first shot phenomenon in MALDI mass spectrometry of protein complexes. Analyst 2007, 132 (3), 199-207.

(34) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal. Chem. 1988, 60 (20), 2299-2301.

(35) Kussmann, M.; Nordhoff, E.; Rahbek-Nielsen, H.; Haebel, S.; Rossel-Larsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskaya, E.; Kroll-Kristensen, A.; Palm ‖ , L.; Roepstorff, P. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry Sample Preparation Techniques Designed for Various Peptide and Protein Analytes. J. Mass Spectrom. 1997, 32 (6), 593-601.

(36) Dai, Y.; Whittal, R. M.; Li, L. Confocal Fluorescence Microscopic Imaging for Investigating the Analyte Distribution in MALDI Matrices. Anal. Chem. 1996, 68 (15), 2494-2500.

(37) Shehadul Islam, M.; Aryasomayajula, A.; Selvaganapathy, P. R. A Review on Macroscale and Microscale Cell Lysis Methods. Micromachines (Basel) 2017, 8 (3), 83.

(38) Branck, T. A.; Hurley, M. J.; Prata, G. N.; Crivello, C. A.; Marek, P. J. Efficacy of a Sonicating Swab for Removal and Capture of Listeria monocytogenes in Biofilms on Stainless Steel. Applied and Environmental Microbiology 2017, 83 (11), e00109-00117.

(39) Ahnrud, G. P.; Mendoza, A. J.; Hurley, M. J.; Marek, P. J. Efficacy of a Sonicating Swab for Removal and Capture of Microorganisms from Experimental and Natural Contaminated Surfaces. Applied and Environmental Microbiology 2018, 84 (9), e00208-00218.

(40) Keeshen, T.; Case, J. B.; Wellehan, J. F.; Dujowich, M. Bacterial recovery using sonication versus swabbing of titanium and stainless steel implants inoculated with Staphylococcus pseudintermedius or Pseudomonas aeruginosa. Vet. Comp. Orthop. Traumatol. 2017, 30 (5), 346-350.

(41) Smyrlaki, I.; Ekman, M.; Lentini, A.; Rufino de Sousa, N.; Papanicolaou, N.; Vondracek, M.; Aarum, J.; Safari, H.; Muradrasoli, S.; Rothfuchs, A. G.; Albert, J.; Högberg, B.; Reinius, B. Massive and rapid COVID-19 testing is feasible by extraction-free SARS-CoV-2 RT-PCR. Nature Communications 2020, 11 (1), 4812.

(42) Wiśniewski, J. R.; Zielinska, D. F.; Mann, M. Comparison of ultrafiltration units for proteomic and N-glycoproteomic analysis by the filter-aided sample preparation method. Anal. Biochem. 2011, 410 (2), 307-309.

(43) Chen, I. H.; Xiao, H.; Daly, T.; Li, N. Improved Host Cell Protein Analysis in Monoclonal Antibody Products through Molecular Weight Cutoff Enrichment. Anal. Chem. 2020, 92 (5), 3751-3757.

(44) Van Huynh, N.; De Backer, O.; Decleire, M.; Colson, C. A procedure for the preparation of bacterial DNA that employs dimethyl sulfoxide to induce the lysis of cells. Anal. Biochem. 1989, 176 (2), 464-467.

(45) Wang, G.; Gong, Y.; Burczynski, F. J.; Hasinoff, B. B. Cell lysis with dimethyl sulphoxide produces stable homogeneous solutions in the dichlorofluorescein oxidative stress assay. Free Radical Research 2008, 42 (5), 435-441.

(46) Yuan, C.; Gao, J.; Guo, J.; Bai, L.; Marshall, C.; Cai, Z.; Wang, L.; Xiao, M. Dimethyl Sulfoxide Damages Mitochondrial Integrity and Membrane Potential in Cultured Astrocytes. PLoS One 2014, 9 (9), e107447.

5.7 Supplementary Information

Table S5. 1 Results and summary table of spike protein S1 and S2 subunit optimisation experiments obtained on the Bruker microflex MALDI-TOF MS.

Experiment	Details	Spike protein S1 Unit			Spike protein S2 Unit			Summary
		Detect (%)	Ave. Intensity (a.u)	Ave. S/N	Detect (%)	Ave. Intensity (a.u)	Ave. S/N	
Mixing methods	Mix on MALDI plate (direct spotting)	0/4 (0%)	N/A	N/A	2/4 (50%)	198.5	9.5	Protein and matrix mixed in 96-well plate provides a more constant detection percentage in both instruments.
	Mix in 96-well plate	0/4 (0%)	N/A	N/A	4/4 (100%)	343 ± 115.87	19.25 ± 6.1	
Protein dilution solvent	H_2O	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	50:50 ACN:0.1% TFA solution is required to dilution the spike protein unit 1 and 2.
	H_2O/0.1% TFA	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	H_2O /0.1% TFA/ACN	0/4 (0%)	N/A	N/A	4/4 (100%)	343 ± 115.87	19.25 ± 6.1	
Matrix optimisation	α-Cyano-4-hydroxycinnamic acid (CHCA)	0/6 (0%)	N/A	N/A	0/6 (0%)	N/A	N/A	Sinapinic Acid (SA) has been reported to provide consistent signal detection for large proteins but Ferulic Acid (FA) matrix provides the best signal for both S1 and S2 units of spike protein in these experiment
	sinapinic acid (SA)	5/6 (83.3%)	93.2 ± 8.1	2	6/6 (100%)	48.7 ± 23.6	1283.1 ± 659.6	
	2-nitrophloroglucinol (2-NPG)	0/2 (0%)	N/A	N/A	2/2 (100%)	6	289.5	
	2,5-Dihydroxybenzoic acid (2,5-	0/6 (0%)	N/A	N/A	0/6 (0%)	N/A	N/A	

