New Tandem Mass Spectrometric Methods of Structure and Sequence Elucidation of Synthetic and Tryptic Peptides

Emilie Viglino

A dissertation submitted in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy

University of Washington

2017

Reading Committee:

František Tureček, Chair

Matthew Bush

Robert E. Synovec

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

New Tandem Mass Spectrometric Methods of
Structure and Sequence Elucidation of Synthetic and Tryptic Peptides

Emilie Aude Viglino

Chair of the Supervisory Committee:
Professor František Tureček
Department of Chemistry

Over the last few decades, technological advances in the physical and life sciences have enabled scientists to look at various systems ever so intricately and to gather evidence for the mechanisms behind these systems, be they at the biological level or the astrophysical level. One of the many tools that has spurred the collection of large amounts of data for a better understanding of the World around us is mass spectrometry. Critical advancements in the engineering of mass spectrometers since their inception over a century ago have secured mass spectrometry' position as a robust technology in the field of Chemistry. The versatility of this tool is illustrated by its application to a variety of scientific disciplines and by its predominance in the scientific literature. Currently, mass spectrometry is utilized routinely in the clinical laboratory as a means to evaluate the potential for certain biomarkers to be disease-causing. Such biomarkers typically are in protein or peptide form, hence the need to develop tools aimed at probing the sequences and structures of these compounds. Gaining insight into the building blocks of proteins and peptides and their overall organization is crucial for determining their functions and biological impact. The body of work presented herein focuses on the development of two separate tools for the study of peptides and proteins in the gas phase; the first tool was aimed at probing the structure of tyrosine-containing peptide cation radicals, which are particularly important in the functioning of redox enzymes, while the second tool was aimed at the synthesis of a unique iodine-containing charge tag in view of increasing the sequence coverage of digested proteins for easier and more reliable identification.

In the first instance, a comprehensive study of collision-induced dissociation (CID) and near-UV photodissociation (UVPD) of a series of tyrosine-containing peptide cation radicals of the hydrogen-deficient and hydrogen-rich types is reported. Hydrogen-rich and hydrogen-deficient peptide cation radicals vary in electronic states and reactivity. Hydrogen-rich cation radicals $[M + nH]^{(n-1)+\bullet}$ are produced when an electron attaches to a multiply charged peptide ion, such as in electron transfer dissociation (ETD). Their backbone fragmentation is dominated by N-Cα cleavage and yields valuable sequencing information. Hydrogen-deficient cation radicals $[M + (n-1)H]^{n+\bullet}$ are produced by an atom abstraction from a protonated peptide or via electron loss from a neutral peptide, as in electron ionization. Formation of the latter highly depends on the presence of both a basic amino acid and a tyrosine or tryptophan residue in the peptide and their fragmentation is dominated by C-Cα backbone cleavage. The former are far less studied and their mechanisms less understood. The generation and photodissociation of both hydrogen-rich and hydrogen-deficient cation radicals in tyrosine-containing pentapeptides was investigated in parallel studies. Using tandem analyses with an LTQ-XL ion-trap equipped with an ETD source and a λ=355nm Nd:YAG laser, crown-ether-peptide and Cu(II)(Terpy)-peptide complexes were formed via electrospray to generate $[M+2H]^{+\bullet}$ and $[M]^{+\bullet}$ cation radicals, respectively. Their subsequent photodissociation was achieved using a range of pulses, each conferring 18 mJ of energy to the cation radicals. The peptides investigated were AAAYR, AAYAR, AYAAR and YAAAR, with special interest in the effect of the tyrosine residue position. A validity experiment was conducted in order to mass-label the fragments obtained, where one alanine residue was substituted with one valine residue. All UVPD spectra obtained were compared to non-zero collision energy CID spectra.

Additionally, the body of work described here reports the first application of UV/Vis photo-dissociation action spectroscopy for the structure elucidation of tyrosine peptide cation radicals produced by oxidative intra-molecular electron transfer in gas-phase metal complexes. Action spectroscopy exploits an optical parametric oscillator (OPO) for the tuning of a given laser frequency. In doing so, the dissociation channels for particular precursor ion are monitored as a function of wavelength, from λ = 200-700 nm in this case. For any intermediate species that may not be studied in solution, the action spectrum may be compared and correlated to putative ion structures obtained from density functional theory (DFT) calculations combined with time-dependent DFT (TD-DFT) calculations of electronic excitation energies and oscillator strengths in the cation radicals. The study showed that the oxidation of YAAAR produces Tyr-O radicals by

combined electron and proton transfer involving the phenol and carboxyl groups. Oxidation of AAAYR produces a mixture of cation radicals involving electron abstraction from the tyrosine phenol ring and N-terminal amino group in combination with hydrogen-atom transfer from the Cα positions of the peptide backbone.

In a second instance, a unique method of C-terminal lysine guanidination coupled to a diiodinated tyrosine insertion allowed the synthesis of a novel mass defect-containing charge tag. The potential for this diiodinated charge tag to enhance the electron transfer dissociation (ETD) efficiency of peptides and to thus increase their sequence coverage was evaluated for both synthetic and tryptic peptides. Iodine is a large atom with a large mass defect; this simply means that the nominal mass of iodine differs from its exact mass more so than is the case for the majority of the other atoms. The addition of iodine's mass-defect is aimed at preventing the loss of information from low-mass cutoff typically seen in ion traps and also at shifting both precursor ions and their fragments to a spectral area with no native peptide fragments. Peptide fragmentation can generate several ion series, which results in complex spectra. Simplifying these spectra is one approach to enhancing peptide sequence coverage. Sequence identification commonly relies on the use of databases for spectral comparisons to known proteins. However, this approach suffers many shortcomings and has catalyzed the need for *de novo* sequencing, made possible by peptide tagging. Our mass defect labeling approach consisted in building the guanidine tag on a solid phase substrate bound to resin beads confined to a fritted syringe. Excess of thiocyanate-modified tyramine was then reacted with the substrate to form a thiourea analog before being oxidized and diiodinated to produce the precursor tag. In a last step, the tag is coupled with lysine-terminated peptides (synthetic or from digest) before being cleaved off the beads and analyzed. The mass spectral analyses of the tagged peptides products and their unmodified counterparts were conducted on at high-resolution. After optimization of the iodination process (3,5-diiodinated product yield >95%), lysine-terminated peptides were tagged. The diiodinated chargeable tag was successfully attached at the ε-amine of the lysine side chain of a mixture of synthetic lysine-terminated peptides (AA**X**AK, where **X** = D, F, H, K, N,). Overall, and across all five peptides, the ETD fragmentation efficiency was enhanced by the availability of multiple charges provided by the guanidine group and the complete coverage of the z-ion series, starting with $\mathbf{z_1}$, was made possible by addition of the iodinated charge tag. This method was applied to the peptides from bovine albumin (BSA) digests using Lys-C and Trypsin proteases in two separate experiments and showed equal success.

Acknowledgements

First and foremost, I would like to acknowledge my research advisor, Frank, without whom none of this work would have been possible. Your unconditional support and your mentorship excellence have allowed me to better myself as a scientist and as a person. The level of faith that you put in me when I was most astray in Graduate School led me to rebuild my confidence, to become passionate about mass spectrometry and to let my abilities shine through. I feel forever indebted to you and I could never adequately express how grateful that I am to have worked under your wing. Thank you for your absolute enthusiasm in science and in the natural World, and for sharing your hiking and climbing experiences with us. I am convinced that the quality of your advising finds no equal, be it at UW or at any other University. From the bottom of my heart, thank you.

Secondly, I would like to credit my Undergraduate advisor, Dr. Sigman, for his encouragements and for convincing me that Graduate School was a path I should consider. I remember your words: "this will be the hardest you ever do and also the most rewarding thing you ever do", and you were right on both counts! Thanks to your recommendation I was able to not only get my degree in Chemistry, but to also live in Seattle, where I always dreamt to move. It pinched my heart a little to be back on UCF grounds for the Alma Mater seminar and you should know that I won't forget where my roots are.

Many people at the University of Washington have helped me see this endeavor through and a few of them deserve special recognition. Priska Von Haller, at the Proteomics Resource Center has amply contributed to the advancement of my knowledge and skills in LC-MS. Thank you for taking the time to train me, thank you for your patience and thank you for all the many life conversations that we have had during idle times. You are an inspiration to me and I strive to become as good as what I do as you are. Jimmy Eng, at the Proteomics Resource Center has made it possible for me to fully exploit the sequencing tool I developed by helping me sort my data and by writing a script that has been invaluable in validating the strength of that tool. Thank you for all of your help and for your contagious smile. Martin Sadilek, at the University of Washington, has been a wonderful resource to me and to many more graduate students over the years. Thank you Martin for all the training on the instruments and thank you for organizing the Pacific Northwest Mass Spectrometry Group meetings and for pushing us to go and mingle with like-minded people. Additionally, thank you for fronting

the Green Chemistry bicycle team and for rallying people to join the Ride in the Rain and the Bike to Work Month year after year. Wing has worked on the *de novo* analysis of peptide undetected by the database search and I am infinitely grateful for his help when my plate was becoming too full to handle.

Finally, Lonnie Gates at Thermo deserves a shout out: you rock! You have followed me through my journey as a young mass spectrometrist and the encouragements that you have showered me with cannot be underrated. I am enlivened by all the conversations we have about science and life in general and I am so glad that you will not be retiring anytime soon. Thanks for being the best service engineer we could hope for.

Last but not least, I would like to acknowledge the National Science Foundation for supporting the projects that I worked on, and which allowed me to fully immerse myself in my scientific investigations.

Dedication

I would like to dedicate this dissertation to my grandmother and my parents. Unlike myself, my grandmother did not have the luxury to go to school and get educated, yet she placed the utmost value in Education. She made sure that her children, including my father, would get educated and make a better life for themselves. I miss you dearly and I carry you in my heart always. This is for you and I know you would be proud of me. My parents made the ultimate sacrifice when, at the age of seventeen, I asked them if they would let me go to the United States to better my English and discover a new World. Seventeen years later, I am still here and the completion of this work means at least as much to them as it does to me. Their unconditional love, support and trust in me have always helped me brave the obstacle that life has thrown at me. I feel strong because I have you and you deserve just as much credit for this work as I do, truly. My sister, Marlene, aka "Gertrude", probably could not agree with me more. Gertrude, je t'aime de tout mon coeur. The bond that unites us is indestructible and every single day I miss you more than I am able to say, but I am looking forward to the next time we'll laugh together. Alexandra and Anais, what can I say? How could I possibly imagine my life without you in it? You are the best nieces an aunt could wish for and I am so grateful that you are always here for me when I need you most. I hope you know you can count on me too. I love you. This dissertation also goes out to my grandparents, and specifically to the exquisite beauty of their bond, evidenced by seventy-three years of marriage. During the best of times and the worst of times you have been there for each other. Now well into your mid-nineties, you are still there for each other and I can only hope that, one day, I too find someone that I can hold onto and who can hold onto me the way you do each other. Marti and Steve, you too certainly have this bond and it is beautiful to see your love for each other abound. You are a second family to me and you have supported me immensely over the years and I can only try to repay you in the very small ways that are available to me. Thank you for including me in your family and for making me a part of it without question. I love you guys.

Lastly, I have to dedicate this work to all the friends that have been in my life and who have kept me sane during the past eight years and beyond, starting with my co-workers, whose friendship has already outlived Graduate School. Andy, seriously, how did I get so lucky to work with you? I respect you and your morals, your mind and your heart. You are a wonderful person and I feel really lucky that I got to be your office mate

and beer buddy for this long. Let's keep it going until you return to California, shall we? Yang, what a beautiful soul you have; I cannot thank you enough for all the conversations and the laughs we have exchanged around the lab. Recently, you have been here for me in ways that I did not expect and I hope that I can preserve this interaction with you for as long as there will be classical music to be heard live! Thank you. Alisa, my life as I know it now would be completely upside down if not for this chance encounter. You have been a blessing to me and I will never quite be able to show my gratitude for everything that you have done for me, from helping me get my dream job to being the most fantastic and reliable co-worker, and for listening and being there. Mariana, "Dude, it was my dream to be here" and that dream is now a reality but we can still laugh out loud when I say it. Amy, I met you at such a strange time in our lives and you're still here, enjoying my company as much as I do yours. Since I have met you I have told you how lucky I feel for what we share and you know how true it is now. What on Earth would have done without you in the past few weeks? I seriously do not know. Thank you for giving me strength, thank you for keeping me lucid and mentally healthy, and thank you for believing in me when I doubted myself most. Sergio, all these nights studying together have come to this, can you believe it? You told me yourself that I was selling my life away when I chose to enter the Graduate Program in Chemistry at UW… and how true was that?! Ah, well, now I can put it in the rear-view mirror and finally take time to visit you and celebrate. Mamita Sonia, you are a strong woman who defeated cancer and who still managed to smile throughout the most trying of times. You are the epitome of poise and wisdom and I thank you for sharing it with me over the years. Lastly, all of you guys born in Montceau-Les-Mines or thereabout, I am so lucky to have you as my friends. I have known most of you between twenty or thirty years and our friendship still stands… so thank you, particularly Delphine, Severine, Jeanne, Jerome and Ben, you keep me grounded and always remind me who I am.

One last dedication goes out to the Pacific Northwest. Your majestic volcanoes and magnificent, breath-taking beauty has made me fall in love with you completely and over the years you have conquered my heart completely. I can only hope to get to know you better over the time that I have left in this area and you will always have a special place in both my heart and memory.

All in all, with all this love in my life, I am a pretty lucky gal. I could not have done it without you. Thank you.

Table of Contents

List of Abbreviations

A	Alanine
ACN	Acetonitrile
BSA	Bovine Serum Albumin
C	Cysteine
Cu^{II}(tpy)	Copper(II) Terpyridine
Cu(tpy)	Copper(II) Terpyridine
CI	Chemical Ionization
CID	Collision-Induced Dissociation
D	Aspartic Acid
DMF	*N*,*N*-dimethylformamide
E	Glutamic Acid
ESI	Electrospray Ionization
ETD	Electron Transfer Dissociation
F	Phenylalanine
FA	Formic Acid
FAB	Fast Atom Bombardment
G	Glycine
H	Histidine
I	Iodine
IAA	Iodoacetamide
K	Lysine
L	Leucine
LC	Liquid Chromatography
[M]	Peptide
MALDI	Matrix-Assisted Laser Desorption Ionization
MS	Mass Spectrometry
N	Asparagine
NaOAc	Sodium Acetate
$Na_2S_2O_3$	Sodium Thiosulfate
NCE	Normalized Collision Energy

PEG	Polyethylene Glycol
Q	Glutamine
S	Serine
SI	Supplemental Information
T	Threonine
TCEP	Tris(2-carboxyethyl) phospine
UVPD	Ultra-Violet Photodissociation
V	Valine
Y	Tyrosine

List of Schemes

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Mass Spectrometry

1.1.1 The Development of Ionization Sources

Over the past century, mass spectrometry (MS) has established itself as one of the most powerful analytical tools for the study of compounds from diatomics to macromolecules. The versatility of mass spectrometry finds its application in the majority of scientific fields, from forensic science[1] to environmental sciences[2–4] and the clinical laboratory,[5] amongst others.[6,7] As technology advanced ever so rapidly following the industrial revolution, mass spectrometry was able to adapt to the growing need for higher-performance instruments. Here is a brief summary of how this technique evolved from a means to study atomic structures to an intricate instrument capable of at analyzing the human proteome.

What exactly is a mass spectrometer? At the most basic level, a mass spectrometer is an instrument that measures the mass-to-charge ratio (*m/z*) of ions. This, in turn, allows for the determination of the molecular mass of that ion. Regardless of how obsolete or modern a mass spectrometer is, three major steps make up the process by which this mass determination is made. Ionization, by which a molecule under study is converted into gas-phase ions is the first, and also the most challenging, step. Separation of the gas-phase ions previously generated ensues, and can be achieved via magnetic or electric fields in a part of the mass spectrometer known as the mass analyzer. Detection of the separated ions of a given *m/z* and their associated abundance follows and completes the general MS process.[8]

Historically, the advent of mass spectrometry is intricately associated with the discovery of the electron by physicist J.J. Thomson in 1899.[9] Thomson developed the first mass spectrograph using a cathode ray tube, as a means to measure the atomic weight of elements.[10] Additionally, the first clue as to the existence of stable isotopes can be attributed to this early approach of mass study by Thomson's students, F.W. Aston and A.J. Dempster, when they proved the existence of the stable isotope of neon, ^{22}Ne.[11] The domino effect that this breakthrough had on the fate of mass spectrometry cannot be underrated. The finding led Alfred Nier to verify that the fissionable isotope of uranium was ^{235}U, and to discover the ^{13}C isotope, which precipitated the ubiquitous use of isotope tracers in both chemical and biological studies, and

subsequently steered the use of mass spectrometry as the detector of these markers.[9,12] After the start of the Second World War, efforts in advancing the field of mass spectrometry stagnated for a time and the instrument was primarily used for the detection of ^{235}U leaks in enrichment chambers, as well as for petroleum purity control as the need for and dependence on fossil fuel grew immensely. Nonetheless, the use of the mass spectrometer for the diagnosis of petroleum quality generated an interest in understanding the processes behind molecular ionization, in order to develop a systematic method that would yield reproducible and diagnostic spectra for fuel quality screening. These efforts, in turn, led a handful of scientists, notably Beynon, Budzikiewicz, Djerassi, and McLafferty,[13–17] to focus on the development of mass spectrometers for the mass measurement and structure assignment of organic, which, until that point, were only crudely made.[18,19] Additionally, Klaus Biemann, then newly appointed chemist at MIT, was particularly interested in the characterization of natural products by mass spectrometry, and he used electron Impact ionization (EI) towards that end.

The technology, developed during the war, used a beam of highly energetic electrons to bombard volatilized molecules; the underlying mechanisms by which EI fragmented these molecules into smaller ions, however, were poorly understood. Biemann was a visionary and he believed that understanding the mechanisms that imparted EI fragmentation would make the *de novo* determination of a molecule's structure possible.[20–22] Towards that goal, he developed a spectral library and a series of algorithms assisting in the interpretation of mass spectra for *de novo* sequencing. To date, no fully reliable system of automated spectral interpretation for *de novo* has been developed. A major advantage of EI was that it allowed the ionization of molecules with inherently low volatility. On the other hand, the fact that EI heavily relied on the derivatization of molecules to achieve their volatilization and the fact that it triggered substantial fragmentation of the molecules subjected to it made the method less than ideal and made the analysis of polar, charged compounds like biomolecules, practically impossible. These undesirable aspects would not be entirely addressed prior to the development of soft ionization techniques, and until then, mass spectrometry heavily depended on derivatization methods for the analysis of compounds with low volatility.

Chemical ionization (CI), a method that used gas-phase ion-molecule reactions between a reagent like CH_4^+ and analytes emerged[23,24] and led way to peptide sequencing.[25] The technique supplied low energy to the molecules as it involved the transfer of only one proton, thereby inducing only little fragmentation in

the newly formed protonated ions and leaving enough intact molecular ion for molecular mass determination; it was the first of the *soft ionization* methods to be developed.[26] Contemporary to this technological advancement, came about the interfacing between gas-chromatography and mass-spectrometry (GC-MS).[27] With their powers combined the two technologies offered the ability to separate the compounds of a mixture and to determine their molecular weights in a single analysis. This advancement was revolutionary as it allowed the mass spectrometric analysis of impure compounds for the first time. This online GC-MS approach was extensively applied by Biemann to the detection and *de novo* sequencing of peptides derivatized with polyamino alcohols and the comparison of these sequences with DNA-derived ones.[28,29] Despite the major landmark of this sequencing method, it was still limited by tedious derivatization reactions, by its sub-par sensitivity and by its inability to volatilize larger peptides. Research in the 1970's produced the automation of the method and the development of a new kind of mass analyzer, the triple quadrupole, which would give birth to tandem mass spectrometry (MS/MS).[30]

The triple quadrupole, developed by Yost and Enke, laid the foundation for a systematic sequencing procedure based on tandem mass spectrometry (MS^n), **Scheme 1.1**. The way it works makes use of a first quadrupole mass analyzer, which allows the selection of a particular ion, which is then guided to a quadrupole collision chamber, where the ion is fragmented upon collision with inert gas molecules for collision-induced dissociation (CID) before entering a second mass analyzer, where the product ions are further mass-analyzed, and finally detected by an electron multiplier.[31] The practical value of the triple quadrupole rests on the structural information that the fragment ions of a MS/MS sequence provides for a variety of compounds.[30–35] While the technique allowed both the combination of a soft ionization and the more controlled and extensive fragmentation of peptides than before the existence of the triple quadrupole, it did not overcome the need for extensive derivatization of the peptides prior to ionization.

The continuous efforts to design new ionization methods, soft ionizations particularly, had fronted the advancement of mass spectrometry for a couple of decades when parallel research brought both fast atom bombardment (FAB) and electrospray ionization (ESI) to light.[36,37] FAB, which used a fast beam of atoms (~8000 eV) to ionize samples from a matrix, could volatilize samples softly and without prior derivatization; this made the technique very attractive for protein and peptide sequencing.[38] The peptides studied by this method could be as large as thousands of Daltons, which was a considerable improvement from past

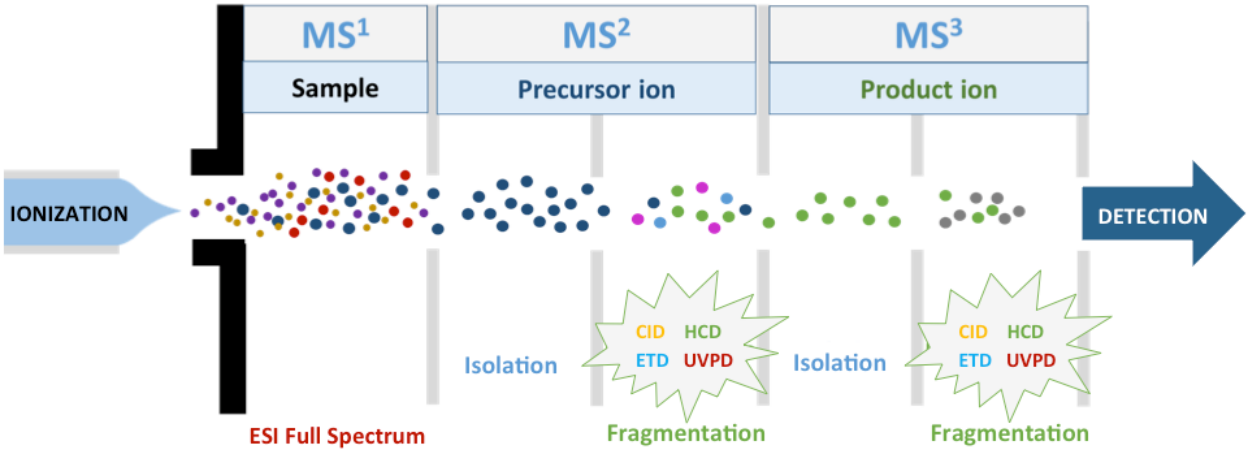
techniques, but yet could not accommodate the larger biomolecules. Due to the little fragmentation that resulted from using FAB, the method was routinely used in combination with tandem mass spectrometry for sequencing and analysis of peptide mixtures.

In the same vein as FAB, another soft ionization method known as matrix-assisted laser desorption ionization (MALDI),[39] developed independently by Tanaka[40] and Karas and Hillenkamp,[41] proved to make the study of large biomolecules such as carbohydrates, proteins and DNA possible.[42] Much like FAB, the method utilizes a beam of photons that interacts with a sample held on a stage.[43] Typically, MALDI was interfaced to time-of-flight analyzers (TOF), which measure the time a compound takes to travel a determined distance.[44] The reason was that no other analyzer type could compete with their capacity to analyze very large mass ranges.

Introduced contemporarily to MALDI, in the 1980's, a newly designed ionization method, known as electrospray ionization (ESI), would soon appear in the field of mass spectrometry and revolutionize the discipline. Illustrating how truly remarkable the technique was, its inventor, John Fenn –who went on to win the 2002 Nobel prize in Chemistry for his invention – said: ESI is "wings for molecular elephants".[45] The image does not fail to describe the incredible contribution that the method made to the ionization of very large molecules. Indeed, by generating multiply charged species, ESI guaranteed that the detectors commonly used at the time it was developed would accommodate the *m/z* range of large molecules. In ESI, samples flow through a capillary, to which a high voltage is applied. At the tip of the capillary is formed a plume of droplets, which desolvate as they move towards the mass spectrometer source.[46] Particularly attractive was the possibility of interfacing a liquid chromatography (LC) separation system to ESI-MS.[47–49] Indeed, the online interfacing of a LC system with a tandem mass spectrometer meant that compounds could easily be separated and analyzed online in a single run and thus, the method represented an incredible increase in time efficiency for sample analysis. With the development of ESI, the field of proteomics took off and LC-MS became a ubiquitous tool in the physical and life sciences for the study of biological samples.[50–53]

Together, MALDI and ESI have permitted the emancipation of mass spectrometry and took it to the masses by providing biologists and chemists alike with an invaluable tool for the study of their biological samples. To this day, these are the two most common methods used for the ionization of biomolecules.

Electrospray ionization (ESI) tandem mass spectrometry is the method that was used throughout the experiments described herein, **Scheme 1.1**.



Scheme 1.1. Overall process of ion isolation and ion fragmentation during the MS^n process in a linear ion trap. Note that on triple quadrupole mass analyzers, isolation and fragmentation of a selected ion can only be done at the MS^2 level.

1.1.2 Tandem Mass Spectrometry: Ion activation methods

In order to gain insight as to the structure of a molecular ion using tandem mass spectrometry, the isolated ion needs to be fragmented in the mass analyzer. For this fragmentation to occur, is required a change in the internal energy of the molecular ion, which can be supplied by any of the several activation methods available. The well-established method of ion activation known as collisional activation (CAD), or collision-induced dissociation (CID), relies on the collision of the molecular ions with inert gas molecules.[54] Known as a "slow-heating" activation process (5×10^{-5} s),[17] the energy causing the fragmentation is more quickly redistributed along the ion and produces mass spectra rich in fragment ions characteristic to that molecular ion, specifically b and y ions in peptides and proteins.[31] Adopted nomenclatures for peptide ions are defined in **section 1.2.1**. Despite some limitations, collision-induced dissociation, which was the first activation method to be developed, is still the most widely used today. The efficiency of CID depends on several parameters, including the amount of kinetic energy supplied to the ion, the nature of the ion itself and the number of collisions allowed, amongst other criteria.[55] This activation method was used in the

experiments described in **Chapters 2-3** as grounds for comparison with two other activation methods (ETD and UVPD) for the generation of peptide cation radicals in the gas phase. These other methods, with more specific applications, have been developed over the years and are briefly described as follows.

Multiply charged ion species can be fragmented under specific mechanisms with methods like electron capture dissociation (ECD) or electron transfer dissociation (ETD). These methods, together referred to as ExD, involve the attachment of an electron to a peptide or protein.[56] An electron source is necessary for the electron capture or transfer to occur; fluoranthene is one of the commonly used radical anion donor.[57] These two methods generate the complementary c and z ion series, and each ion can be observed in a spectrum when it carries the charge with it. It has been well-documented that the ETD efficiency is increased as the number of charges is increased, thereby resulting in greater fragmentation and hence, in more sequencing information for a given peptide.[58] One major advantage of ETD rests on the fact that it does not affect post-translational modifications, tagging or labeling of a protein, unlike CID.[59-65] Electron transfer dissociation was used in the work described in **Chapter 2** for the generation of hydrogen-rich peptide cation radicals and in **Chapters 4-5** for the sequence coverage analysis using a novel iodinated charge-tag.

Photodissociation (PD) is a method that utilizes a laser for the dissociation of gas phase ions; both infrared (IR) and ultraviolet (UV) lasers have been widely used for this purpose. Infrared Multiphoton photodissociation (IRMPD) involves the stepwise vibrational excitation of a gas-phase ion by absorption of several photons to trigger dissociation, and as such, is classified as a slow heating process.[66] The fragmentation patterns observed in IRMPD spectra show similarities to those obtained via CID and the mechanisms behind these two processes have been investigated.[67, 68] The nature of the IRMPD dissociation process has been covered extensively reviewed.[69,70] Ultraviolet photodissociation (UVPD)utilizes higher energy lasers and targets chromophores –light-absorbing groups such as aromatics, or radicals newly formed via ETD or CID.[71-74] In the case of peptide and protein studies, the chromophore may be inserted either covalently or non-covalently[75] for absorption in the near-UV wavelength range. The advantages of the PD approach are multiple; it is possible to couple this activation type with both time-of-flight (TOF) and ion trap instruments, and the energy that is deposited to the peptide or protein under study can be modulated. The coupling of an optical parametric oscillator (OPO) to a laser pump allows the frequency tuning of the laser for absorption at a wide range of wavelengths[76-79] and for action spectroscopy studies.[80-82] Action spectroscopy

consists in the monitoring of all the dissociation channels resulting from the exposure of a precursor ion to laser pulses as a function of wavelength. The interpretation of the spectra obtained by action spectroscopy is typically assisted by time-dependent density functional theory (TD-DFT) calculations, and the sound candidate structures for the dissociation fragments are proposed. This method was used in the work described in **Chapters 2-3**.

While only these three aforementioned ion-activation methods were utilized to obtain data in the work described in the following chapters, it is worth briefly mentioning a few other commonly used activation methods. Infrared multiphoton dissociation (IRMPD), a photon-driven activation method, yields spectra with an abundance of fragment ions, but is a non-selective method.[83] Typically used with quadrupole ion traps and Fourier transform ion cyclotron resonance instruments, this method can be employed with both high-power pulsed laser or with continuous wave (slow heating) power.[17] Blackbody radiation, another slow heating method, has been used to study the folding and unfolding of peptide ions in the gas phase[84,85] and to determine both the activation energies and various critical parameters for the characterization of fragmentation pathways.[83] Lastly, higher collision dissociation (HCD), is found in hybrid instruments, such as the Thermo LTQ-Orbitrap Fusion mass spectrometer, and operates the ion dissociation in a C-trap with radial ion ejection from a curved RF-only quadrupole ion trap.[86] HCD is comparable to CID operated on a triple-quadrupole or time-of-flight instrument (TOF); it can be used alone for the identification and quantification of proteins using reporter ions from labeled peptides[87] or it can be used in conjunction with CID, with no effect on the duty cycle, in order to achieve a higher number of protein identifications and a greater number of quantifiable proteins with better accuracy.[88,89]

As previously mentioned, CID and ETD were used as ways to generate tyrosine-containing peptide cation radicals, of both hydrogen rich and deficient types, and UVPD was utilized as a probe to the structure of these peptide cation radicals, introduced below.

1.2 Peptide Cation Radicals

1.2.1 Peptides

Natural peptides are composed of any combination of the twenty naturally occurring amino acids joined together at their C- and N-termini via an amide bond. The study of natural peptides often requires the

cleavage of a protein into smaller units, which is achieved enzymatically.[90] For greater ease of sequencing of these peptides, a variety of modification methods such as phosphorylation,[91] guanidination,[60,92-94] methylation[59] or yet ubiquitination[95] are available. Oftentimes, modifications are performed on the side chain of lysine (Lys, K) residues,[96] but they also commonly target either terminus of a peptide.[91,97] The major common advantage to the many peptide modifications available is a reduction in spectral complexity, be it for its steering of the peptide fragmentation towards a single ion series[98] or for its shifting of the masses of natural peptides into an unoccupied space of the spectrum,[99,100] as observed for mass defect labeling, described thoroughly in **Chapters 4-5**. Non-naturally occurring peptides are made synthetically and can be modified in identical ways as for natural peptides. Additionally, some peptide modifications, such as photo-labels using diazarine,[82,101,102] have been devised for purposes other than peptide sequencing and may help reveal structural information about the peptide. The structural information of a peptide can also be elucidated using tandem mass spectrometry (MS/MS) by analysis of the fragment ions composition resulting from activation of a precursor ion.

Peptide fragment ions are conventionally named according to the types of broken backbone bonds (C-C, C-N, or N-C_α) from which they result. Those are labeled a_n, b_n , or c_n ions if the ions contain the N-terminus of a peptide, and their complementary x_n, y_n, or z_n ions are named for C-terminus-containing fragments. A subscript number is assigned to each identified ion in a spectrum, which indicates its location on the peptide backbone. The convention is illustrated in **Figure 1.1**. It is worth noting that different ion activation methods will generate different types of fragment ions. For example, the slow-heating collision-induced dissociation (CID) mostly generates b and y types ions, while electron transfer dissociation (ETD) generates almost exclusively c and z ions. The aforementioned nomenclature was devised[20,103] for ease of spectrum peak labeling, but with no intention of proposing a given structure for these ions.[104] This nomenclature is used in **Chapters 4-5**.

Furthermore, a recently introduced nomenclature,[105] which builds upon the Biemann naming convention, is also used[74] and aims at uniformly describing fragment ions from various cation radical sources. This new and comprehensive naming method differs from Biemann's in that it assigns charge, radical and the number of hydrogen atoms comprising the given fragment ions. This nomenclature is used and detailed in **Chapter 2** for the study of tyrosine-containing peptide cation radicals.

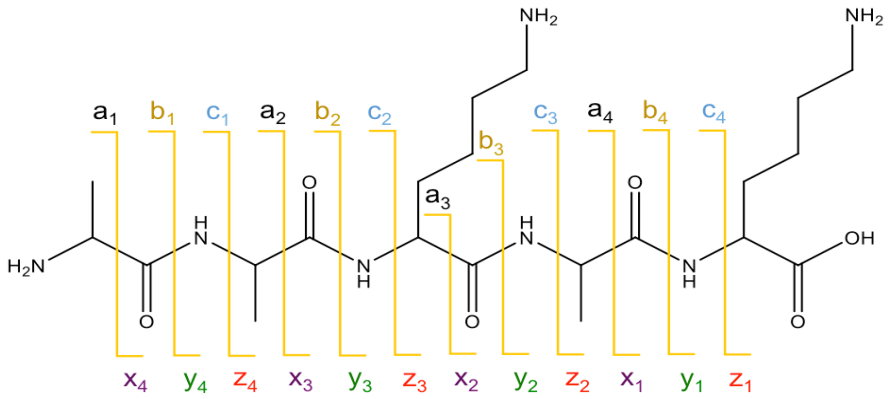


Figure 1.1. Illustration of peptide fragment ion nomenclature for pentapeptide AAKAK (Ala-Ala-Lys-Ala-Lys).

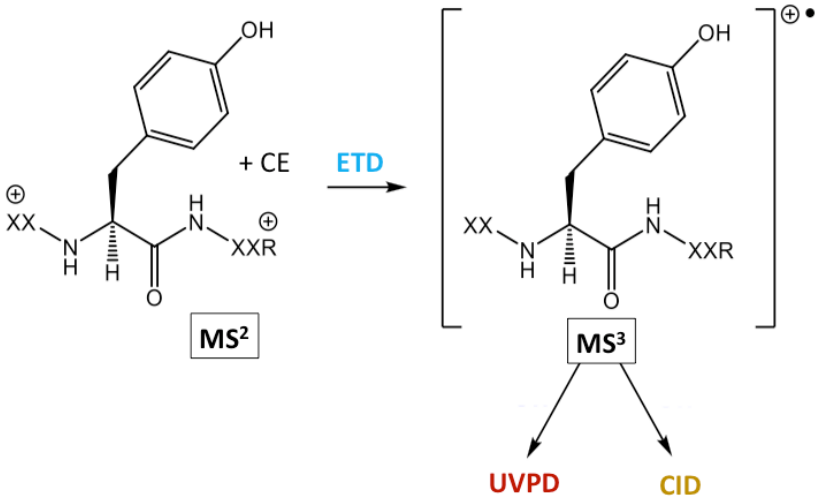
1.2.2 Cation Radicals

With the advent of electron impact, which abundantly generated cation radicals upon ionization of organic molecules, the field of mass spectrometry turned its attention to studying these intermediate species.[104] The tools available to them allowed mass spectrometrists to investigate the nature of these intermediate radicals and to simultaneously study the underlying mechanisms that cause fragmentation.[106] Many biological processes of research interest are driven by a number of intermediary steps that involve reactive radicals. Examples of such important transient and reactive intermediates are the aromatic cation radicals and the tyrosyl O-radical, which can result from the one-electron oxidation of the tyrosine residue.[74]

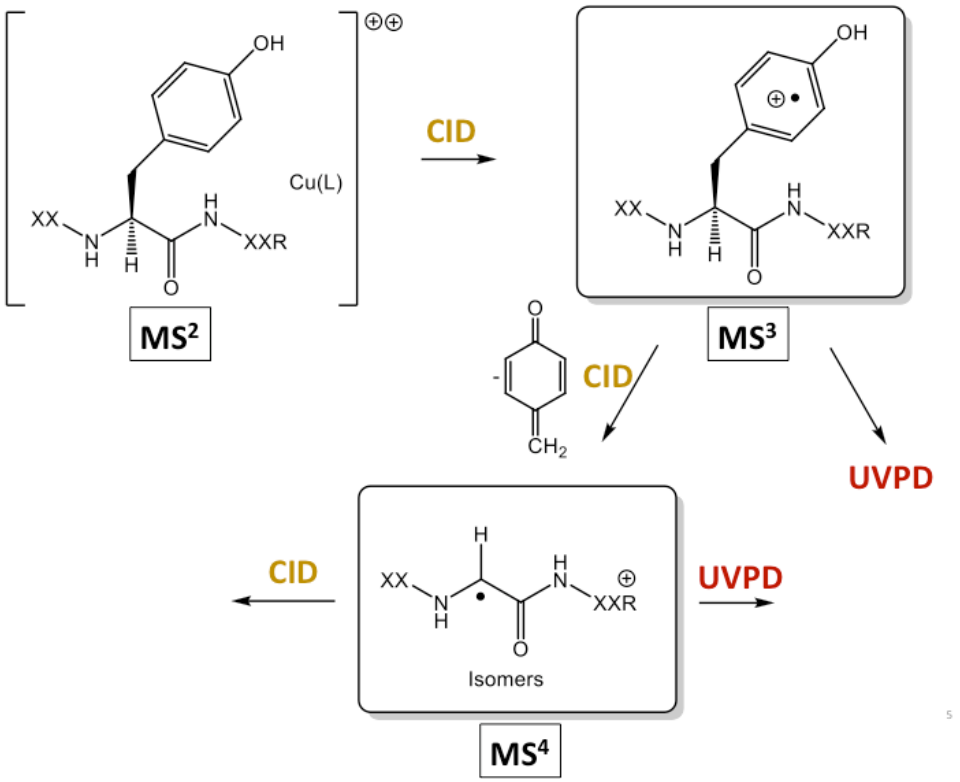
There are two main types of peptide cation radicals that can be generated in mass spectrometry, namely hydrogen-rich peptide cation radicals and hydrogen-deficient cation radicals; each of these two types of odd-electron ions displays very different fragmentation patterns and supplies complementary structural information regarding the peptide. Characteristic of the peptide cation radicals, relative to their even-electron counterparts, is the uncertainty of the location of the backbone and side-chain radicals, shown to migrate by hydrogen transfer of radical isomers.[82] Investigations on the location of the radical center was at the core of the projects presented herein (**Chapters 2-3**).[107] There are several ways in which each type of cation radical can be formed, as discussed below.