Experiment	Details	Spike protein S1 Unit			Spike protein S2 Unit			Summary
		Detect (%)	Ave. Intensity (a.u)	Ave. S/N	Detect (%)	Ave. Intensity (a.u)	Ave. S/N	
	DHB)							s. However, FA requires manual control on MALDI as uneven co-crystallisation is observed with FA. With SA, co-crystallisation of sample and matrix is more homogenous so automatic run is possible.
	2-(4-Hydroxyphenylazo)benzoic acid (HABA)	2/2 (100%)	169.5	2	2/2 (100%)	289.5	4	
	2,5-DHB+ CHCA	0/2 (0%)	N/A	N/A	1/2 (50%)	171	2	
	Ferulic acid (FA) (4-hydroxy-3-methoxycinnamic acid)	2/2 (100%)	225	6.5	2/2 (100%)	79	2263.5	
	2,6-dihydroxyacetophenone (DHAP)	0/2 (0%)	N/A	N/A	0/2 (0%)	N/A	N/A	
	Super-DHB (a mixture of 2,5-DHB and the additive 2-hydroxy-5-methoxybenzoic acid.)	0/6 (0%)	N/A	N/A	0/6 (0%)	N/A	N/A	
Sinapinic acid (SA) matrix conc. optimisation (matrix: sample)	Matrix conc.:							The best signal obtained for both S1 and S2 units is mixing the 500 fmol/µL protein
	1:1 ratio, 20 mg/mL SA	0/3 (0%)	N/A	N/A	3/3 (100%)	204 ± 66.6	9.3 ± 3.7	
	2:1 ratio, 20	1/3 (33%)	71	2	3/3 (100%)	271.6 ± 125.8	12.6 ± 6.5	

Experiment	Details	Spike protein S1 Unit			Spike protein S2 Unit			Summary
		Detect (%)	Ave. Intensity (a.u)	Ave. S/N	Detect (%)	Ave. Intensity (a.u)	Ave. S/N	
	mg/mL SA							solutions with 40 mg/mL SA matrix in 2:1 ratio.
	2:1 ratio, 40 mg/mL SA	3/3 (100%)	81.8 ± 0.08	2 ± 0	3/3 (100%)	619.3 ± 380.4	23.0 ± 8.8	
	2:1 ratio, 60 mg/mL SA	3/3 (100%)	89.7 ± 7.0	2.33 ± 0.47	3/3 (100%)	345.6 ± 164.6	15.7 ± 7.3	
Ferulic acid (FA) matrix conc. optimisation	1:1 ratio, 30 mg/mL FA	2/2 (100%)	427	12.5	2/2 (100%)	2615	80	The best signal for both S1 and S2 units are achieved by mixing 30 mg/mL FA solution with 500 fmol/µL protein in 1:1 ratio.
	2:1 ratio, 60 mg/mL FA	0/2 (0%)	N/A	N/A	2/2 (100%)	238	2.5	
Effect of matrix reaction time and temperature	Method:							Incubation for 30 mins at 37°C can improve S1 protein detection using SA matrix but destroys the signal for both S1 and S2 using FA matrix. The best signal is achieved by using FA matrix and spots immediately on MALDI plate after mixing.
	SA: Immediately	5/6 (83.3%)	93.2 ± 8.1	2 ± 0	6/6 (100%)	1283.1 ± 659.6	48.7 ± 23.6	
	SA: Incubated 30 mins	6/6 (100%)	146.3 ± 29.0	4.3 ± 1.2	6/6 (100%)	784.6 ± 388.0	33.8 ± 15.7	
	FA: Immediately	2/2 (100%)	427	12.5	2/2 (100%)	2615	80	
	FA: Incubated 30 mins	0/2 (0%)	N/A	N/A	0/2 (0%)	N/A	N/A	

Experiment	Details	Spike protein S1 Unit			Spike protein S2 Unit			Summary
		Detect (%)	Ave. Intensity (a.u)	Ave. S/N	Detect (%)	Ave. Intensity (a.u)	Ave. S/N	
Limit of detection	Matrix conc. (fmol/µL):							1.The purity of S1 unit provided by the company is ~ 80%, thus the LOD achieved is 100 fmol/µL * 0.8 = 80 fmol/µL. 2.The purity of S2 unit provided by the company is ~ 95%, thus the LOD achieved is 75 fmol/µL * 0.95 = 71.3 fmol/µL. 3.FA matrix increases the sensitivity of detecting S1 and S2 units.
FA	50	2/12 (16.7%)	97.7	2	4/12 (33.3%)	107.4 ± 11.3	2.75 ± 0.4	
	75	8/12 (66.7%)	99.3 ± 9.9	2.25 ± 0.43	12/12 (100%)	139.8 ± 24.0	4.6 ± 1.3	
	100	12/12 (100%)	112.8 ± 10.2	2.91 ± 0.3	12/12 (100%)	329.7 ± 116.3	13.8 ± 5.3	
	500	12/12 (100%)	308.1 ±71.4	9.75 ± 2.8	12/12 (100%)	1618.3 ± 902.5	64.7 ±32.5	
SA	50	0/6 (0%)	N/A	N/A	0/6 (0%)	N/A	N/A	
	250	0/6 (0%)	N/A	N/A	8/12 (66.7%)	142 ± 58.5	5.2 ± 3.7	
	500	10/12 (83.3%)	105.7 ± 9.1	3.1 ± 0.5	12/12 (100%)	359.2 ± 86.6	15.4 ± 4.2	

Table S5. 2 Results and summary table of the N-protein and the S-protein optimisation experiments obtained on the Bruker microflex MALDI-TOF MS.

Experiment	Details	N-Protein			S-Protein			Summary
		Detect (%)	Ave. Intensity (a.u)	Ave. S/N	Detect (%)	Ave. Intensity (a.u)	Ave. S/N	
Mixed detection of N and S protein mix (0.5 µM) in SA (40 mg/mL)	Protein:							N-protein is easier to detect in the mixed sample as S-protein is at a much higher *m/z* ratio and there is also the ion suppression effect resulting from increased signal detection of the N-protein.
	N-Protein Only (0.5 µM)	4/4 (100%)	1919.97 ±285.76	602.21 ±19.18	0/4 (0%)	N/A	N/A	
	S-Protein Only (0.5 µM)	0/4 (0%)	N/A	N/A	4/4 (100%)	12.017 ± 1.56	4.64 ± 1.01	
	N- & S-Protein Mixed (1:1) (0.5 µM)	4/4 (100%)	3816.56 ± 617.12	783.89 ± 72.1	0/4 (0%)	N/A	N/A	
Further mixing method optimisation of N and S protein mixture (0.5 µM):	Matrix and order of the matrix:							Best N-protein signal detected using SA (40 mg/mL) with mixing in 96-well plate before spotting on MALDI target plate.
Mix in 96-well plate	sinapinic acid (SA)	4/4 (100%)	4136.08 ± 1584.05	649.82 ± 209.55	0/4 (0%)	N/A	N/A	
	Ferulic acid (FA)	4/4 (100%)	710.80 ± 172.72	221.49 ± 42.06	0/4 (0%)	N/A	N/A	
	SA+FA (1:1)	4/4 (100%)	3449.16 ± 356.62	670.82 ± 48.41	0/4 (0%)	N/A	N/A	
Dried sandwich spotting method	SA (bottom layer) -> sample (middle	4/4 (100%)	1162.12 ± 508.56	303.72 ± 135.62	0/4 (0%)	N/A	N/A	

	layer) ->FA (top layer)							
	FA (bottom layer) -> sample (middle layer) ->SA (top layer)	4/4 (100%)	2076.67 ± 195.26	514.32 ± 72.64	0/4 (0%)	N/A	N/A	
	SA (bottom layer) -> sample (middle layer) ->SA (top layer)	4/4 (100%)	1558.13 ± 311.6	255.80± 30.52	0/4 (0%)	N/A	N/A	
	FA (bottom layer) -> sample (middle layer) ->FA (top layer)	4/4 (100%)	1753.61 ± 164.98	478.92 ± 124.52	0/4 (0%)	N/A	N/A	
	FA+SA (bottom layer) -> sample (middle layer) -> FA+SA (top layer)	4/4 (100%)	1335.33 ± 641.87	381.26 ± 204.11	0/4 (0%)	N/A	N/A	
N and S protein mix in SA (40 mg/mL) TFA % optimisation	TFA (%):							0.05 % TFA gives highest signal detection of N-protein in N and S protein mix. However, variation is high (st.
	0.05	4/4 (100%)	5003 ± 1005.4	151.6 ± 14.7	0/4 (0%)	N/A	N/A	
	0.1	4/4 (100%)	2473.8 ± 415.1	93.8 ± 11.7	0/4 (0%)	N/A	N/A	
	0.5	4/4 (100%)	1537 ± 339.7	67.2 ± 10.6	0/4 (0%)	N/A	N/A	
	1	4/4	1070.8 ±	49.7	0/4			

		(100%)	200.3	± 15.2	(0%)			dev. 1005.4) so 0.1% TFA (standard TFA concentration in most MALDI publications) is used.
N and S protein mix (0.5 µM) in SA (40 mg/mL) matrix solvent variation	Solvents (50/50) with 0.1% TFA:							Highest signal intensity of N protein in N and S mixture with SA is in the matrix solvent of water/acetonitrile (50/50).
	Ammonium phosphate (10 mM) / Acetonitrile	4/4 (100%)	3610.93 ± 1581.03	810.28 ± 84.82	0/4 (0%)	N/A	N/A	
	Methanol / Acetonitrile	4/4 (100%)	544.86 ± 182.1	221.67 ± 73.56	0/4 (0%)	N/A	N/A	
	Ethanol / Acetonitrile	4/4 (100%)	430.56 ± 155.84	176.57 ± 63.05	0/4 (0%)	N/A	N/A	
	Chloroform / Methanol	4/4 (100%)	1051.71 ± 629.5	324.08 ± 149.18	0/4 (0%)	N/A	N/A	
	Dichloromethane/ Methanol	4/4 (100%)	2907.28 ± 1292.42	695.64 ± 140.05	0/4 (0%)	N/A	N/A	
	Toluene / Methanol	4/4 (100%)	2908.05 ± 350.72	727.26 ± 96.11	0/4 (0%)	N/A	N/A	
	Hexane / Ethanol	4/4 (100%)	220.41 ± 47.1	75.64± 17.6	0/4 (0%)	N/A	N/A	
	Acetone / Methanol	4/4 (100%)	2192.07 ± 887.69	657.33 ± 157.66	0/4 (0%)	N/A	N/A	
	Water /Acetonitrile	4/4 (100%)	3614.89 ± 518.6	813.80 ± 72.53	0/4 (0%)	N/A	N/A	
Addition of	Polyme							No protein

polymer to the matrix ferulic acid (FA) 30 mg/mL	r type and concentration:							detected because the addition of the polymers resulted in the sample solution to run on MALDI target plate. The hydrophobic ring around the barrier of each MALDI spot is broken due to the polymers added. Final concentration of analyte may be too low for detection.
	TWEEN 80 0.01%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	TWEEN 80 0.1%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	TWEEN 80 1%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	Triton-X 100 0.01%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	Triton-X 100 0.1%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	Triton-X 100 1%	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
Detection of N- & S-Protein in sputum and oral fluid with matrix SA (40 mg/mL)	Sample dilution media:							N S-protein detected but N-protein was detected at low intensity in the oral fluid only.
	Standard Sample (no dilution with media)	4/4 (100%)	2821.44 ± 381.5	829.11 ± 89.33	0/4 (0%)	N/A	N/A	
	Diluted with Sputum	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	Diluted with Oral Fluid	4/4 (100%)	46.86 ± 20.09	12.71 ± 4.5	0/4 (0%)	N/A	N/A	
N and S protein mixture in synthetic sputum solvent extraction with matrix SA (40	Solvent used for extraction:							Highest signal detection of N-protein in toluene aqueous layer. No/very
	chloroform aqueous layer	4/4 (100%)	3658.5 ± 521.9	130.8 ± 14.5	0/4 (0%)	N/A	N/A	