Hydrogen-rich peptide cation radicals, $[M+2H]^{+\bullet}$, also known as charge-reduced cation radicals, can be formed by the transfer of an electron to a multiply charged peptide via ETD.[108] While the direct transfer of an electron to a doubly-charged peptide may only result in its backbone fragmentation, the peptide can first

be mixed with a 18-crown-6-ether solution (CE) to form a non-covalent complex that is isolated and broken apart by the incoming electron, leaving behind the peptide cation radical to be studied, **Scheme 1.2.**[74] Once formed, the cation radical can be probed by various methods, including CID, ion-molecule reactions[109] and UVPD at λ=355 nm,[73] which selectively targets the radical chromophore, while the rest of the peptide remains transparent to the laser light.[102] Upon dissociation, hydrogen-rich peptide cation radicals are subjected to homolytic bond cleavages that produce closed-shell and radical backbone fragments that provide sequence information.[62,64]



Scheme 1.2. Hydrogen-rich peptide cation radical formation via ETD using 18-crown-6-ether



Scheme 1.3. Hydrogen-deficient peptide cation radical formation via CID using Cu^{II}(tpy) and peptide alpha radicals.

Hydrogen-deficient peptide cation radicals, $[M]^{+\bullet}$, on the other hand, lack one hydrogen atom relative to their fully protonated, even-electron peptide ion counterpart.[110] These peptide cation radicals can be generated by various methods, including electron-induced dissociation of peptide ions,[111] photolysis,[112] and collision-induced dissociation,[76,113] among others;[114] however, only the method utilized in the work presented here is described in details below. The method,[115-118] developed by the Siu lab at York University, involves the non-covalent binding of the desired peptide to a copper (II) ligand, i.e. Cu(II) terpy, where terpy = 2,2':6',2"-terpyridine, in solution.[118] The complex formed at the free peptide carboxylate[119-121] is ionized and its doubly charged ion isolated before being dissociated under CID. During the dissociation process, the complex is broken apart and, rather than yielding a spectrum dominated by b and y ions, a one-electron transfer reaction occurs, which generates the hydrogen-deficient peptide cation radical in high yield, **Scheme 1.3**. Several isomers of these peptide cation radicals exist, notably the delocalized distonic cation radicals, where the charge and radical are spatially separated, the localized zwitterion cation radicals, where the charge and the

radical are close together and both termini are oppositely charged, and finally the localized canonical form, where the charge and the radical are close together and no other charge resides on the cation radical. Determining which structures contribute to a given peptide cation radical is greatly valuable in understanding the underlying mechanisms that cause the dissociation patterns observed via tandem mass spectrometry.

The overall motivation for the study of peptide cation radicals is that the dissociation of these odd-electron species provides reliable and informative insight as to the structure of a peptide.[106] Indeed, peptide cation radicals are far less subject to partial sequence scrambling due to backbone rearrangements than their even-electron counterparts.[122] Additionally, the bond connectivity (structure) and geometry of peptide cation radicals is identical to the precursor molecule, which implies that parallel conclusions can be drawn about the two species.[123] Consequently, knowledge of the mechanisms that regulate the formation and dissociation of peptide cation radicals is essential to learn the fundamental biological processes that cause, for example, protein damage,[124] and which explains the current wealth of MS-based research on peptide radicals.[125]

1.3 Protein and Peptide Sequencing via Mass Spectrometry

1.3.1 Charge-tags in Protein and Peptide sequencing

Peptide tagging has been recently used for quantitative purposes;[88] however, the major impetus for developing tags toward protein and peptide sequencing rests on the complexity of reliable peak assignments. While some methods allow intact protein studies,[126-128] proteins of interest are commonly analyzed in a bottom-up approach, whereby the protein is first enzymatically digested into its smaller peptide units before rebuilding the entire identified sequence. Tagging the resulting complex mixture of peptides can prove useful in achieving lower spectral complexity and better sequence assignment. With the emergence of ETD in proteomics as an ideal ion activation method that does not alter post-translational modification (PTMs),[129] several charge-tags have been developed.[130-135] Indeed, the efficiency of ETD was shown to increase with the increasing charge-state of a peptide,[136] hence the motivation in the development of these fixed-charged tags. Several synthetic methods of charge-tags have been utilized,[133] and the one used in the described work involves the guanidination of the lysine side chain at the ε-amino group.[93] The guanidination of this side-chain

renders the lysine residue more basic by modifying it into a homoarginine, **Figure 1.2**, which localizes a fixed charge on the derivatized peptide[137] and thus, increases the charge state of the peptide for a more efficient dissociation by ETD. Though valuable by itself, another dimension can be added to the guanidination of the lysine side-chain by inserting unique components to the guanidine moiety, which distinguishes it from any other non-derivatized peptide. One such unique component currently exploited in peptide sequencing is the addition of a mass defect, described below.

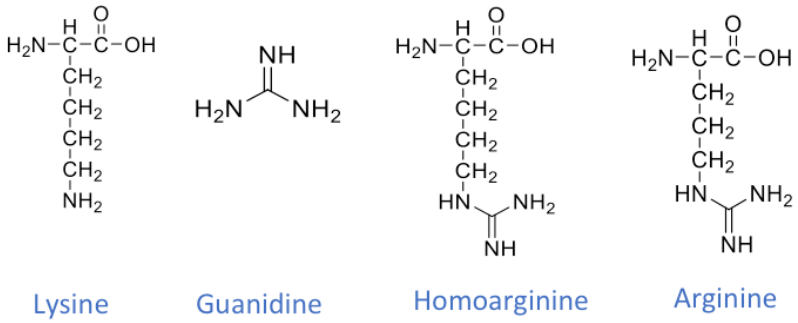


Figure 1.2. Structures of lysine and its guanidinated counterpart, homoarginine, with one additional carbon atom relative to arginine

1.3.2 Mass Defect Labeling and Filtering

A certain amount of energy, known as the binding energy, is supplied to a nucleus during its formation. It can be thought of as the amount of energy released if a nucleus was to be formed from separate particles coming together. This release of energy translates into a "loss of mass" as a portion of the mass of a nucleus is converted into that binding energy.[138-140] Thus, the mass of any given atom or molecule is typically slightly less than the sum of its individual protons, neutrons, and electrons. This discrepancy is referred to as the mass defect (MD) and it is greatly illustrated in Einstein's equation $E = mc^2$, in which mass is related to energy. The following equation gives the mass defect of an atom:

$$\Delta m = [\, Z\,(m_p + m_e) + (A - Z)\, m_n \,] - m_{atom}$$

where Δm is the mass defect (amu), m_p is the mass of a proton (1.007277 amu), m_n is the mass of a neutron (1.008665 amu), m_e is the mass of an electron (0.000548597 amu), m_{atom} is the mass of the nuclide (amu), Z is the atomic number (number of protons) and A is the mass number (number of nucleons). For example, iodine, with Z = 53, a nominal mass A = 127 –defined as the integer mass of an atom, and an atom mass of 126.90004 amu has a mass defect Δm = 1.1574. The definition of mass defect is in the field of mass spectrometry slightly varies from that above and simply refers to the difference in the nominal mass of an atom (i.e. 127 amu for iodine) and its exact mass (126.90447 amu), which gives a mass defect of -0.09553 for iodine. By convention, carbon was assigned a mass defect of zero, while any other existing isotope has either a positive or negative mass defect relative to its binding energy to ^{12}C, **Table 1.1**. The mass defect can be exploited as a valuable marker in protein sequencing. Proteins and peptides are made up of a small subset of elements, the most abundant of which –carbon (MD = 0.000), hydrogen (MD = +0.0078), nitrogen (MD = +0.0031) and oxygen (MD = -0.0051)- have negligible mass defect relative to ^{12}C.[141] Other elements found in proteins have larger mass defects, but these are also less abundant in the protein makeup; they include sulfur (MD = -0.0279) and phosphorus (MD = -0.0262). In the typical mass spectrum of a naturally occurring protein or peptide, the distribution of masses from a given combination of the twenty available amino acids is heterogeneous and distributed in clusters separated by spaces devoid of spectral information, also known as "gaps" or "forbidden zones", usually at around one Dalton's distance from one another, where no common elemental composition would exist, even in noisy or complex spectra.[140-142] These empty spectral areas, in which only either modified peptides or non-peptide compounds may exist, can be taken advantage of in protein sequence identification with high accuracy mass measurements. Seminal work by Mann *et al*[144] on the mass distribution of all possible theoretical peptides first reported the existence of this clustered distribution of peptide masses and laid grounds for research on the potential benefits of those vacant zones; the work was expanded by Nefedov *et al*[143,145] for tryptic peptides as large as 3.5 kDa. Two methods emerged from utilizing the mass defect as an analytical tool for complex data; namely, mass defect filtering and mass defect labeling.

Mass defect filtering refers to a peak identification method that uses a restricted range of possible mass defect that encompasses a class of compounds.[142] Any identified compound having a mass defect outside of that range is essentially filtered out during data processing, post-acquisition. The less common the mass defect range, the more easily the data of interest is extracted from background noise or signal from

other compounds. The work of Taguchi *et al*[146] illustrates successful mass filtering for the identification of an unknown series of polychlorinated compounds using a Kendrick plot of mass defect (y-axis) versus nominal mass (x-axis). Other prominent applications of high-resolution mass defect filtering reviewed by Sleighter *et al*[140,147] include drug metabolite detection[148] and natural organic matter studies, such as humic acid comparisons based on the mass defect trends of compound classes.

Element	Isotope	Atomic Mass (Da)	Mass Defect	% isotopic composition
Hydrogen	^{1}H	1.00783	0.00783	99.9885
	^{2}H	2.01410	0.01410	0.0115
Carbon	^{12}C	12.00000	0.00000	98.93
	^{13}C	13.00335	0.00335	1.07
Nitrogen	^{14}N	14.00307	0.00370	99.632
	^{15}N	15.00011	0.00011	0.368
Oxygen	^{16}O	15.99491	-0.00509	99.757
	^{17}O	16.99913	-0.00087	0.038
	^{18}O	17.99916	-0.00084	0.205
Fluorine	^{19}F	18.99840	-0.00160	100
Phosphorus	^{31}P	30.97377	-0.02623	100
Sulfur	^{32}S	31.97207	-0.02790	94.93
	^{33}S	32.97146	-0.02854	0.76
	^{34}S	33.96787	-0.03213	4.29
Chlorine	^{35}Cl	34.96885	-0.03115	75.78
	^{37}Cl	36.96590	-0.03410	24.22
Arsenic	^{75}As	74.92160	-0.07840	100
Bromine	^{79}Br	78.91834	-0.08166	50.69
	^{81}Br	80.91629	-0.08371	49.31
Yttrium	^{89}Y	88.90585	-0.09415	100
Iodine	^{127}I	126.90447	-0.09553	100
Terbium	^{159}Tb	158.92534	-0.07466	100
Holmium	^{165}Ho	164.93032	-0.06968	100

Figure 1.3. Table of mass defects for the various isotopes of elements of interest[140]

The second approach, known as mass defect labeling, consists in tagging either an intact protein or the peptides resulting from a digested protein with the goal of increasing their detectability. Using accurate mass measurement, mass-defect labeling explores ways to incorporate an uncommon molecule or atom -such as chlorine,[141] 1 bromine[149] or iodine -with a singular mass defect that is not typically found in natural amino acid sequences.[140,141] The purpose of the method, as described by Bajrami et al, is to force a given protein, peptide or peptide fragment ion into the "forbidden zones" of the spectrum by shifting its mass.[100] In doing so, the peptides are moved in a less noisy spectral space, thereby making their detectability more likely.[140] This approach is particularly attractive for the analysis of complex bottom-up proteomics data that result from the digestion of a protein by a given protease and subsequently matched to potential sequences using a spectral database. The greatest shortcoming of the reliance on spectral databases is the propensity for false positive results, which catalyzed the need for *de novo* studies.[151] Several applications of mass defect labeling can be found in the literature and is becoming an attractive method to simplify spectral complexity.[139,149, 152-159] Moreover, this approach finds value in *de novo* sequencing, which relies on the analytical process of deriving a peptide's amino acid sequence from its experimentally obtained mass spectrum (MS/MS) with no reference to sequence databases.[159] An example of *de novo* application is the analysis of a novel or unknown protein for which no known spectrum exists. The mass defect utilization will be further discussed as we propose our newly developed tag as a tool for *de novo* sequencing.

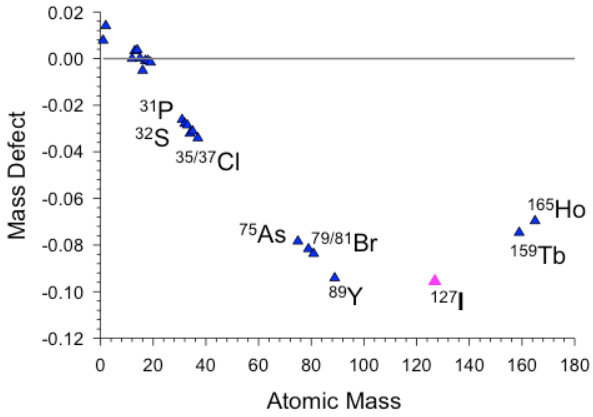


Figure 1.4. Plot of mass defect versus atomic mass for elements of interest*1.3.2 De Novo Sequencing*

The major advantage of *de novo* sequencing lies in the possible identification of both known and entirely novel peptides. The use of tandem mass spectrometry, MS^n, (whereby the product ion from a fragmented precursor ion can itself be isolated and further fragmented) is crucial for *de novo* analysis, where the full sequence of a peptide is reconstituted from its individual fragments. *De novo* sequencing makes use of the mass difference between two fragment ions to identify the corresponding amino acid residue in the peptide backbone.[160] If one ion series can be identified, the mass difference between two adjacent ions in that series can be calculated and assigned to a known amino acid until the complete sequence is identified. It should be noted that *de novo* sequencing is simplified when a single ion series stands out of a complex spectrum and that, reversely, it is greatly complicated when more than one ion series appear in the spectrum, hence the interest in steering the fragmentation towards a single ion series. *De Novo* sequencing can not only be useful in the analysis of a novel peptide, for which no spectral comparison via database is available, it can also be used when a database is available for validation of a peptide identification obtained via direct spectral comparison or when the presence of novel peptide mutations is suspected.[161,162] Additionally, though the use of manual *de novo* sequencing alone is a rather tedious and slow method of peptide determination, commercially available software can assist in the automation of *de novo* analysis and help speed the analytical process.[151] One such software, PEAKS, has established itself has a powerful tool in de novo peptide sequencing[162] and has spurred the development of the UniNovo algorithm, a decade later, which can be applied to spectra obtained via CID, ETD and HCD, or a combination thereof.[163] These algorithms have helped bring *de novo* analyses back to the table and the method shows a promising future as a fierce competitor of spectral database searches.

Based on all the factors stated above, we proposed the development of a novel iodinated charge tag aimed at decreasing the complexity of data obtained from protein digests and aimed at enhancing the ETD backbone cleavage and thus, the sequence information of peptides for *de novo* analysis. The results of this work are thoroughly discussed in **Chapters 4 and 5**.

Chapter 2 Ground and Excited-Electronic-State Dissociations of Hydrogen-Rich and Hydrogen-Deficient Tyrosine Peptide Cation Radicals

2.1 Introduction

Whereas isolated peptide molecules and ions are inherently stable species with closed-electronic shells, perturbation of their electronic structure results in novel, and often, unexpected chemistry.[104] There are now several ways of generating peptide radicals and cation-radicals in the gas phase to be investigated by mass spectrometry. Resonant multiphoton ionization of gas-phase peptide molecules is the most direct method that was reported as early as in 1986.[164-166] One-electron oxidation of peptide ligands upon collision-induced dissociation of ternary complexes of transition metal ions represents another useful method for generating peptide cation-radicals.[115,116,118,167,168] Thermal[113] or light-induced[79] homolytic bond dissociations in suitably derivatized peptide ions have also been developed into methods of peptide cation-radical generation. All these techniques produce peptide cation radicals that are stoichiometrically equivalent to neutral peptide molecules and because they do not contain additional hydrogens from protonation; they are called "hydrogen deficient".[104] Of note is that the structures of hydrogen-deficient peptide cation-radicals, wherever the charge and radical sites are known to be, depend on the particular activation method.

A fundamentally different method of generating peptide cation-radicals relies on electron attachment to multiply protonated peptides via capture of a slow free electron[61] or collisional transfer from a suitable neutral[60,169-173] or anionic donor.[19] Peptide cation-radicals of this type are called "hydrogen rich". They are often unstable and undergo extensive dissociation by multiple channels, forming backbone fragment ions and providing sequence information for protein analysis.[104] The majority of cation-radical fragment ions produced by electron attachment arises by cleavage of bonds between an amide nitrogen and the Cα atom of the adjacent amino acid residue, N—C_α cleavage for short, retaining the radical site in C-terminal residues (z-type ions).[20] When singly protonated, z ions are stoichiometrically equivalent to deaminated hydrogen-deficient peptide cation-radicals and have been considered as representing a connecting bridge between the chemistries of both types of peptide cation-radicals.[111]

Selected peptide cation-radicals of either type have been studied with the aim of determining their structure and elucidating dissociation mechanisms.[110,174-183] However, despite there being a large body of data on peptide cation-radicals and substantial progress in understanding their chemistry, there has not yet been a study exploring side by side the properties of both hydrogen- deficient and hydrogen-rich peptide cation-radicals derived from the same peptide molecules. Here, we report a comparative study of the generation and unimolecular dissociations of a series of tyrosine containing peptides in their hydrogen-rich and hydrogen-deficient radical forms.[74] We adopt the standard approach of systematically varying the amino acid sequence in Ala-Ala-Ala-Ala-Tyr-Arg (AAAYR) to generate and study both types of derived cation-radicals, as well as those from sequence isomers AAYAR, AYAAR, and YAAAR. The C-terminal arginine residue is employed throughout to anchor the charging proton at a specific site. Additionally, substitution of Val for Ala is used in AVAYR, AAYVR, AYAVR, and YAAVR to aid fragment ion identification through predictable mass shifts. Specific methods of generating these peptide cation-radicals are reported and their dissociations are studied in ground electronic states using collisional activation and in excited electronic states using UV-VIS photodissociation at wavelengths targeting the radical chromophores.