mg/mL)	dichloromethane aqueous layer	4/4 (100%)	3954.3 ± 425.4	135.7 ± 11.9	0/4 (0%)	N/A	N/A	low protein detected in solvent organic layer.
	hexane aqueous layer	4/4 (100%)	1797.8 ± 261.4	79.5 ± 11.4	0/4 (0%)	N/A	N/A	
	toluene aqueous layer	4/4 (100%)	4254.5 ± 286.3	144.4 ± 12.5	0/4 (0%)	N/A	N/A	
COVID-19 test kit spiked with N and S protein	Protein detection from:							Difficult to detect N and S protein spiked using COVID-19 test kit. Only cotton bud sonication with ACN - able to detect very low intensity of N-protein. Further enrichment steps required.
	1. Cotton Bud directly	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	2. Cotton Bud sonicated with ACN	yes (low)	N/A	N/A	0/4 (0%)	N/A	N/A	
	3. ACN sonicated solution	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	
	4. Testing kit solution	0/4 (0%)	N/A	N/A	0/4 (0%)	N/A	N/A	

Table S5. 3 Table of instrument parameters for the Bruker Microflex MALDI-TOF MS experiments.

Detection parameters	**Values**
Laser Power	90%
Mass Range	5k – 300k
Detector Gain	x20
Attenuator offset	35%
Number of scans	2,000
Sampling	Random
Shots/raster spot	50

Table S5. 4 Peak assignment table for the standard N-protein tryptic peptide $[GFYAEGSR]^{2+}$ CAD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_1	1+	$C_6 H_{15} N_4 O_2$	175.118952	175.118857	-0.54
a_2	1+	$C_{10} H_{13} N_2 O_1$	177.102239	177.102239	0.00
b_2	1+	$C_{11} H_{13} N_2 O_2$	205.097154	205.09715	-0.02
YA	1+	$C_{12} H_{15} N_2 O_3$	235.107719	235.107567	-0.65
y_2	1+	$C_9 H_{20} N_5 O_4$	262.150981	262.150872	-0.42
y_3	1+	$C_{11} H_{23} N_6 O_5$	319.172444	319.172444	0.00
a_3-NH_3	1+	$C_{19} H_{19} N_2 O_3$	323.139019	323.139038	0.06
y_6	2+	$C_{28} H_{44} N_9 O_{11}$	341.661378	341.661349	-0.08
b_3	1+	$C_{20} H_{22} N_3 O_4$	368.160483	368.160491	0.02
a_4	1+	$C_{22} H_{27} N_4 O_4$	411.202682	411.202763	0.20
MH-H_2O	2+	$C_{22} H_{27} N_4 O_4$	434.701035	434.701022	-0.03
b_4	1+	$C_{23} H_{27} N_4 O_5$	439.197596	439.1978	0.46
MH	2+	$C_{39} H_{57} N_{11} O_{13}$	443.706317	443.706305	-0.03
y_4	1+	$C_{16} H_{30} N_7 O_8$	448.215037	448.215361	0.72
y_5	1+	$C_{19} H_{35} N_8 O_9$	519.252151	519.252051	-0.19
b_5	1+	$C_{28} H_{34} N_5 O_8$	568.240189	568.239784	-0.71
y_6	1+	$C_{28} H_{44} N_9 O_{11}$	682.31548	682.315488	0.01
Average error					-0.07
Absolute average error					0.24
Standard deviation					0.27

Table S5. 5 Peak assignment table for the standard N-protein tryptic peptide [AYNVTQAFGR]$^{2+}$ CAD MS/MS spectrum.

Assignment	Charge state	Elemental composition	Theoretical *m/z*	Observed *m/z*	Mass error (ppm)
y_1	1+	$C_6 H_{15} N_4 O_2$	175.118952	175.118842	-0.63
a_2	1+	$C_{11} H_{15} N_2 O_2$	207.112804	207.112804	0.00
y_2	1+	$C_8 H_{18} N_5 O_3$	232.140416	232.140459	0.19
b_2	1+	$C_{12} H_{15} N_2 O_3$	235.107719	235.107819	0.43
a_3	1+	$C_{15} H_{21} N_4 O_4$	321.155732	321.15571	-0.07
b_3	1+	$C_{16} H_{21} N_4 O_5$	349.150646	349.150712	0.19
y_3	1+	$C_{17} H_{27} N_6 O_4$	379.20883	379.208839	0.02
a_4	1+	$C_{20} H_{30} N_5 O_5$	420.224146	420.224174	0.07
b_4	1+	$C_{21} H_{30} N_5 O_6$	448.21906	448.219046	-0.03
y_4	1+	$C_{20} H_{32} N_7 O_5$	450.245944	450.245892	-0.12
b_5	1+	$C_{25} H_{37} N_6 O_8$	549.266739	549.266608	-0.24
MH-H_2O	2+	$C_{50} H_{74} N_{15} O_{14}$	554.780347	554.779933	-0.75
MH	2+	$C_{50} H_{77} N_{15} O_{15}$	563.78563	563.785546	-0.15
y_5	1+	$C_{25} H_{40} N_9 O_7$	578.304521	578.304949	0.74
b_6	1+	$C_{30} H_{45} N_8 O_{10}$	677.325316	677.325601	0.42
y_6	1+	$C_{29} H_{47} N_{10} O_9$	679.3522	679.351951	-0.37
b_7	1+	$C_{33} H_{50} N_9 O_{11}$	748.36243	748.36291	0.64
y_7	1+	$C_{34} H_{56} N_{11} O_{10}$	778.420613	778.420516	-0.12
y_8	1+	$C_{38} H_{62} N_{13} O_{12}$	892.463541	892.463551	0.01
b_8	1+	$C_{42} H_{59} N_{10} O_{12}$	895.430844	895.430329	-0.58
Average error					-0.02
Absolute average error					0.29
Standard deviation					0.26