2.2 Experimental Section

2.2.1 *Materials and Methods*

2.2.1.1 Hydrogen-Deficient Peptide Cation Radicals

All peptides were synthesized on Wang resin (Bachem Americas, Torrance, CA, USA) using commercially available Fmoc peptides (Life Technologies, Rockford, IL, USA) and purified by ion-exchange chromatography. Fmoc-(O-methyl)-L-tyrosine was purchased from Santa Cruz Biotechnology (Paso Robles, CA, USA) and used in standard solid-state peptide synthesis of AAY(OCH3)R. The metal-ligand 2,2':6',2''-terpyridine (tpy) and Cu(NO3)2 were obtained from Sigma-Aldrich. When generating peptide radicals using metal complexes, 600 μM CuII(tpy)(NO3)2 was added to the peptide stock solutions (50 μM), without acetic acid, such that peptide radical ions ($M^{+\bullet}$) would be generated through one-electron transfer from the neutral peptide to the metal center in the metal–tpy–peptide ternary complex. For the experiments performed with the triple-quadrupole instrument, $[CuII(tpy)M]^{2+\bullet}$ was first introduced to the ion source by the electrospray

with an ion spray voltage of 3.5 kV. The $M^{+\bullet}$ species was then generated through in-source fragmentations of the complexes using N_2 as the collision gas under a declustering potential of 35 eV. The resulting $M^{+\bullet}$ species was selected in the first quadrupole (Q1) and further dissociated in the collision cell under the MS^2 mode.[184] The specific α-carbon-centered radicals were formed through multistage CID of $[CuII(tpy)(M)]^{2+\bullet}$, in which peptide canonical radical cations $M^{+\bullet}$ were generated in the first stage of CID of metal-tpy-peptide complexes and then underwent a subsequent stage of CID to generate the α-carbon-centered radical through tyrosine side-chain loss. For the experiments performed with the LTQ-XL linear ion trap mass spectrometer, electrospray ionization with a home-built microspray source was used to generate gas-phase $[CuII(tpy)peptide]2^{+\bullet}$ ions according to the above-described procedure. The ^{65}Cu or ^{63}Cu isotopologues were selected by mass and subjected to collision-induced dissociation (CID) at collision energies that were tuned to optimize peptide cation-radical formation, typically at normalized collision energies (NCE) set to 15-18 instrument units. The $[peptide]^{+\bullet}$ ions at the corresponding m/z were selected by mass and subjected to CID or photodissociation (UVPD).

2.2.1.2 Hydrogen-Rich Peptide Cation Radicals

When generating peptide radicals using crown-ether complexes, 1 μM solution of 18-crown-6-ether was added in equal volumes (1:1) to the peptide stock solutions (50 μM), before being electrosprayed and dissociated by electron transfer. These electron transfer dissociation mass spectra were obtained on a modified LTQ-XL linear ion trap mass (LIT) spectrometer (ThermoElectron Fisher, San Jose, CA, USA) equipped with a laser system. Doubly charged peptide ions or their crown-ether complexes were produced by electrospray ionization, mass-selected in the ion trap and allowed to react with fluoranthene anions at typical reaction times of 200 ms. The CID and UVPD mass spectra of mass-selected peptide cation-radicals were measured on the LTQ-XL. High-resolution mass spectra were measured on an LTQ-Orbitrap (ThermoElectron Fisher, San Jose, CA, USA). The peptide cation-radicals were prepared by CID in the LTQ and transferred to the Orbitrap for high-resolution measurements using Fourier-Transform treatment of the time-domain signal. The mass resolution was set to 100,000.

2.2.2 *Photodissociation*

Photodissociation of trapped ions in the LIT was performed as reported previously.[102] The typical experimental set up consists of selecting the ion to be photodissociated and storing it in the LIT for a chosen time period. For example, 400-ms storage time can accommodate up to 7 laser pulses, each spaced by a 50 ms time interval. This allows one to vary the number of pulses, which are also normalized, as the number of pulses used is varied according to the degree of dissociation observed. The pulse-dependent UVPD measurements were performed with the 355 nm line from the laser source at 15 mJ/pulse laser power, as described previously.[199]

2.2.3 *Calculations*

Standard *ab initio* and density functional theory (DFT) calculations were carried out using the Gaussian 09 suite of programs.[185] Ion structures were gradient-optimized with the B3LYP,[186] M06-2X,[187] and ωB97X-D[188] hybrid DFT methods all using the 6-31+G(d,p) basis set, and local energy minima were confirmed by harmonic frequency analysis. Single-point energy calculations were conducted with DFT and Møller-Plesset perturbational treatment[189] (MP2, frozen core) using the 6-311++G(2d,p) basis set.

2.2.4 *Note on Nomenclature*

To describe in a uniform way the fragment ions originating from different types of peptide cation-radicals, we adopt the all- inclusive nomenclature system introduced recently.[190] Briefly, this uses the original Biemann system of naming peptide fragments according to the backbone bond being broken,[20] but explicitly assigning the charge, radical, and number of hydrogen atoms included in the fragment ion.[190] The equivalent names for singly charged fragment ions are as follows (n = number of residues): $[a_n]^+ = a_n$, $[b_n]^+ = b_n$, $[c_n + 2H]^+ = c_n$, $[c_n + H]^{+\bullet} = c_n - 1$, $[x_n + H]^{+\bullet} = x_n$, $[y_n + 2H]^+ = y_n$, $[y_n]^+ = y_n - 2$, $[z_n + H]^{+\bullet} = z_n$, $[z_n + 2H]+ = z_n + 1$. Residues in fragment ions arising by side-chain loss from Tyr are named $G\alpha^{\bullet}$.

2.3 Results & Discussion

2.3.1 *Hydrogen-Rich Peptide Cation-Radicals*

2.3.1.1 CID of Hydrogen-Rich Peptide Cation Radicals

Electron transfer dissociation of the doubly charged tyrosine peptide ions resulted in extensive dissociation, forming abundant $[z1\text{-}z4 + H]^{+\bullet}$ sequence fragment ions, thereby leaving very weak survivor $[M + 2H]^{+\bullet}$ peptide cation radicals that overlapped with residual ^{13}C isotope satellites of the $[M + H]+$ fragment ions (**Figure 2.7a-d, SI**). To increase the $[M + 2H]^{+\bullet}$ yield upon ETD, we resorted to a previously developed technique in which doubly charged non-covalent peptide were complexes (1:1) with 18-crown-6-ether (CE) and treated by ETD.[101,190] This resulted in charge-reduction and loss of the CE ligand, forming $[M + 2H]^{+\bullet}$ cation-radicals at relative abundances that allowed them to be further investigated by tandem (MS^n) experiments. In addition to the ligand loss upon reduction, ETD also produced CE-coordinated backbone fragment ions of the N-terminal (c) and C-terminal (z) types (**Figure 2.8a-d, SI**). These allowed us to estimate the populations of peptide complexes in which the crown ether was attached to the arginine guanidinium and N-terminal ammonium, which are the charged groups in these ions.

The data (**Table 2.1, SI**) indicated somewhat preferential coordination to the arginine charged group (56-73%) except in AAYAR and AAYVR where the N-terminal coordination was slightly more prevalent (57% and 52%, respectively). The figures for N-terminal coordination should be regarded as lower limits, because the ETD spectra of the complexes showed very abundant $[M + 2H - NH3]^{+\bullet}$ fragment ions that could originate from charge-reduced N-terminal CE-H3N- complexes. More discussion of this aspect will be given later in the text.

Collision-induced dissociation of $[M + 2H]^{+\bullet}$ ions generated by ETD of the CE complexes showed a marked dependence on the amino acid sequence (Figure 2.1). In general, the MS3/ETD-CID spectra were substantially different from the ETD spectra of $[M + 2H]^{2+}$ ions (direct dissociation, no crown ether used) as is obvious from the comparison of the **Figure 2.7a-d (SI)** and **Figure 2.1a-d** spectra. The CID fragmentation pattern of $[AAAYR + 2H]^{+\bullet}$ (m/z 552) showed a dominant loss of a hydrogen atom (m/z 551) (**Figure 2.1a**). The other abundant dissociation channels were loss of water (m/z 534), loss of CH5N3 from the arginine side chain (m/z 493), and the formation of the $[z1 + 2H]^+$ backbone fragment ion at m/z 160. In contrast,

backbone fragment ions of the standard $[zn + H]^{+\bullet}$ type (n = 1-4) were not formed. The minor presence of a fragment ion at m/z 444, which corresponds to a loss of 106 Da from the tyrosine side chain (-C7H6O, neutral fragment), is also seen. This loss is commonly observed for hydrogen-deficient tyrosine-containing peptide cation radicals,[115,116,118,167,168] and the ion is denoted as $[AAAG_{\alpha}{}^{\bullet}R + H]^{+}$, indicating a conversion of the Tyr residue to a glycine C_{α}-radical.[190]

As tyrosine was moved away from the C-terminus by one residue, the resulting $[AAYAR + 2H]^{+\bullet}$ cation radical showed a different fragmentation pattern upon CID (**Figure 2.1b**). In addition to the same four major channels observed for $[AAAYR + 2H]^{+\bullet}$, the $[AAYAR + 2H]^{+\bullet}$ fragmentation displayed a greater complexity in the formation of backbone fragments. The major ones were identified as $[b3]^{+}$ and $[b4]^{+}$ ions along with $[y3 + 2H]^{+}$ and $[y2 + 2H]^{+}$ ions, a prominent $[y1 + 2H]^{+}$ ion, as well as $[z2]^{+}$ and $[z3]^{+}$ ions, and an $[a4]^{+}$ fragment ion.

The dissociation patterns of $[AYAAR + 2H]^{+\bullet}$ and $[YAAAR + 2H]^{+\bullet}$ lack backbone fragments. CID of $[AYAAR + 2H]^{+\bullet}$ resulted in five distinct reactions, with a dominating loss of an H atom and formation of prominent fragment ions at m/z 444, m/z 493, m/z 160 and m/z 534, which were analogous to those described above (Figure 2.1c). The same dissociations were observed for $[YAAAR + 2H]^{+\bullet}$, which showed a more abundant elimination of C7H6O compared to the above-described sequences (Figure 2.1d). This dissociation is usually associated with the presence of a tyrosyl-O radical moiety.[168,80] However, the CID spectra do not allow us to distinguish if Tyr-O radicals were present to a different extent in the stable $[M + 2H]^{+\bullet}$ cation-radicals or if they were produced as reactive intermediates in the course of dissociation. The identity of the peptide cation radicals and the assignments of their fragments following CID were corroborated by mass shifts in the CID spectra of cation-radicals derived from AVAYR, AAYVR, AYAVR, and YAAVR. The pertinent spectra are presented in the Supplement (**Figure 2.9a-d**).

The dominating dissociation pathways for this sequence were loss of water (m/z 534), the formation of $[z1 + 2H]^{+}$ (m/z 160), loss of an H atom (m/z 551) and loss of CH5N3 from the arginine side chain. Also worth noting is the presence of the fragment ion at m/z 444 that results from the tyrosine side chain loss (-C7H6O), which is significantly more abundant than for $[AAAYR + 2H]^{+\bullet}$.

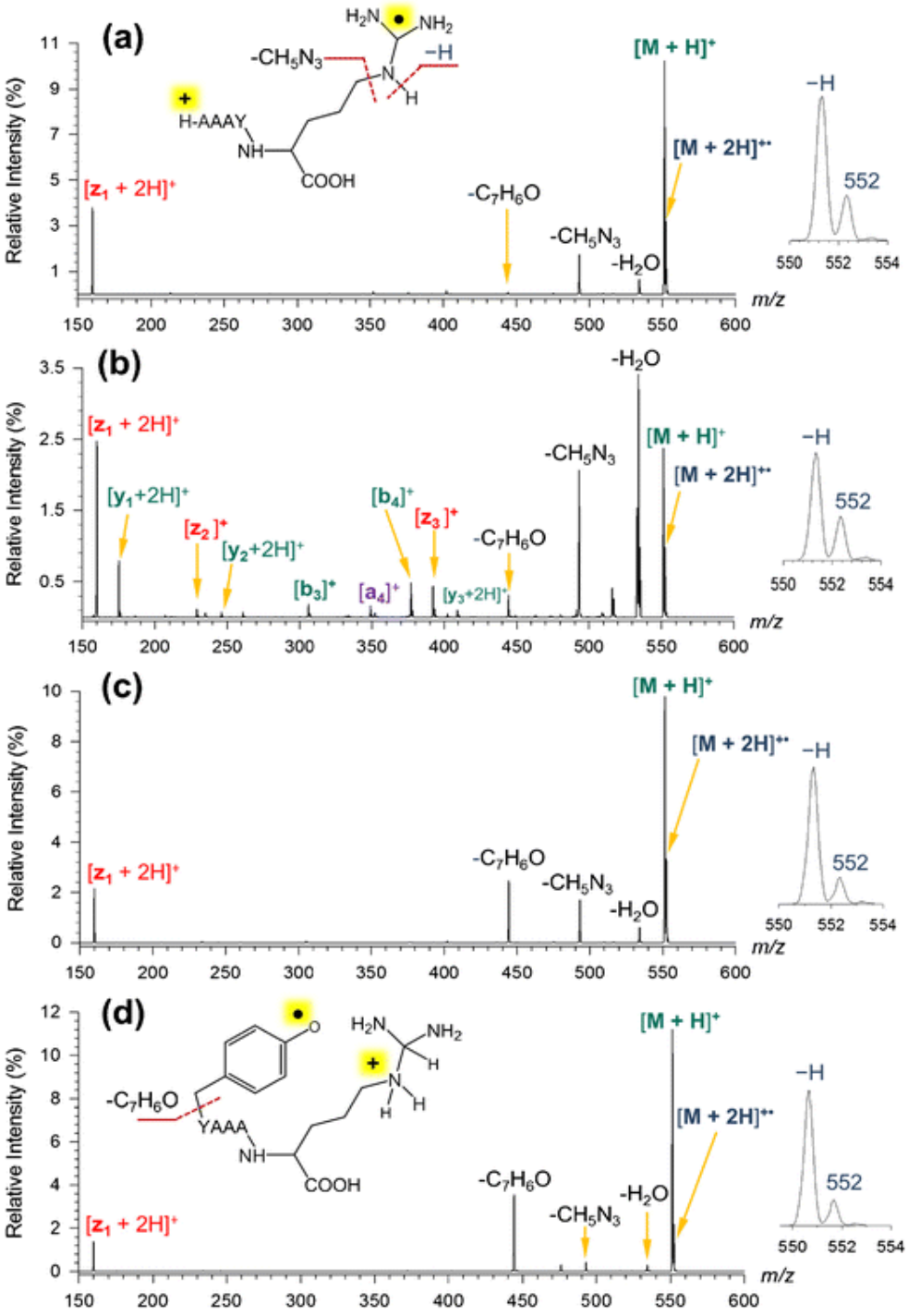


Figure 2.1. CID-MS3 mass spectra of hydrogen-rich peptide cation-radicals (m/z 552) generated by ETD of crown-ether complexes: (a) $[AAAYR + 2H]^{+\bullet}$, (b) $[AAYAR + 2H]^{+\bullet}$, (c) $[AYAAR + 2H]^{+\bullet}$, (d) $[YAAAR + 2H]^{+\bullet}$.

2.3.1.2 UVPD of Hydrogen-Rich Peptide Cation Radicals

Photodissociation (UVPD) of hydrogen-rich peptide cation-radicals was studied at 355 nm, wavelength at which excitation of peptide radical chromophores is selectively targeted.[199] UVPD of $[AAAYR + 2H]^{+\bullet}$ through $[YAAAR + 2H]^{+\bullet}$ (**Figure 2.2a-d**) gave rise to two major fragment ions, one (m/z 551) being formed by photoinduced loss of a hydrogen atom and the other (m/z 160) corresponding to a backbone $[z1 + 2H]^{+}$ ion (see inset structure in **Figure 2.2a**).

Again, no standard backbone fragment ions of the $[z + 2H]^{+\bullet}$ type were formed, indicating that even the excited electronic states of the long-lived cation-radicals were different from the electronic states accessed by electron transfer to the doubly charged ions in the ETD mode. The UVPD fragment ions were photostable at 355 nm and their relative intensity reached a steady state after four laser pulses (**Figure 2.10a-d, SI**).

Photodissociation of the peptide cation-radicals at 355 nm must be associated with a chromophore formed by electron attachment, as the natural amino acid residues Ala, Tyr, and Arg do not absorb light at this wavelength. Likewise, the even-electron fragment ions at m/z 551 and 160 were transparent at 355 nm and did not undergo further photodissociation. The slight decrease of the m/z 160 relative intensity at long trapping times needed to accommodate multiple laser pulses (**Figure 2.10a-d, SI**) is probably caused by a less efficient trapping at the edge of the LTQ stability region corresponding to 160/552 = 0.29 of the selected precursor ion m/z.

UVPD showed a substantial depletion of the precursor cation-radical populations after one laser pulse that exceeded 50% for all four peptide ions. However, all of them showed populations of residual photo-inactive cation-radicals that did not undergo dissociation, asymptotically converging to 23, 23, 16, and 6% for $[AAAYR + 2H]^{+\bullet}$, $[AAYAR + 2H]^{+\bullet}$, $[AYAAR + 2H]^{+\bullet}$ and $[YAAAR + 2H]^{+\bullet}$, respectively (**Figure 2.10a-d, SI**). Hence, the pulse-dependent experiments indicated that the populations of each of these cation-radicals were not homogeneous and contained isomers of different light absorption properties.

The valine peptide analogues showed a similar behavior. UVPD resulted in a loss of H as the predominant fragment ion (**Figure 2.11a-d, SI**). Formation of the $[z1 + 2H]^{+}$ ion, albeit expected, could not be confirmed because the ion's m/z was below the low-mass cutoff of the LTQ, 160/580 = 0.276. UVPD of $[YAAVR + 2H]^{+\bullet}$ also produced a minor fragment by loss of C7H6O from the Tyr side chain. The pulse

dependence of the charge-reduced AVAYR, AAYVR, AYAVR, and YAAVR cation-radicals (**Figure 2.12a-d, SI**) showed bimodal behavior consisting of rapid exponential depletion of the precursor ion within four laser pulses and leaving a substantial fraction of photo-inactive ions. The nature of the photoactive and photo-inactive isomers will be discussed later in the text.

2.3.2 *Hydrogen-Deficient Peptide Cation-Radicals*

2.3.2.1 CID of Hydrogen-Deficient Peptide Cation Radicals

Peptide cation-radicals of the hydrogen-deficient type were generated by CID of $[CuII(tpy)(peptide)]^{2+\bullet}$ complexes, according to the previously reported method.[118] The hydrogen-deficient peptide cation-radicals are denoted as $[M]^{+\bullet}$ to distinguish them from the hydrogen-rich analogues. Further collisional activation of $[M]^{+\bullet}$ was accomplished under a variety of conditions, including slow heating by resonant excitation in 3D and linear ion traps and acceleration of a mass-selected ion beam in a tandem quadrupole mass spectrometer. These CID experiments yielded qualitatively similar results; the LTQ spectra are presented and discussed in the main text, the other data are in the Supplement (**Figure 2.13a-d, SI**). Ion assignments in the CID spectra were corroborated by accurate mass measurements (**Table 2.2**, **SI**).

CID of $[M]^{+\bullet}$ produced MS3 spectra that showed a dependence on the position of the Tyr residue in the peptide sequence (**Figure 2.3a-d**). As the Tyr residue was moved toward the N-terminus, CID loss of C7H6O from the Tyr side chain formed fragment ions of increasing prominence that amounted to 8, 39, 75, and 77% of the total ion intensity for $[AAAYR]^{+\bullet}$, $[AAYAR]^{+\bullet}$, $[AYAAR]^{+\bullet}$, and $[YAAAR]^{+\bullet}$, respectively. The loss of C7H6O is usually associated with the presence of a Tyr O-radical,[115] and a recent UV-action spectroscopy study has identified a Tyr O-radical moiety in $[YAAAR]^{+\bullet}$.[80] Sequence fragment ions were observed for $[AAAYR]^{+\bullet}$ where the $[y4 + 2H]^{+}$ ion (m/z 480) was a dominant product, accompanied by ions of the z-series (m/z 159, 322, 392, and 464). It is of note that some fragment ions of the c and z series from $[AAAYR]^{+\bullet}$ are exactly isobaric, for example, both the $[c4]^{+}$ and $[z3]^{+}$ ions have the same theoretical m/z of 392.1928 for C18H26N5O5, consistent with the experimental m/z (**Table 2.2, SI**).