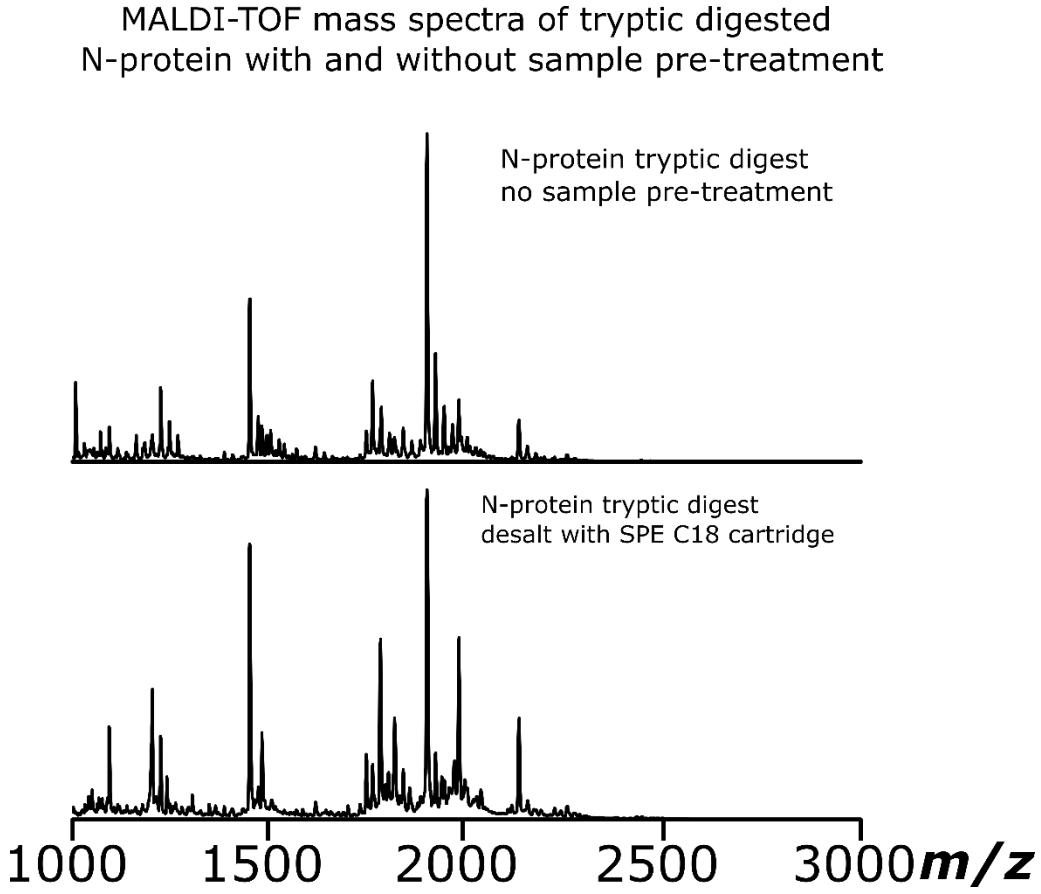


Figure S5. 1 MALDI-TOF MS of tryptic digested N-protein with no sample pre-treatment (top spectrum) and the tryptic digested N-protein desalted with the SPE C18 cartridge (bottom spectrum).

6. Conclusions and Future Work

The work presented in this thesis has demonstrated the implementation of mass spectrometry (MS) together with various advanced fragmentation methods for the differentiation and relative quantification of biologically significant isomeric species. In addition, MS studies and viral enrichment optimisation experiments were conducted on SARS-CoV-2 proteins for the improved detection of viral biomarkers obtained from human swab samples. This chapter aims to provide a summary, conclusion, and final outlook for each of the experimental results chapters previously discussed.

Chapter 2: Differentiation and Relative Quantification of the Isomeric Products of Deamidation using ECD and UVPD Tandem Mass Spectrometry

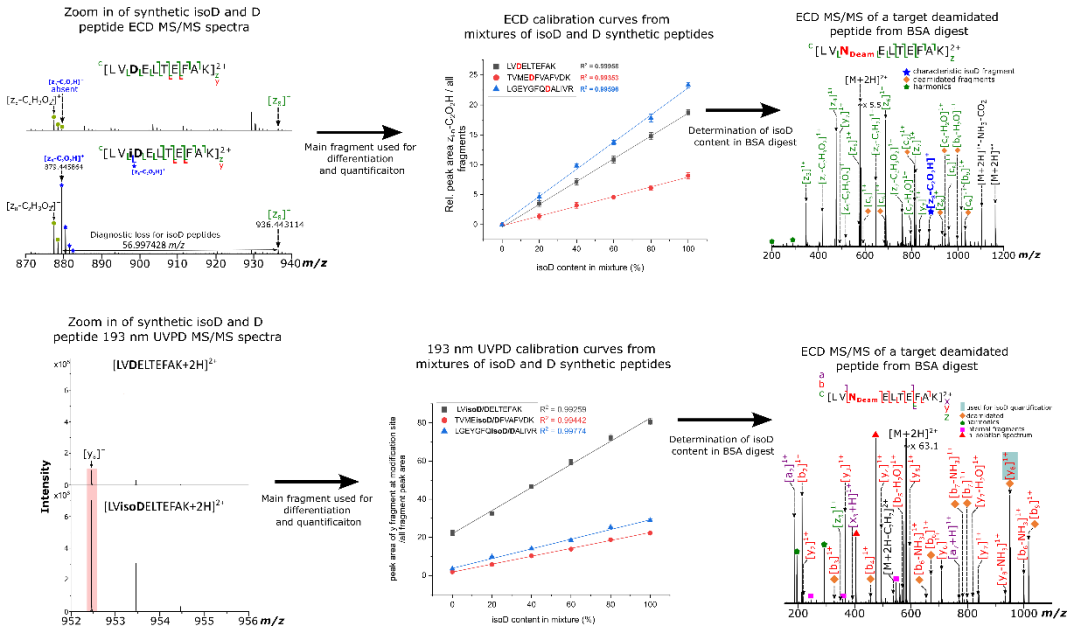


Figure 6. 1 Summary figure for chapter 2, illustrating the fragments used for isoD and D peptide differentiation and quantification using ECD MS/MS and 193 nm UVPD MS/MS, resulting in the determination of the isoD percentage content of three target peptides in deamidated BSA digest mixture samples.