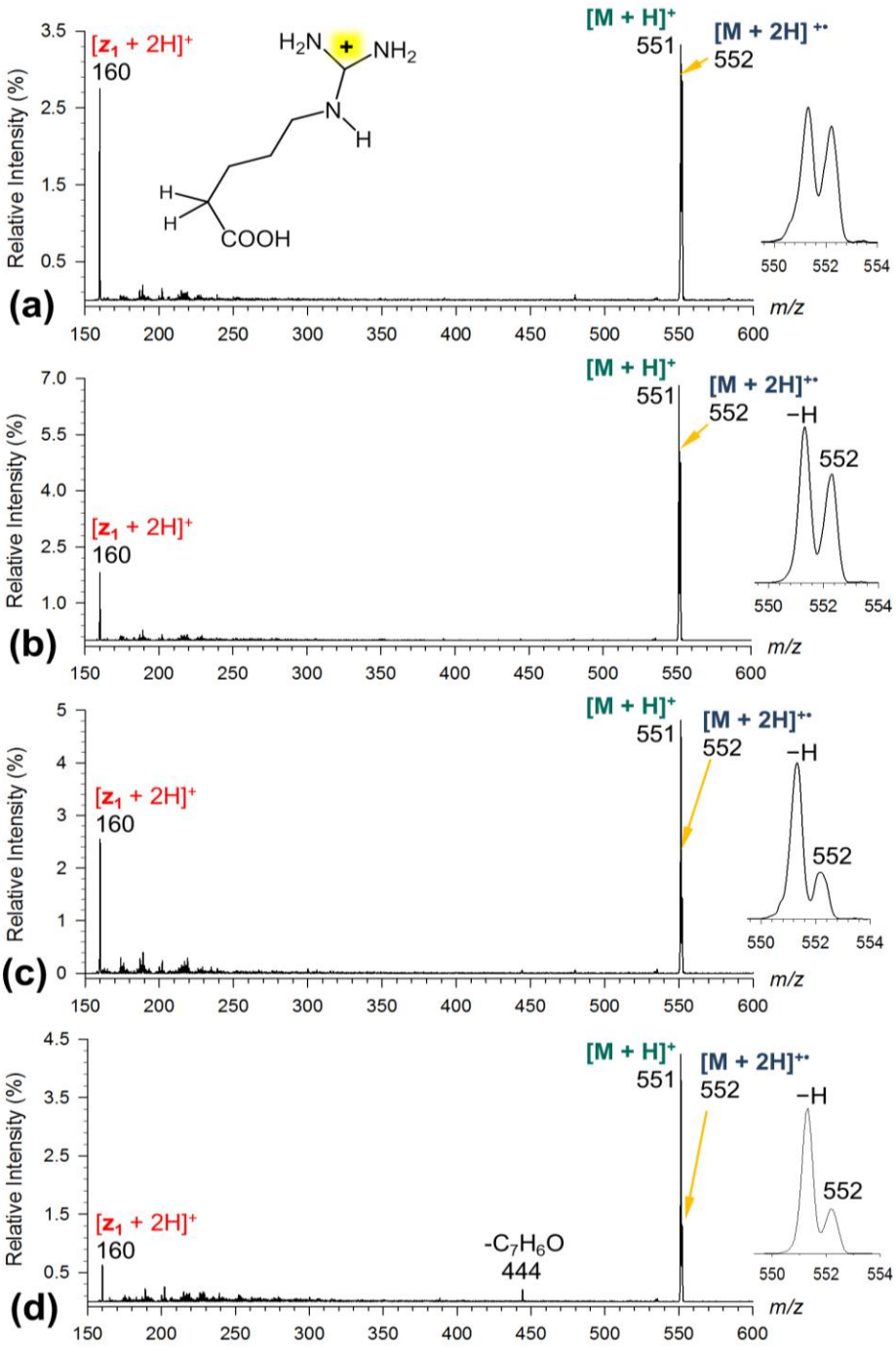


Figure 2.2 UVPD-MS3 mass spectra of hydrogen-rich peptide cation-radicals (m/z 552) generated by ETD of crown-ether complexes: (a) $[AAAYR + 2H]^{+\bullet}$, (b) $[AAYAR + 2H]^{+\bullet}$, (c) $[AYAAR + 2H]^{+\bullet}$, (d) $[YAAAR + 2H]^{+\bullet}$.

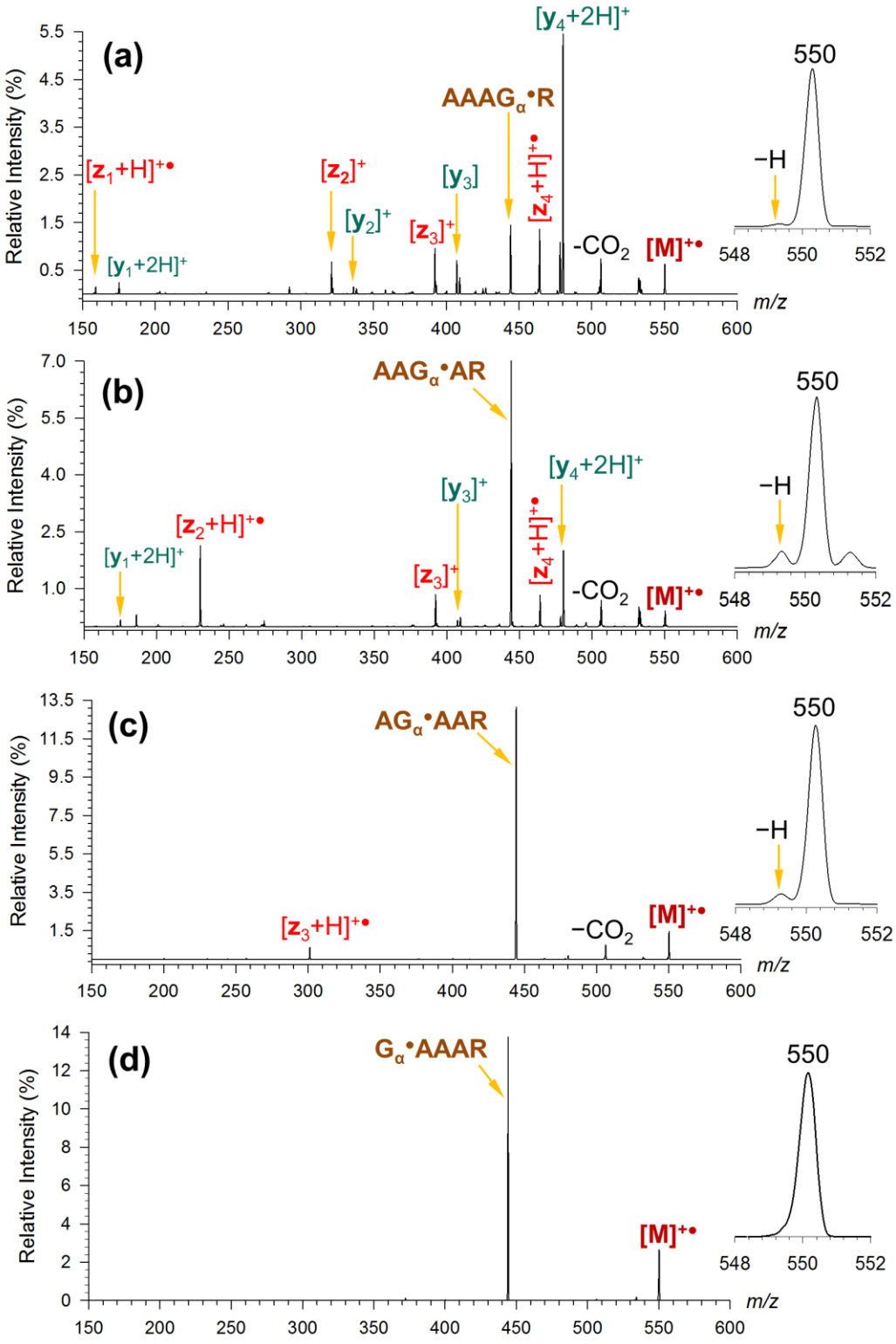


Figure 2.3. CID-MS3 mass spectra of hydrogen-deficient peptide cation-radicals (m/z 550) generated from Cu(tpy) complexes: (a) $[AAAYR]^{+\bullet}$, (b) $[AAYAR]^{+\bullet}$, (c) $[AYAAR]^{+\bullet}$, (d) $[YAAAR]^{+\bullet}$.

A definite assignment of the fragment ion type was achieved by considering mass shifts in the homologous series, $[AVAYR]^{+\bullet}$ and $[AAAYR]^{+\bullet}$, where there is no shift in the [z2]+ (m/z 321) and [z3]+ (m/z 392) ion masses (**Figure 2.14a-d, SI**) whereas a 28 Da shift would be expected for the equivalent $[c3]^+$ and $[c4]^+$ ions. Sequence ions of the z type were also observed for $[AAYAR]^{+\bullet}$, e.g., $[z2 + H]^{+\bullet}$, $[z3]^+$ and $[z4 + H]^{+\bullet}$, that were assigned by appropriate mass shifts in the spectrum of $[AAYVR]^{+\bullet}$ (**Figure 2.14b, SI**).

2.3.2.2 UVPD of Hydrogen-Deficient Peptide Cation Radicals

UVPD of $[M]^{+\bullet}$ showed some novel features. All sequence variants of $[M]^{+\bullet}$ were photoactive at λ=355 nm, pointing to radical-associated chromophores (**Figure 2.4a-d**).[102] UVPD of $[AAAYR]^{+\bullet}$ and $[AAVYR]^{+\bullet}$ showed dominant loss of H which was accompanied by minor loss of C3H4NO and formation of the $[y4 + 2H]^+$ fragment ions. Upon moving the Tyr residue to the next position in $[AAYAR]^{+\bullet}$ and $[AAYVR]^{+\bullet}$, photodissociation resulted in loss of H and C7H6O. Loss of C7H6O was dominant in the UVPD spectra of $[AYAAR]^{+\bullet}$ and $[YAAAR]^{+\bullet}$ (**Figure 2.4c,d**) as well as in the spectra of their valine homologues (**Figure 2.15a-d, SI**). In addition to differences in the photofragmentation patterns, the isomeric peptide cation radicals also differed in their photodepletion efficiencies, as illustrated by the pulse-dependent UVPD spectra (**Figure 2.16a-d, SI**). Thus, $[YAAAR]^{+\bullet}$ showed the most efficient photodepletion of the precursor ion at 50% after the first laser pulse (**Figure 2.16d, SI**). The main primary product ion, $[G_\alpha{}^\bullet AAAR]^+$ undergoes further photodissociation by loss of H and forming the $[y4]^+$ fragment ion as the main sequence product. Note that $[y4]^+$ in the current notation corresponds to a $(y4 - 2H)^+$ species by the previously used nomenclature and represents a common photodissociation product of peptide C_α-radical ions.[102] The formation of $[AG_\alpha{}^\bullet AAR]^+$ and $[AAG_\alpha{}^\bullet AR]^+$ intermediates is also seen in the photodepletion curves of the pertinent $[AYAAR]^{+\bullet}$ and $[AAYAR]^{+\bullet}$ ions (**Figure 2.16b,c, SI**) and their Val homologues (**Figure 2.17a-d, SI**). The isomeric Gly C_α-radical ions (m/z 444) were further investigated by CID and UVPD.

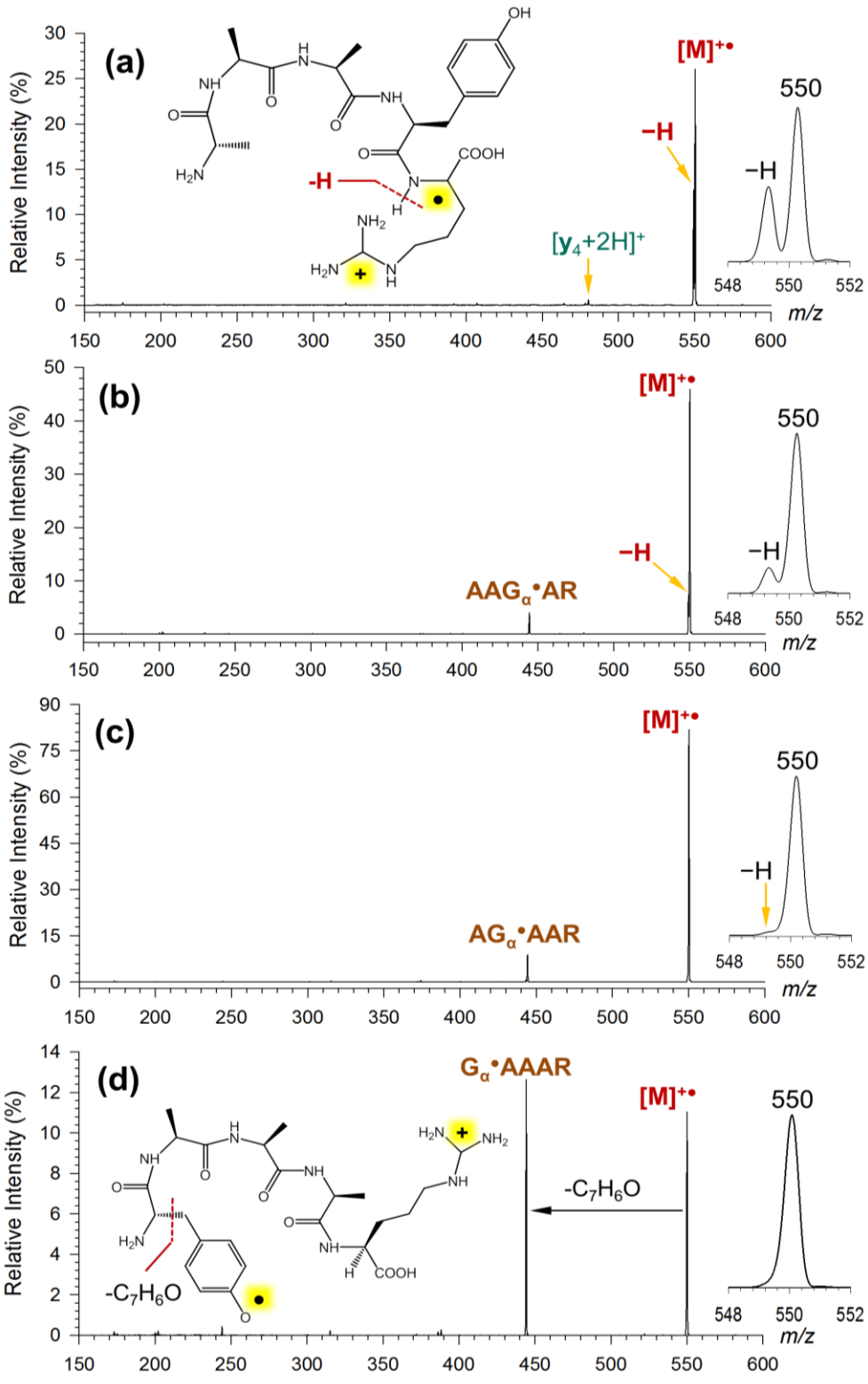


Figure 2.4. UVPD-MS3 mass spectra of hydrogen-deficient peptide cation-radicals (m/z 550) generated from Cu(tpy) complexes: (a) $[AAAYR]^{+\bullet}$, (b) $[AAYAR]^{+\bullet}$, (c) $[AYAAR]^{+\bullet}$, (d) $[YAAAR]^{+\bullet}$.

2.3.2.3 Effects of Gly C_α Radical Position on UVPD fragmentation

The Gly C_α radicals (m/z 444) were generated by CID-MS3 of $[M]^{+\bullet}$, isolated by mass and subjected to CID-MS4 and UVPD-MS4. The UVPD-MS4 spectra showed dissociations that were highly specific of the position of the Gly C_α radical site (**Figure 2.5a-d**). Starting with $[AAAG_\alpha{}^\bullet R]^+$, UVPD resulted in the formation of $[y2]^+$ (m/z 230, GR) and $[y3]^+$ (m/z 301, AGR) sequence fragment ions in addition to loss of H which was another abundant dissociation channel (**Figure 2.5a**). The formation of the $[y2]^+$ ion proceeds by β-fission of the CO–NH bond adjacent to the Gly_α radical and does not require a hydrogen migration. To form the $[y3]^+$ ion, the Hα from the adjacent Ala residue must migrate to form an Alaα radical that promotes CO–NH bond dissociation at the adjacent position.[102] Hydrogen atom migration in $[AAAG_\alpha{}^\bullet R]^+$ to the Gα radical from the proximate Arg C_α or C_β positions can also be considered to explain the loss of H forming a product ion with an α, β -dehydro-Arg moiety. Loss of H is either absent or much less abundant in the other isomers where the G_α radical is flanked by Ala residues. The laser pulse-dependent measurements of $[AAAG_\alpha{}^\bullet R]^+$ (**Figure 2.18a, SI**) indicated an exponential depletion of the $[AAAG_\alpha{}^\bullet R]^+$ ion intensity whereas the [y2] to [y3] abundance ratio was essentially constant at 4.1 to 1.

Photodissociation of the $[AAG_\alpha{}^\bullet AR]^+$ ion was even more specific, inducing CO-NH bond cleavage in the position adjacent to the Gly_α radical and producing the $[y3]^+$ ion (m/z 301, GAR) as the major product (**Figure 2.5b**). Noteworthy is the absence of loss of H in this case. These features are reflected by the photodepletion curve that shows rapid decrease of the $[AAG_\alpha{}^\bullet AR]^+$ relative intensity and formation of the $[y3]^+$ (GAR) fragment ion (**Figure 2.18b, SI**).

UVPD of $[AG_\alpha{}^\bullet AAR]^+$ also showed substantial specificity in breaking the CO–NH bond adjacent to the G_α radical and forming the $[y4]^+$ ion (m/z 372, GAAR)(Figure 2.5c). As a side reaction, Ala H_α atom migration moved the radical and triggered CO–NH bond dissociation between the Gly and Ala residues to produce the $[y3]^+$ (AAR) fragment ion at m/z 315. The overall photodepletion curve for $AG_\alpha{}^\bullet AAR]^+$ showed an exponential decay resulting in virtually complete dissociation after 7 laser pulses (**Figure 2.18c, SI**).

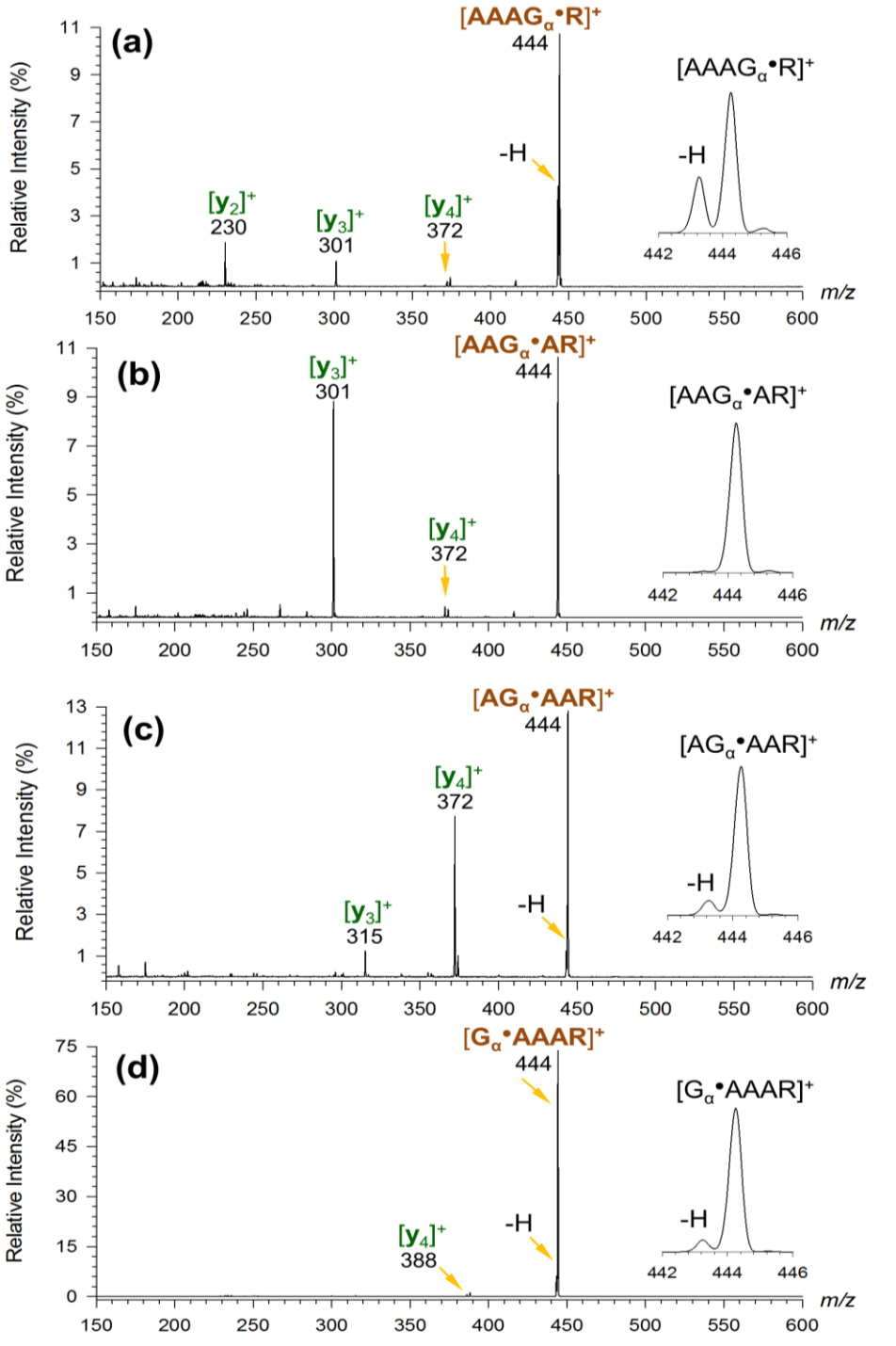


Figure 2.5. UVPD-MS4 mass spectra of m/z 444 ions for (a) [AAAGα•R]+, (b) [AAGα•AR]+, (c) [AGα•AAR]+, (d) [Gα•AAAR]+.

2.3.2.4 Effects of Gly Cα Radical Position on CID fragmentation

Finally, UVPD of $[G_{\alpha}^{\bullet}AAAR]^{+}$ was the least efficient of the four isomers (**Figure 2.18d, SI**) and resulted in loss of H, C_2H_2NO and C_2H4NO neutral molecules (**Figure 5d**). These minor backbone fragmentations indicate an H atom migration to and from the adjacent Ala residue, followed by CO–NH bond dissociations forming the $[y4 + 2H]^{+}$ and $[y4]^{+}$ ions. Loss of H can proceed by β-fission from the N-terminal amine group without need for H-atom migration.

In contrast to UVPD, CID-MS4 was much less specific to the radical position. The CID-MS4 spectra (**Figure 2.6a-d**) show a dominant formation of the $[y4 + 2H]^{+}$ ions from all four isomers, appearing at m/z 374 for $[AAAG_{\alpha}^{\bullet}R]^{+}$, $[AAG_{\alpha}^{\bullet}AR]^{+}$, $[AG_{\alpha}^{\bullet}AAR]^{+}$, and m/z 388 from $[G_{\alpha}^{\bullet}AAAR]^{+}$, indicating hydrogen atom migrations and preferential cleavage of the CO–NH bond between the N-terminal and adjacent residues. In spite of this dominant general feature, the CID-MS4 spectra show some differences in the fragment ion relative intensities. A conspicuous feature is the formation of even-electron $[z1]^{+}$-$[z3]^{+}$ ions by N—C_{α} bond cleavage accompanied by H atom migration onto the neutral fragment. In contrast, N—C_{α} cleavage of the bond connecting the N-terminal residue leads to $[z4 + H]^{+\bullet}$ cation-radical fragments in all sequences. In the absence of more detailed mechanistic investigation, these effects are difficult to relate to the ion structure.

2.3.2.5 Mechanistic Study of Hydrogen Loss

The prominent loss of H upon UVPD of $[AAAYR]^{+\bullet}$ and its virtual absence in CID raised questions as to both the ion structure and the origin of the hydrogen atom. To address these topics, we resorted to deuterium labeling in a smaller homologue, $[AAYR]^{+\bullet}$, which showed a rather similar behavior upon CID and UVPD, namely, abundant photo-induced loss of H, (**Figures 2.19, 2.20a, SI**), indicating that the main structure and reactivity features of $[AAYR]^{+\bullet}$ were analogous to those governing the dissociation of $[AAAYR]^{+\bullet}$. The prime advantage in using $[AAYR]^{+\bullet}$ rested in its lower number of exchangeable protons in H/D exchange experiments, and hence allowed greater conversion into the fully deuterated species. UVPD of completely H/D exchanged $[d_{11}\text{-}AAYR]^{+\bullet}$ showed a highly specific (94%) loss of D from one of the exchangeable positions (**Figure 2.20b, SI**). However, distinction among hydrogen atoms in the exchangeable positions is a daunting task that has been achieved using special techniques[191,192] that are not applicable to peptide ions. An

alternative method is to block by methylation a specific position in the ion, as previously applied in a study of peptide electron transfer dissociations.[59] To this end, we synthesized a AAY(OCH_3)R peptide in which the Tyr residue was replaced by O-methyl tyrosine, thus blocking the phenol hydroxyl group. UVPD of the [AAY(OCH_3)R]$^{+\bullet}$ ion (m/z 493, **Figure 2.21a, SI**) showed abundant loss of H in contrast to CID of the ion that generated backbone [z1 + H]$^{+\bullet}$, [z2]$^+$, and [y3 + 2H]$^+$ fragment ions, as well as an ion from CO2 loss (**Figure 2.22, SI**). H/D exchange in [AAY(OCH_3)R]$^{+\bullet}$ produced a d_{10} isotopologue (m/z 503) which upon UVPD showed a major loss of D forming the fragment ion at m/z 501 (**Figure 2.21b, SI**).

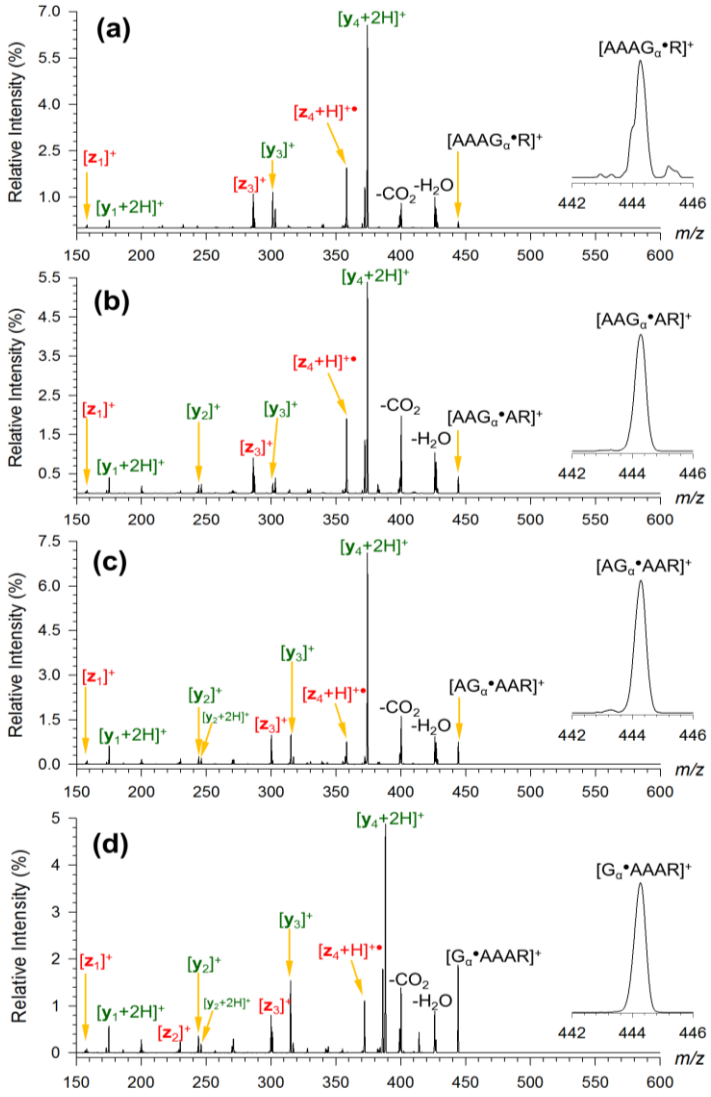


Figure 2.6 CID-MS4 mass spectra of (a) [AAAGα•R]$^+$, (b) [AAGα•AR]$^+$, (c) [AGα•AAR]$^+$, (d) [Gα•AAAR]$^+$.

These experiments indicated that the photo-induced loss of H from $[AAYR]^{+\bullet}$, and, by implication also from $[AAAYR]^{+\bullet}$, can originate from exchangeable positions other than the Tyr hydroxyl. It should be noted, however, that methylation of the Tyr residue changes its electronic and chemical properties in addition to blocking the phenol hydroxyl. This is indicated by the UVPD spectrum of $[AAY(OCH_3)R]^{+\bullet}$ which shows a specific formation of the $[y1]^+$ fragment ion (m/z 158, **Figure 2.21a, SI**), which is much less abundant when produced from $[AAYR]^{+\bullet}$ (**Figure 2.20a, SI**). A pulse dependence experiment showed exponential depletion of the $[AAY(OCH_3)R]^{+\bullet}$ intensity (**Figure 2.22, SI**), indicating that all the components of the ion population absorbed light at 355 nm. We note that the absorption spectra of peptide cation-radicals depend on the position of the Cα radical and ion conformation,[80,102] and elucidating this effect would require a dedicated study.

2.3.3. *Discussion*

The experimental data using CID and UVPD showed distinct features that can be related to the nature of the peptide cation-radicals and their mode of activation. It is useful to first reiterate the differences in the activation modes because they apply to both hydrogen-rich and -deficient peptide cation-radicals. Collision-induced dissociation in the ion trap is performed by resonant translational excitation of the mass-selected ion under slow heating conditions,[17] which upon collisions with the bath gas (He at 3 mTorr in the LTQ), results in vibrational excitation of the ground electronic state of the precursor cation-radical. The fragment ions are not accelerated and their collisions with the bath gas result in cooling that occurs on the 10-20 ms time scale.[193] In contrast, UVPD is a "sudden" excitation method that generates an electronically excited state of the peptide ion with an internal energy (E_{int}) given by the sum of the photon energy (3.493 eV, 337 kJ mol^{-1}) and the thermal rovibrational energy (H_{rovib}) of the ion at the ion trap temperature. For $[AAAYR + 2H]^{+\bullet}$ and related ions, H_{rovib} has a mean value of 105 kJ mol^{-1} at 310 K, giving Eint = 337 + 105 = 442 kJ mol^{-1}. The photoexcited peptide ion can dissociate on the same potential energy surface, utilizing the electronic part of the excitation energy, or undergo vibronic conversion to a vibrationally excited ground electronic state with E_{int}. Both these processes compete with collisional de-excitation of the excited electronic state and vibrational cooling of the hot ground state.

Starting with the hydrogen-rich cation-radicals, both the CID and UVPD spectra indicated that the stable, long-lived, cation radicals were distinctly different from the reactive intermediates produced by electron attachment to peptide dications. This raises the question of the cation-radical structure and formation upon ETD of the CE complexes. Crown ethers bind to charged groups by forming strong hydrogen bonds in gas-phase ions.[194] The ETD spectra of the doubly charged peptide-CE complexes indicate rather non-specific binding to the N-terminal ammonium and Arg guanidinium charged groups. This finding is at odds with the bonding energies of 18-crown-6-ether to singly protonated amino acids, reported by Chen and Rodgers,[195] where N-terminal bonding was found to be stronger than side-chain bonding in arginine. The probable reason for this apparent discrepancy is that the competitive binding of 18-crown-6-ether to a charged group in a doubly protonated peptide disrupts this group's internal solvation and this loss of attractive interactions can compensate for the stronger binding to the ligand. In addition, disruption of internal solvation of the charged group by the peptide amide groups affects the ion conformation.[102]

Another large effect, which is related to electron transfer, is that of coordination with CE, which substantially lowers the intrinsic recombination energy of the charged group; for example, from 4.31 eV in isolated $CH_3NH_3^+$ to 1.74 eV in the $CH_3NH_3^+$-CE complex.[196] The calculated recombination energy of the ethylguanidinium cation also showed a drop from 380 kJ mol-1 (3.94 eV) in the free ion to 231 kJ mol-1 (2.40 eV) in the CE-complex (**Table 2.3, SI**), illustrating the effect of CE-coordination on the electronic structure of the cation and radical.

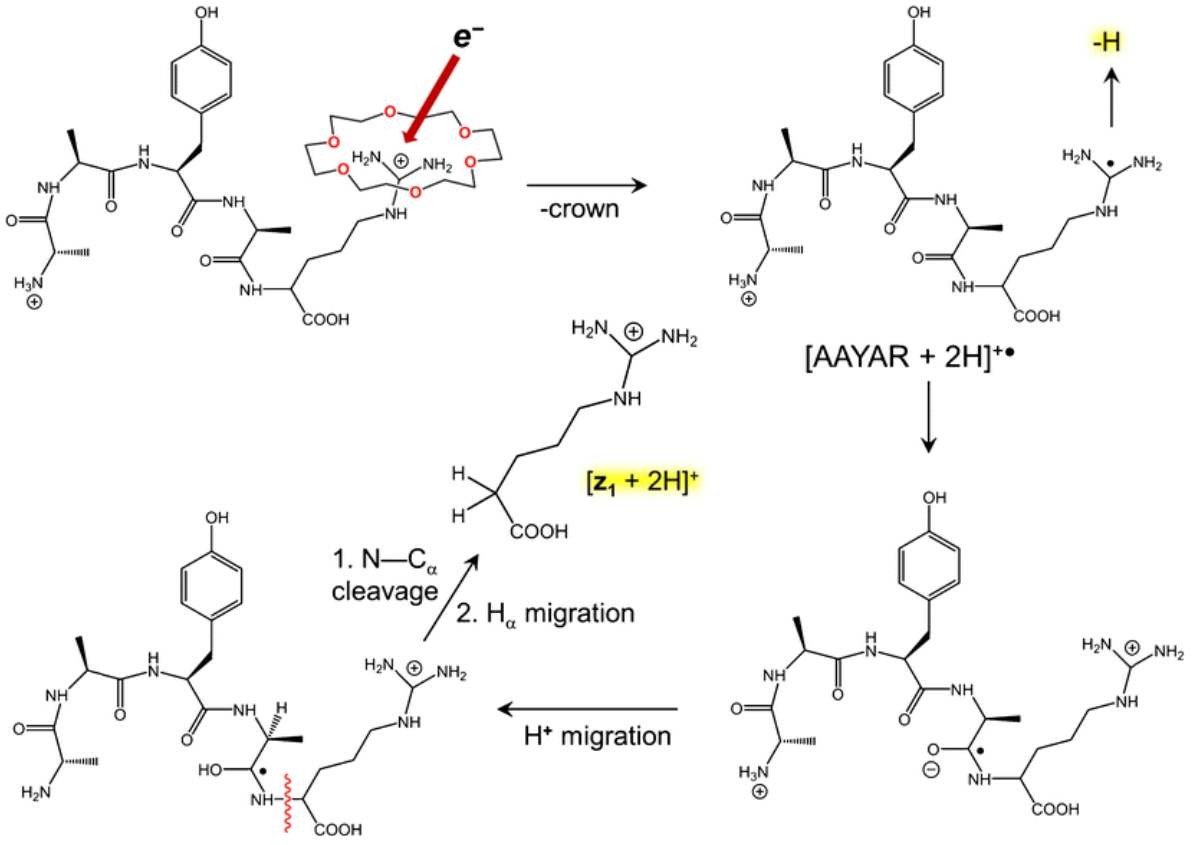
Charge reduction by electron attachment to the charged group substantially lowers its CE-bonding energy; for example, from 273 kJ mol^{-1} in the $CH3NH3^+$-CE complex to 35 kJ mol-1 in its charge-reduced counterpart.[196] Similar effects were indicated by calculations of N-ethylguanidinium ion and radical CE complexes that stand for the Arg-CE bonding (**Table 2.3, SI**). The 0 K bonding energy to CE was calculated to decrease from $\Delta H_{0,diss}$ = 197 kJ mol^{-1} in the ion to 56 kJ mol^{-1} in the radical. The calculated CE-bonding free energy for ethylguanidinium radical at the experimental temperature of 310 K, $\Delta G_{310,diss}$ = 16 kJ mol^{-1}, indicated that binding of the CE ligand to the charge-reduced Arg guanidinium as well as N-terminal ammonium functional groups is weak, and hence electron transfer targeting the coordinated charged group is expected to result in a facile loss of the CE ligand.

The peptide cation radicals resulting from electron attachment and CE loss have the odd electron in the charge-reduced group, which is an ammonium radical at the N-terminus or a guanidinium radical in the Arg side chain. This electron distribution is fundamentally different from that accessed by direct electron attachment to the isolated peptide dication, where the reactive electronic states are charge-stabilized amide π* states.[56,197] This interpretation is consistent with the stark differences in the peptide cation-radical reactivity depending on their mode of formation. The isomeric charge-reduced cation-radicals from peptide-CE complexes also have different stabilities, whereby the ammonium radicals are only weakly bound and can rapidly dissociate by competing losses of H and ammonia.[198] This may account for the abundant $[M + 2H - CE - NH_3]^{+\bullet}$ fragment ions in the ETD spectra of the complexes (m/z 535, **Figure 2.8a-d, SI**). In contrast, arginine guanidinium radicals are intrinsically stable.[192,199] Vibrational excitation, such as in CID, results in competitive loss of H and guanidine,[200] accounting for the pertinent fragment ions, m/z 551 and m/z 493, respectively, in the CID spectra (**Figure 2.1**). Arginine radicals formed by electron transfer have been shown to be moderately efficient hydrogen atom acceptors.[199,201] This may account for the formation of the $[z1 + 2H]^+$ fragment ions, as sketched for $[AAYAR + 2H]^{+\bullet}$ in **Scheme 2.1**, as well as loss of C_7H_6O which is triggered by transfer of the phenol hydrogen. Detailed mechanisms of the hydrogen transfer and bond cleavage steps leading to these ions have not been established yet.