The process and products of asparagine (N) deamidation, aspartic acid (D) and isoaspartic acid (isoD) are associated with significant neurological disorders, such as Alzheimer's and Parkinson's disease. The main aim of chapter 2 was to fully utilise ECD MS/MS and UVPD MS/MS on a 12 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) to distinguish between the isomeric deamidation products, isoD and D, via generation of diagnostic fragments, which can then be used for the relative quantification of isoD in a deamidated tryptic digested protein sample. Thus, improved methods for the relative quantification using diagnostic MS/MS fragments were developed to determine the percentage content of isoD, which has been previously identified and implicated as a biomarker of the neurological diseases mentioned herein.

Tryptic digested bovine serum albumin (BSA) was used as a model protein for this study and a series of deamidated BSA peptides were observed in the MS, with three deamidated peptides chosen as the key deamidation target peptides. This is because the selected peptides were sufficiently deamidated during the accelerated ageing experiments, which could be used for the relative quantification experiments. Synthetic isoD and D peptide standards were purchased for each target BSA peptide and diagnostic z_{n-i}-C_2O_2H fragment ions were detected in the ECD MS/MS spectra for all isoD peptides, which were absent for the corresponding D peptides. Although the z_{n-i}-C_2O_2H fragment ion was not detected in the UVPD MS/MS spectra for the isoD peptides, a significant difference in the intensities of the y fragment generated at the specific isoD and D positions in the peptide sequence by UVPD can be used to discriminate between the isomeric peptides. This is due to the isoD peptides containing higher y-ion intensities at the deamidated sites compared to the D peptides.

A modified and improved relative quantification method for isoD using ECD and UVPD was demonstrated. Herein, we improved the ECD quantification method, determining the percentage isoD content based on the peak area of the characteristic z_{n-i}-C_2O_2H fragment ion fragment divided by the sum of all the fragment peak areas. Furthermore, the UVPD relative quantification method was based on the ratio of the peak area of the y fragment generated at the deamidation modification site to the sum of all the fragment peak areas. A good linearity ($R^2 > 0.99$) was obtained in all calibration curves of synthetic peptides using ECD and UVPD.

The relative direct infusion ECD and UVPD quantification results were also compared to those obtained via nano-LC ECD MS/MS; and overall, the direct infusion ECD provides the best results due to the detection of diagnostic fragments as for direct

infusion UVPD, although the difference in relative intensities of isoD and D peptides is significant, other factors such as instrument stability and sample infusion concentration may affect the observed intensities, which would ultimately affect the reliability of the quantification results. Nano-LC ECD MS/MS analysis of the deamidated peptides on the other hand, requires time for optimisation of LC parameters such as the gradient and column conditions to ensure sufficient separation of the isomeric forms for relative quantification.

In conclusion, although no modifications prior to analysis of the deamidated peptides were necessary, synthetic standards of the deamidated target peptides aided the direct infusion MS/MS quantification experiments, with direct infusion ECD proving to be a fast and reliable fragmentation method and UVPD as an alternative and easily applied fragmentation method for the relative quantification of isoD in the tryptic peptides of BSA.

Chapter 3: Distinguishing between methylated histidine isomers generated as a post-translational modification of actin

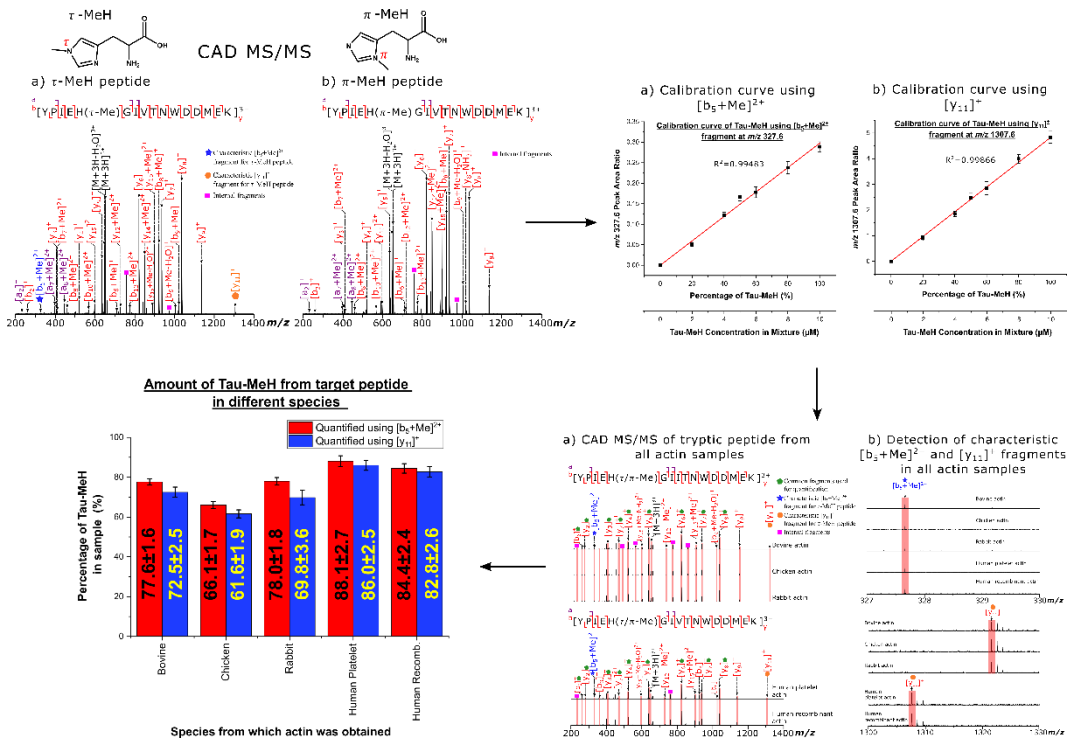


Figure 6. 2 Summary figure for chapter 3, demonstrating the CAD MS/MS detection of diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions fragments used for τ-MeH and π-MeH peptide differentiation and quantification, resulting in the determination of the τ-MeH percentage content of the target actin peptide in the mixture samples of actin digest obtained from different species.