Guanidinium radicals are strong chromophores absorbing light at 320, 341, and 357 nm, with the corresponding oscillator strength factors of 0.07, 0.18, and 0.01.[78] This is consistent with the efficient photodepletion of the charge-reduced hydrogen-rich cation radicals at 355 nm (**Figure 2.10a-d, SI**). Excited states of arginine radicals have been shown to undergo extremely fast loss of hydrogen,[200] which is consistent with the dominant photodissociation channels in the UVPD spectra.

The 6-23% fractions of photo-inactive hydrogen-rich peptide cation-radicals (**Figure 2.10a-d, SI**) belong to ion structures that are difficult to directly assign and likely differ for the sequence isomers. Based on the analysis of peptide radical chromophores,[78] dihydrophenol radicals produced by H-atom migration to the Tyr side chain, and $[a + x]^{+\bullet}$ ion-molecule complexes are possible candidates for structures not absorbing at 355 nm.

The hydrogen-deficient $[M]^{+\bullet}$ cation radicals show collision-induced and near-UV dissociations that depend on both the activation mode and position of the Tyr residue in the peptide sequence. CID and UVPD of $[YAAAR]^{+\bullet}$ are most straightforward in that they both result in a dominant loss of the Tyr side chain. This is consistent with the Tyr-O radical structure of this ion where loss of C_7H_6O is the lowest energy process in CID and the Tyr-O radical is the absorbing chromophore for UVPD.[80] We cannot distinguish whether UVPD occurs from an excited electronic state or after vibronic transition to a vibrationally hot ground electronic state because both are expected to display similar reactivity. In contrast, the other peptide sequences show substantial differences between CID and UVPD, indicating large differences in the excited-state and ground-state reactivity.



Scheme 2.1. Proposed reaction sequence for the loss of H and formation of [z1 + 2H]+ fragment ions.

With $[AAAYR]^{+\bullet}$, CID induces hydrogen atom migrations forming intermediates that dissociate by backbone cleavages and loss of C_7H_6O, forming the various fragment ions seen in the spectrum. In contrast, UVPD proceeds by only two major channels, one being loss of protogenic exchangeable H and the other being a backbone cleavage that leads to the $[y4 + 2H]^+$ ion. The latter reaction can be interpreted as proceeding from a hot vibrational state accessed by vibronic transition from the excited electronic state. In contrast, the loss of H does not have an analogy in CID and is likely proceeding from an excited electronic state of the ion. A previous UV photodissociation action spectroscopy study has concluded that $[AAAYR]^{+\bullet}$ was a mixture of isomeric radicals,[80] and thus, it is possible that the observed photodissociations originated from different radical isomers.

2.4 Conclusions

The results reported in this comparative study allow us to arrive at the following conclusions. Both the hydrogen-rich peptide cation radicals produced by electron-transfer reduction, and their hydrogen-deficient counterparts produced by electron-transfer oxidation contain chromophores associated with the radical moieties that result in photodissociation at 355 nm. The hydrogen-rich peptide cation-radicals contain major fractions of structures with Arg guanidinium radical groups. Consequently, both collision-induced dissociation and photodissociation of the long-lived hydrogen-rich cation radicals are diametrically different from electron-transfer dissociation of the corresponding peptide dications. CID and UVPD of hydrogen-deficient peptide cation-radicals shows dependence on the position on the Tyr residue. Ions in which the Tyr residue is in sequence positions remote from the charged Arg group show larger proportions of Tyr-O radicals that undergo collision-induced and photodissociative loss of C_7H_6O from the Tyr side chain.

2.5 Acknowledgments

Research at University of Washington received support from the Chemistry Division of the National Science Foundation (Grant CHE-1359810). I.K.C. thanks support from the Hong Kong Research Grants Council (project nos. HKU 701613P and 173306015), the University of Hong Kong (Seed Funding Programme for Basic Research 201411159067 and 201511159023); X.Y.M. thanks the Hong Kong RGC for supporting studentship. F.T. thanks the Royal Society Kan Tong Po Professorship for supporting his stay at the University of Hong Kong in September-December 2014. Support from the Klaus and Mary Ann Saegebarth Endowment is gratefully appreciated.

2.6 Supplemental Information (SI)

Table 2.1. Relative Intensities of CE-Coordinated ETD Fragment Ions from Peptide-CE Complexes.

Ion m/z	Assignment[a]	Relative Intensity[b]
AAAYR		
N-terminal Fragment Ions		
282	$[CE + NH_4]^+$	9.7
353	$[\mathbf{c_1}+2H+CE]^+$	0.1
424	$[\mathbf{c_2}+2H+CE]^+$	0.5
495	$[\mathbf{c_3}+2H+CE]^+$	0.8
658	$[\mathbf{c_4}+2H+CE]^+$	8.9
757	$[M + 2H + CE - CH_5N_3]^+$	5.8
773	$[M + 2H + CE - CH_3N_2]^+$	18.0
C-terminal Fragment Ions		
423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	17.1
586	$[\mathbf{z_2}+H+CE]^{+\bullet}$	9.2
657	$[\mathbf{z_3}+H+CE]^{+\bullet}$	20.9
728	$[\mathbf{z_4}+H+CE]^{+\bullet}$	9.1
AAYAR		
N-terminal Fragment Ions		
282	$[CE + NH_4]^+$	13.0
353	$[\mathbf{c_1}+2H+CE]^+$	0.0
424	$[\mathbf{c_2}+2H+CE]^+$	0.7
587	$[\mathbf{c_3}+2H+CE]^+$	2.3
658	$[\mathbf{c_4}+2H+CE]^+$	9.4
757	$[M + 2H + CE - CH_5N_3]^+$	9.2
773	$[M + 2H + CE - CH_3N_2]^+$	21.6
C-terminal Fragment Ions		
423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	10.5
494	$[\mathbf{z_2}+H+CE]^{+\bullet}$	8.9
657	$[\mathbf{z_3}+H+CE]^{+\bullet}$	13.8
728	$[\mathbf{z_4}+H+CE]^{+\bullet}$	10.6

AYAAR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	8.6
353	$[\mathbf{c_1}+2H+CE]^+$	0.1
516	$[\mathbf{c_2}+2H+CE]^+$	0.8
587	$[\mathbf{c_3}+2H+CE]^+$	1.6
658	$[\mathbf{c_4}+2H+CE]^+$	8.5
757	$[M + 2H + CE - CH_5N_3]^+$	7.0
773	$[M + 2H + CE - CH_3N_2]^+$	16.7

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	11.0
494	$[\mathbf{z_2}+H+CE]^{+\bullet}$	9.2
565	$[\mathbf{z_3}+H+CE]^{+\bullet}$	27.9
728	$[\mathbf{z_4}+H+CE]^{+\bullet}$	8.4

YAAAR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	5.1
445	$[\mathbf{c_1}+2H+CE]^+$	0.3
516	$[\mathbf{c_2}+2H+CE]^+$	0.5
587	$[\mathbf{c_3}+2H+CE]^+$	0.7
658	$[\mathbf{c_4}+2H+CE]^+$	8.9
757	$[M + 2H + CE - CH_5N_3]^+$	5.4
773	$[M + 2H + CE - CH_3N_2]^+$	11.7

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	7.1
494	$[\mathbf{z_2}+H+CE]^{+\bullet}$	6.8
565	$[\mathbf{z_3}+H+CE]^{+\bullet}$	22.8
636	$[\mathbf{z_4}+H+CE]^{+\bullet}$	30.6

AVAYR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	9.1
353	$[\mathbf{c_1}+2H+CE]^+$	0.1
452	$[\mathbf{c_2}+2H+CE]^+$	0.2
523	$[\mathbf{c_3}+2H+CE]^+$	1.1
686	$[\mathbf{c_4}+2H+CE]^+$	9.8
785	$[M + 2H + CE - CH_5N_3]^+$	7.0
801	$[M + 2H + CE - CH_3N_2]^+$	12.9

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	17.2
586	$[\mathbf{z_2}+H+CE]^{+\bullet}$	9.5
657	$[\mathbf{z_3}+H+CE]^{+\bullet}$	25.9
756	$[\mathbf{z_4}+H+CE]^{+\bullet}$	7.2

AAYVR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	11.3
353	$[\mathbf{c_1}+2H+CE]^+$	0.1
424	$[\mathbf{c_2}+2H+CE]^+$	0.7
587	$[\mathbf{c_3}+2H+CE]^+$	1.0
686	$[\mathbf{c_4}+2H+CE]^+$	9.9
785	$[M + 2H + CE - CH_5N_3]^+$	6.9
801	$[M + 2H + CE - CH_3N_2]^+$	21.4

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	12.9
522	$[\mathbf{z_2}+H+CE]^{+\bullet}$	6.1
685	$[\mathbf{z_3}+H+CE]^{+\bullet}$	15.4
756	$[\mathbf{z_4}+H+CE]^{+\bullet}$	14.2

AYAVR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	7.5
353	$[\mathbf{c_1}+2H+CE]^+$	0.1
516	$[\mathbf{c_2}+2H+CE]^+$	0.6
587	$[\mathbf{c_3}+2H+CE]^+$	0.8
686	$[\mathbf{c_4}+2H+CE]^+$	6.6
785	$[M + 2H + CE - CH_5N_3]^+$	5.4
801	$[M + 2H + CE - CH_3N_2]^+$	16.5

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	11.0
522	$[\mathbf{z_2}+H+CE]^{+\bullet}$	5.5
593	$[\mathbf{z_3}+H+CE]^{+\bullet}$	32.4
756	$[\mathbf{z_4}+H+CE]^{+\bullet}$	13.5

YAAVR

N-terminal Fragment Ions

282	$[CE + NH_4]^+$	4.3
445	$[\mathbf{c_1}+2H+CE]^+$	0.3
516	$[\mathbf{c_2}+2H+CE]^+$	0.5
587	$[\mathbf{c_3}+2H+CE]^+$	0.0
686	$[\mathbf{c_4}+2H+CE]^+$	7.2
785	$[M + 2H + CE - CH_5N_3]^+$	3.5
801	$[M + 2H + CE - CH_3N_2]^+$	11.1

C-terminal Fragment Ions

423	$[\mathbf{z_1}+H+CE]^{+\bullet}$	8.7
522	$[\mathbf{z_2}+H+CE]^{+\bullet}$	4.1
593	$[\mathbf{z_3}+H+CE]^{+\bullet}$	23.9
664	$[\mathbf{z_4}+H+CE]^{+\bullet}$	36.2

---

[a]See the ion nomenclature in the main text.

[b]Relative to the sum of these fragment ion intensities.

Table 2.2. Accurate Mass Measurements on a LTQ_Orbitrap at 100,000 resolving power.

Ion *m/z* Calculated	Elemental Composition	Neutral Loss		Neutral Error (millim.u.) Assignment	
AAAYR					
550.2846	550.2858	$C_{24}H_{38}N_8O_7$			
480.2554	480.2565	$C_{21}H_{34}N_7O_6$	70.0292	C_3H_4NO	−0.1
478.2397	478.2397	$C_{21}H_{32}N_7O_6$	72.0449	C_3H_6NO	−0.0
464.2368	464.2378	$C_{21}H_{32}N_6O_6$	86.0478	$C_3H_6N_2O$	−0.2
444.2430	444.2439	$C_{17}H_{32}N_8O_6$	106.0416	C_7H_6O	−0.3
407.2030	407.2037	$C_{18}H_{27}N_6O_5$	143.0816	$C_6H_{11}N_2O_2$	−0.5
392.1922	392.1928	$C_{18}H_{26}N_5O_5$	158.0924	$C_6H_{12}N_3O_2$	−0.6
321.1553	321.1557	$C_{15}H_{21}N_4O4$	229.1293	$C_9H_{17}N_4O_3$	−0.8
AAYAR					
550.2814	550.2858	$C_{24}H_{38}N_8O_7$			
480.2527	480.2565	$C_{21}H_{34}N_7O_6$	70.0287	C_3H_4NO	−0.6
464.2342	464.2378	$C_{21}H_{32}N_6O_6$	86.0472	$C_3H_6N_2O$	−0.8
444.2405	444.2439	$C_{17}H_{32}N_8O_6$	106.0409	C_7H_6O	−1.0
392.1901	392.1928	$C_{18}H_{26}N_5O_5$	158.0913	$C_6H_{12}N_3O_2$	−1.6
230.1357	230.1373	$C_9H_{18}N_4O_3$	320.1457	$C_{15}H_{20}N_4O_4$	−2.8
AYAAR					
550.2849	550.2858	$C_{24}H_{38}N_8O_7$			
444.2433	444.2439	$C_{17}H_{32}N_8O_6$	106.0416	C_7H_6O	−0.3
YAAAR					
550.2853	550.2858	$C_{24}H_{38}N_8O_7$			
444.2436	444.2439	$C_{17}H_{32}N_8O_6$	106.0417	C_7H_6O	−0.2

Table 2.3. Relative Energies

	Relative Energy[a,b]			
Species/Reaction	B3LYP[c]	M06-2X[c]	ωB97X-D[c]	MP2[c,d]
[*N*-Ethylguanidinium + CE]$^+$	0	0	0	0
N-Ethylguanidinium$^+$ + CE	158	202	199	205
[*N*-Ethylguanidinium + CE]$^\bullet$	0	0	0	0
N-Ethylguanidinium$^\bullet$ + CE	9.4 (−31)[e]	54 (13) [e]	56 (16)[e]	57 (17)[e]
IE(*N*-Ethylguanidinium$^\bullet$)	393 (4.07)[f]	380 (3.94)[f]	385 (3.99)[f]	380 (3.94)[f]
IE([*N*-Ethylguanidinium + CE]$^\bullet$)	244 (2.52)[f]	237 (2.46)[f]	237 (2.45)[f]	217 (2.24)[f]

[a]Energies in kJ mol^{-1}. [b]Including B3LYP/6-31+G(d,p) zero-point vibrational energies and referring to 0 K. [c]From single-point energy calculations with the 6-311++G(2d,p) basis set on DFT/6-31+G(d,p) fully optimized geometries. [d]Spin projected energies for radicals. [e]Relative free energies at 310 K. [f]Adiabatic ionization energies in eV.

Figure 2.7. ETD mass spectra (fluoranthene, 200 ms ion-ion reaction time) of (a) $[AAAYR + 2H]^{2+}$ (m/z 276), (b) $[AAYAR + 2H]^{2+}$, (c) $[AYAAR + 2H]^{2+}$, and (d) $[YAAAR + 2H]^{2+}$ The fragment ion assignment follows the all-inclusive nomenclature (ref [40] in main text).

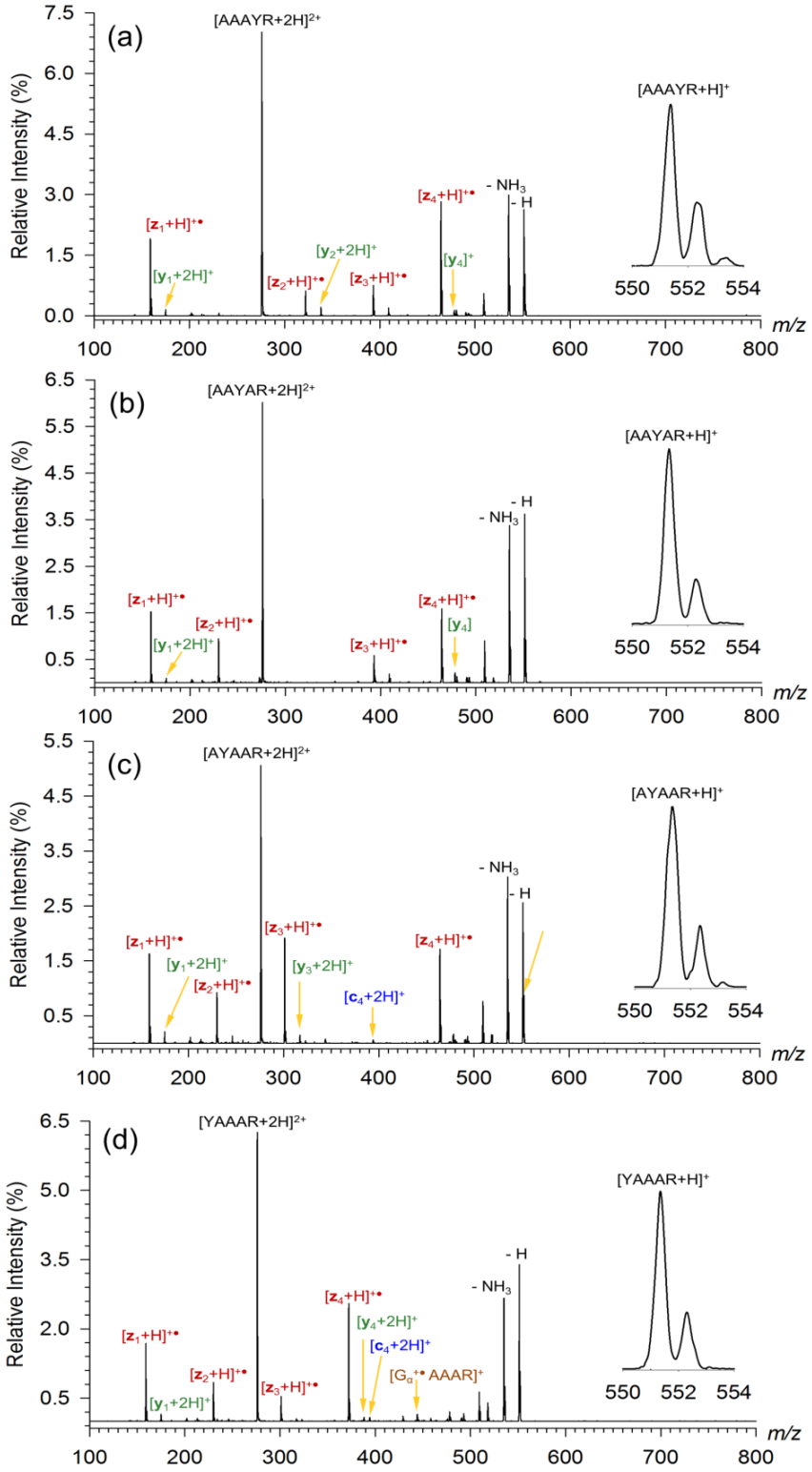


Figure 2.8. ETD mass spectra (fluoranthene, 200 ms ion-ion reaction time) of (a) $[AAAYR + CE + 2H]^{2+}$ (m/z 408), (b) $[AAYAR + CE + 2H]^{2+}$, (c), $[AYAAR + CE + 2H]^{2+}$, and (d) $[YAAAR + CE + 2H]^{2+}$. The fragment ion assignment follows the all-inclusive nomenclature.