Chapter 3 focuses on the application of various fragmentation methods, available on a 12 T FT-ICR MS, for the differentiation and relative quantification of isomeric N-methylated histidine containing peptides from the cytoskeletal protein, actin.

Actin methylation, specifically N-methylation at the histidine-73 (H73) residue in actin, has been identified as a regulatory mechanism, which contributes to the function of the cytoskeletal protein. Previous studies have shown that H73 N-methylation of the isolated target actin peptide (YPIEH(Me)GIVTNWDDMEK) can result in the formation of tele- or pros-methylhistidine (τ-MeH or π-MeH), which are difficult to differentiate between using analytical techniques due to the minor changes in their structures and zero mass difference between the residues, as they are isomers.

In this study, the τ-MeH and π-MeH target actin peptides were subjected to various fragmentation methods, including CAD, ExD, and photodissociation methods. The experimental results have successfully demonstrated the differentiation of the isomeric τ-MeH and π-MeH actin peptides via the generation of diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ MS/MS fragment ions for the τ-MeH peptide, which were absent in the π-MeH peptide CAD, IRMPD, and UVPD MS/MS spectra.

Based on the detection of the fragments unique to the τ-MeH peptide, a relative quantification method was developed using CAD MS/MS. Isomeric mixtures of the target peptide were prepared and a linear trend was observed between the relative intensity of the diagnostic fragment ions and the τ-MeH content in the synthetic τ-MeH and π-MeH peptide mixtures. A ratio was taken of each diagnostic fragment peak area to the sum of the common fragments in the MS/MS spectra, which were plotted against the τ-MeH percentage content in the synthetic peptide mixtures. Calibration curves were achieved with good linearity ($R^2 > 0.99$) using the diagnostic $[b_5+Me]^{2+}$ and $[y_{11}]^{+}$ fragment ions.

The linear calibration curves were then applied to quantify the relative τ-MeH and π-MeH content in the target peptide of 5 types of actin samples obtained from different species including bovine, chicken, rabbit, and human actin. The relative τ-MeH quantification results show that the τ-MeH form is the dominant isomeric form in all mammalian actin samples studied herein, with the highest percentage of the τ-MeH content in the target peptide detected in the human actin samples (> 80 %). Overall, the results also highlight the benefits of utilising direct infusion fragmentation approaches for relative quantification of isomeric species, which can also be applied to the complex actin digest mixtures obtained from various species.

Chapter 4: Exploring tandem mass spectrometry methods for the analysis of dihydroxylated vitamin D_3 isomers

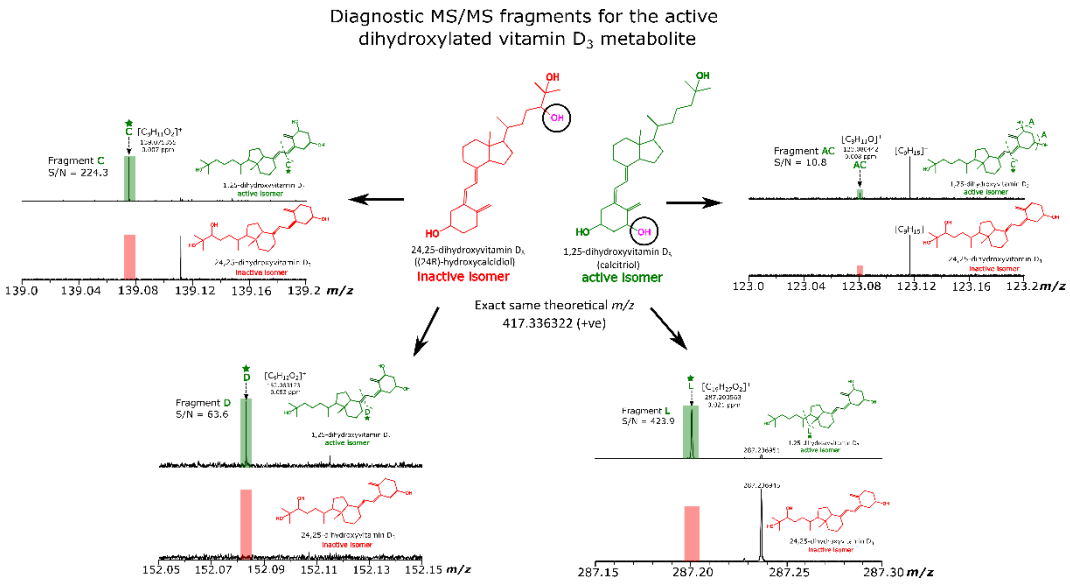


Figure 6. 3 Summary figure for chapter 4, showing the structures of the dihydroxylated vitamin D_3 isomers and the comparative IRMPD MS/MS spectra for both metabolites, highlighting four of the various diagnostic fragments obtained for 1,25-dihydroxyvitamin D_3 (biologically active), which were absent in the 24,25-dihydroxyvitamin D_3 (inactive form).

Chapter 4 explores the application of all available fragmentation methods on the 12 T FT-ICR MS such as CAD, IRMPD, UVPD and EID MS/MS for the differentiation of the isomeric dihydroxylated vitamin D_3 compounds.

Vitamin D compounds are a group of secosteroids derived from cholesterol, which are vital for maintaining bone health in humans. Recent studies have shown extraskeletal effects of vitamin D, involving vitamin D metabolites such as the dihydroxylated vitamin D_3 compounds 1,25-dihydroxyvitamin D_3 and 24,25-dihydroxyvitamin D_3. Differentiation and characterization of these isomers by mass spectrometry can be challenging due to the zero-mass difference and subtle structural differences between them. The isomers usually require separation by liquid chromatography (LC) prior to mass spectrometry, which adds extra complexity to the analysis.

Isomer-specific fragments were observed for the 1,25-dihydroxyvitamin D_3, which were absent in the 24,25-dihydroxyvitamin D_3 MS/MS spectra using all

fragmentation methods mentioned herein. The structure-specific fragments generated due to cleavage of the C-6/C-7 bond in the 1,25-dihydroxyvitamin D_3 compound successfully demonstrate the retention of the fragile hydroxyl groups during dissociation using all the available fragmentation methods.

It should be noted that the loss of the hydroxyl groups and series of hydrocarbon chain decompositions for both vitamin D_3 metabolites dominate all the MS/MS spectra obtained and therefore detailed analysis of the MS/MS spectra is required. Nevertheless, multiple diagnostic fragments were detected and assigned with high confidence, aided by the high resolving power and high mass accuracy capabilities provided by FT-ICR MS.

In summary, diagnostic fragments generated via all MS/MS methods were observed for 1,25-dihydroxyvitamin D_3, enabling differentiation between the two dihydroxylated vitamin D_3 isomers, without the need for prior chromatographic separation or derivatization of the molecules.

The development of vitamin D quantification methods are essential for the determination of vitamin D status in humans. Preliminary experiments for the quantitative analysis of 1,25-dihydroxyvitamin D_3 were also carried out. Herein, a quantification method was developed, which uses the peak area of a selected diagnostic fragment of 1,25-dihydroxyvitamin D_3 divided by the sum of the fragments peak areas to reduce the fluctuation caused by a single fragmented peak. A calibration curve using the diagnostic fragments of 1,25-dihydroxyvitamin D_3 was established with good linearity ($R^2 > 0.99$).

For future work, application of this direct infusion MS/MS quantification method has the potential to be applied to the vitamin D_3 metabolites detected in matrices such as serum or urine, which are routinely found in low concentrations and often masked by other endogenous material. Hence, chromatographic separation prior to MS/MS analysis may be beneficial whilst the confirmatory characteristic fragments highlighted in chapter 4 can be used to identify and quantify the biologically active 1,25-dihydroxyvitamin D_3 compound.

Chapter 5: Development of matrix-assisted laser desorption ionisation mass spectrometry for the detection of SARS-CoV-2 proteins

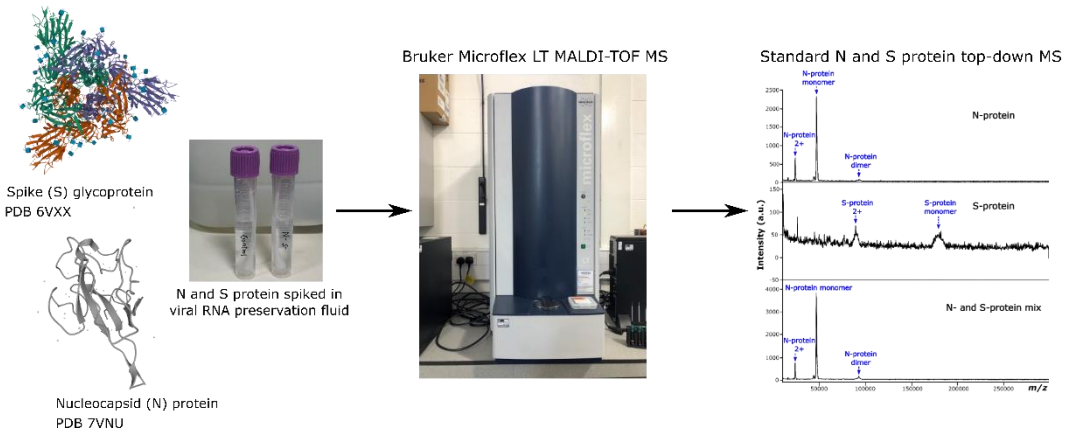


Figure 6. 4 The summary figure for chapter 5, highlights examples of standard SARS-CoV-2 biomarker proteins and samples prepared for the viral enrichment optimisation experiments, subjected to MALDI-TOF MS and the resulting top-down mass spectra obtained for data analysis.

In chapter 5, different methods for protein enrichment and extraction were investigated, which were subsequently applied to the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) standard proteins, including the SARS-CoV-2 positive and negative patient swab samples. Analytical results were obtained on a Bruker MALDI-TOF MS platform as well as the 12 T Bruker SolariX FT-ICR MS, where possible.

The establishment and use of linear MALDI-TOF MS for the analysis of clinical microorganisms has rapidly evolved over the years. The available commercial instruments have demonstrated the capability of providing high speed detection and identification of large molecular weight proteins, bacteria, and other microbial species. As a result of global spread and significant impact of the SARS-CoV-2 virus, there is an urgent need for the development of rapid screening methods.

SARS-CoV-2 has four main structural proteins, which include the spike (S) glycoprotein, envelope (E) glycoprotein, membrane (M) glycoprotein, and the nucleocapsid (N) protein. The N and S protein is of special interest, as the N protein is essential for viral genome packaging and the S protein is responsible for mediating attachment and cellular entry of the virus. Thus, in this work, focus is placed on the

optimisation of the detection of the N and S proteins in the standard proteins and in the SARS-CoV-2 negative and positive patient samples.

The different approaches explored in chapter 5 identified sample preparation methods such as high molecular weight cut-off filters and the use of detergents often used in cell lysis for enrichment and extraction of proteins present in the patient samples. The various sample preparation techniques were firstly applied to the standard viral SARS-COV-2 proteins, which demonstrated enrichment of the proteins and improved intensity of the detected species. However, due to the complex nature of the patient samples, the same enrichment and extraction approaches did not result in detection of the N and S viral biomarker proteins in the SARS-COV-2 positive patient swab samples.

Overall, the experimental results demonstrate that it is crucial to optimise sample preparation techniques for the detection and enrichment of SARS-CoV-2 proteins in the patient samples prior to MALDI-TOF analysis. This is supported by the information obtained from the experiments conducted, which resulted in a total of 107 variables tested on the standard COVID-19 proteins and the patient swab samples to improve the detection of the proteins of interest. In future, the sample preparation methods discussed herein for viral protein enrichment and extraction can be expanded to include different assays, such as affinity capture beads, which target the critical biomarkers of SARS-CoV-2. Thus, further studies are required to explore the full potential of MALDI-TOF MS as a clinically useful tool for the screening of SARS-CoV-2.