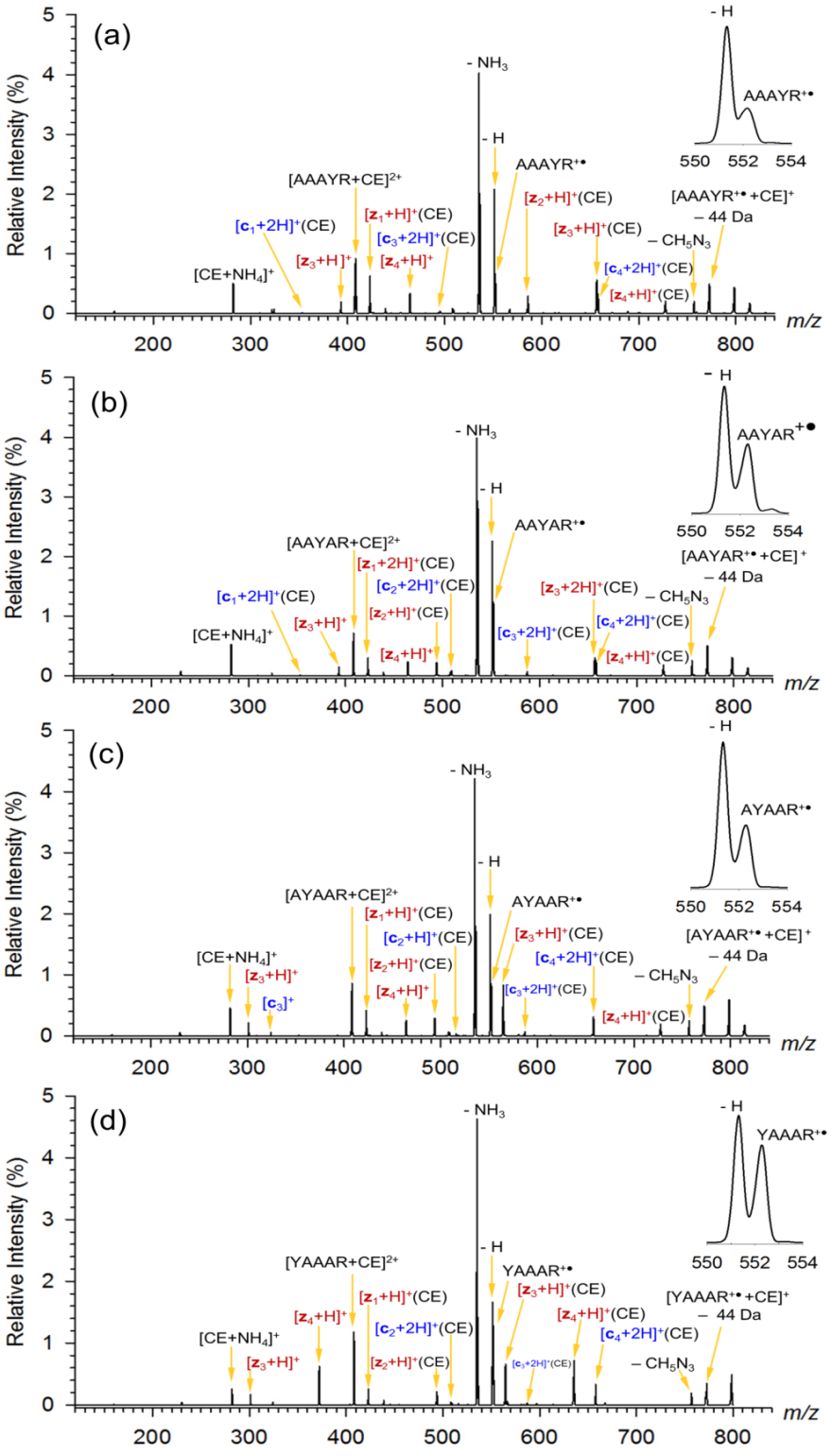


Figure 2.9. CID mass spectra of cation-radicals generated by ETD of crown-ether complexes: (a) [AVAYR + 2H]$^{+\bullet}$ (m/z 580), (b) [AAYVR + 2H]$^{+\bullet}$, (c) [AYAVR + 2H]$^{+\bullet}$, and (d) [YAAVR + 2H]$^{+\bullet}$.

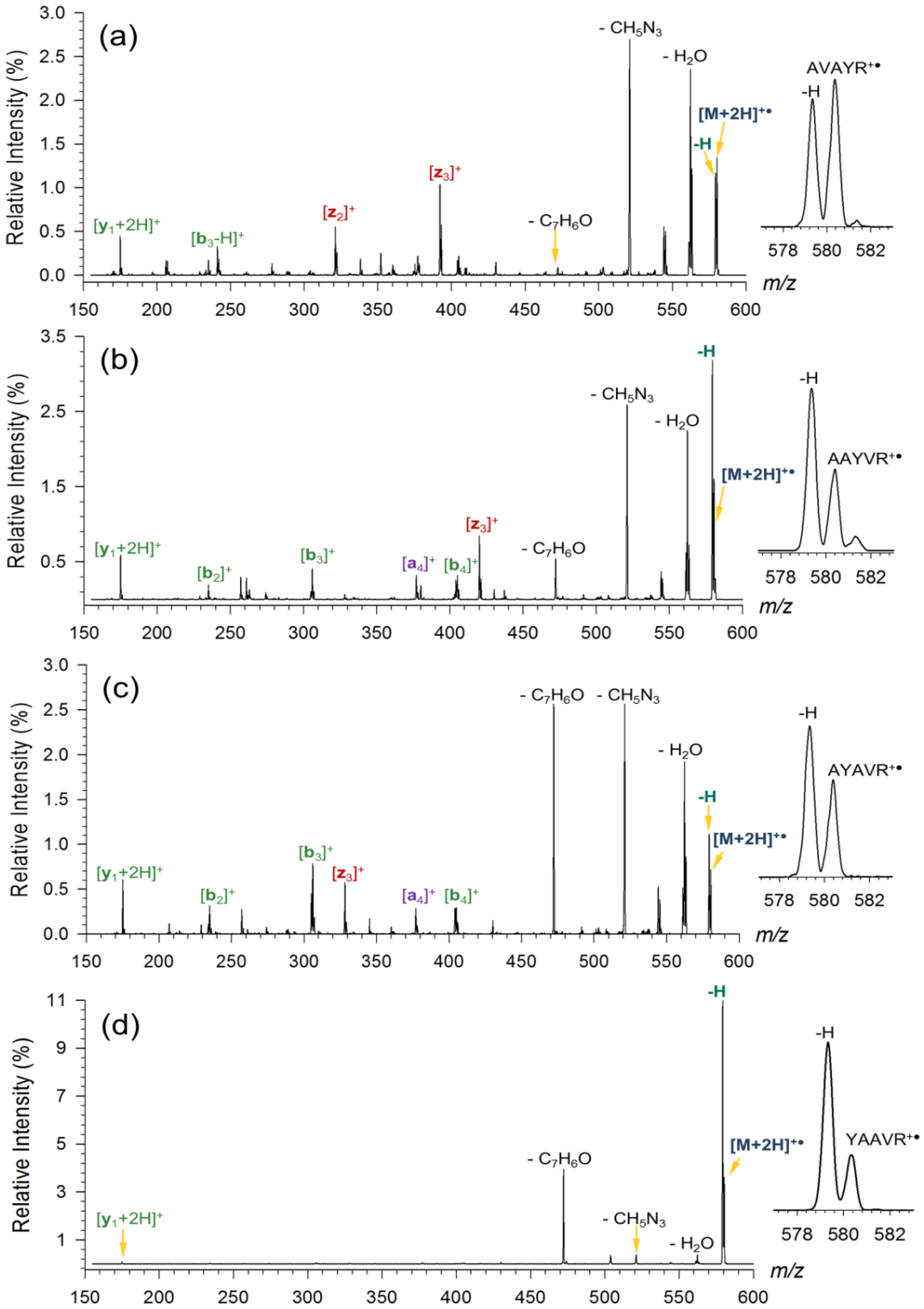


Figure 2.10. Photodepletion curves at 355 nm as a number of laser pulses of (a) $[AAAYR + 2H]^{+\bullet}$, (b) $[AAYAR + 2H]^{+\bullet}$, (c) $[AYAAR + 2H]^{+\bullet}$, (d) $[YAAAR + 2H]^{+\bullet}$.

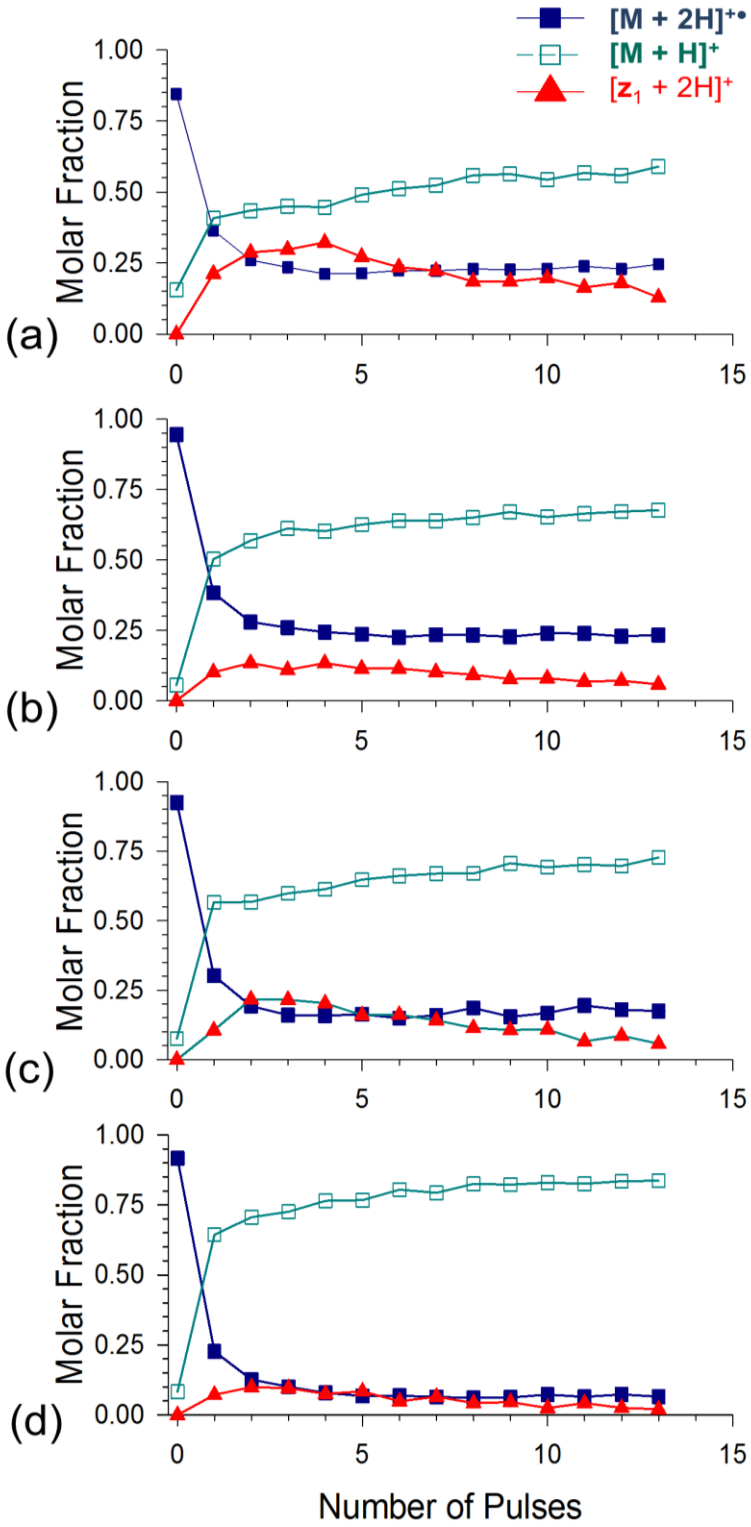


Figure 2.11. UVPD mass spectra of cation-radicals generated by ETD of crown-ether complexes: (a) [AVAYR + 2H]$^{+\bullet}$ (m/z 580), (b) [AAYVR + 2H]$^{+\bullet}$, (c) [AYAVR + 2H]$^{+\bullet}$, and (d) [YAAVR + 2H]$^{+\bullet}$.

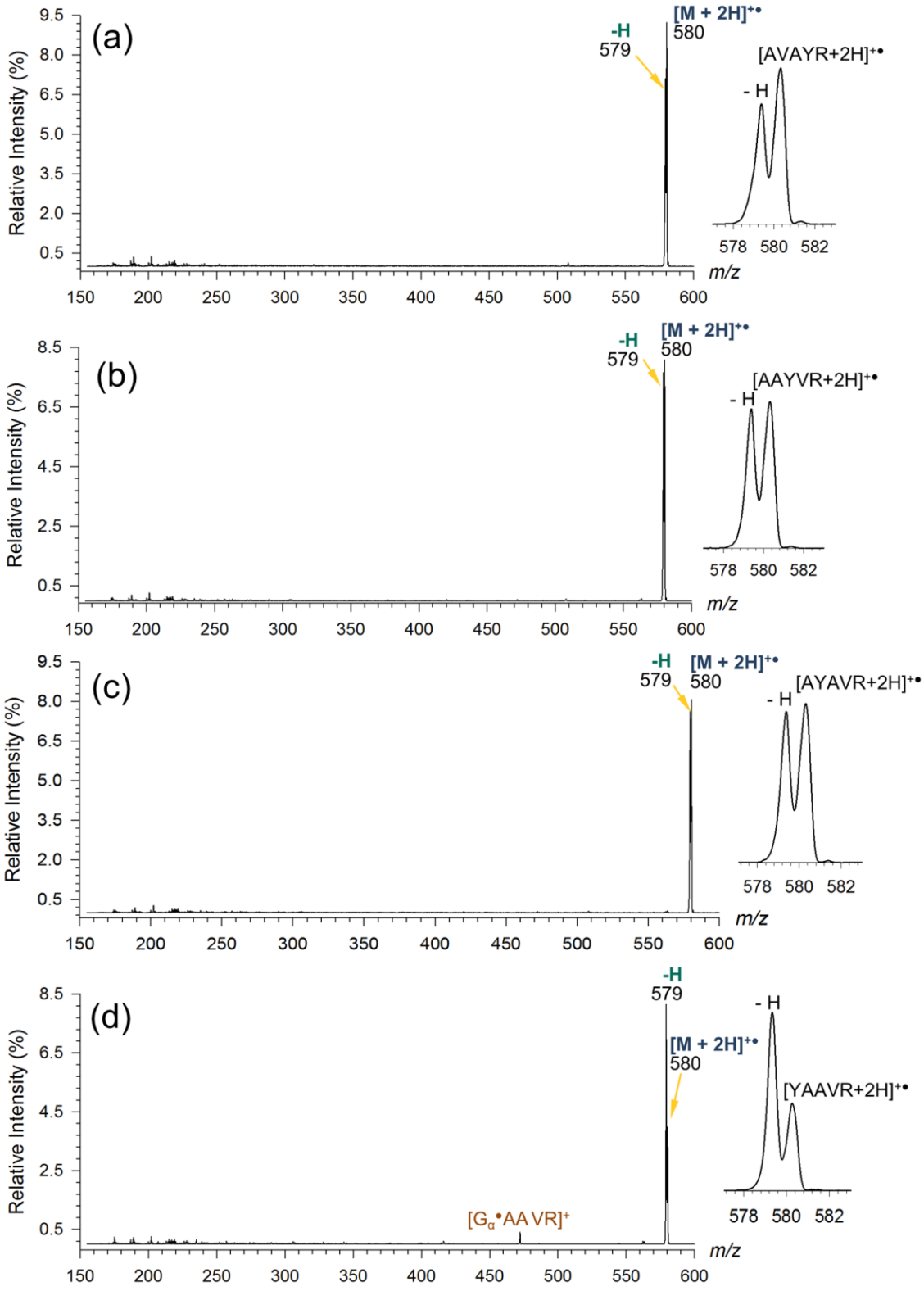


Figure 2.12. Photodepletion curves at 355 nm as a number of laser pulses of (a) $[AVAYR + 2H]^{+\bullet}$, (b) $[AAYVR + 2H]^{+\bullet}$, (c) $[AYAVR + 2H]^{+\bullet}$, (d) $[YAAVR + 2H]^{+\bullet}$.

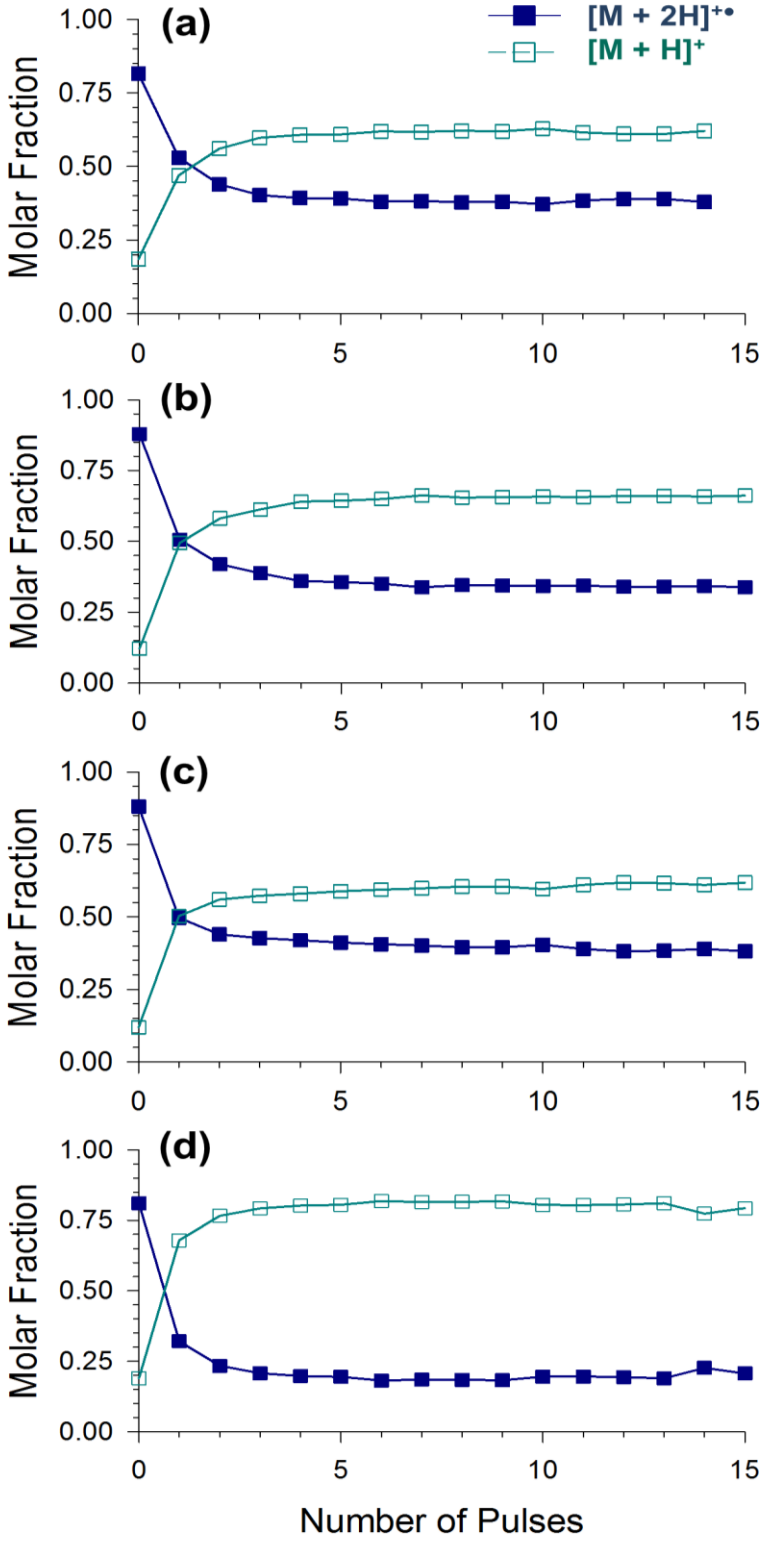


Figure 2.13. CID mass spectra of (a)[AAAYR]$^{+\bullet}$ (m/z 550, ELAB = 20 eV)); (b) [AAYAR]$^{+\bullet}$ (ELAB = 15 eV); (c) [AYAAR]$^{+\bullet}$ (ELAB = 15 eV); (d) [YAAAR]$^{+\bullet}$ (ELAB = 15 eV).

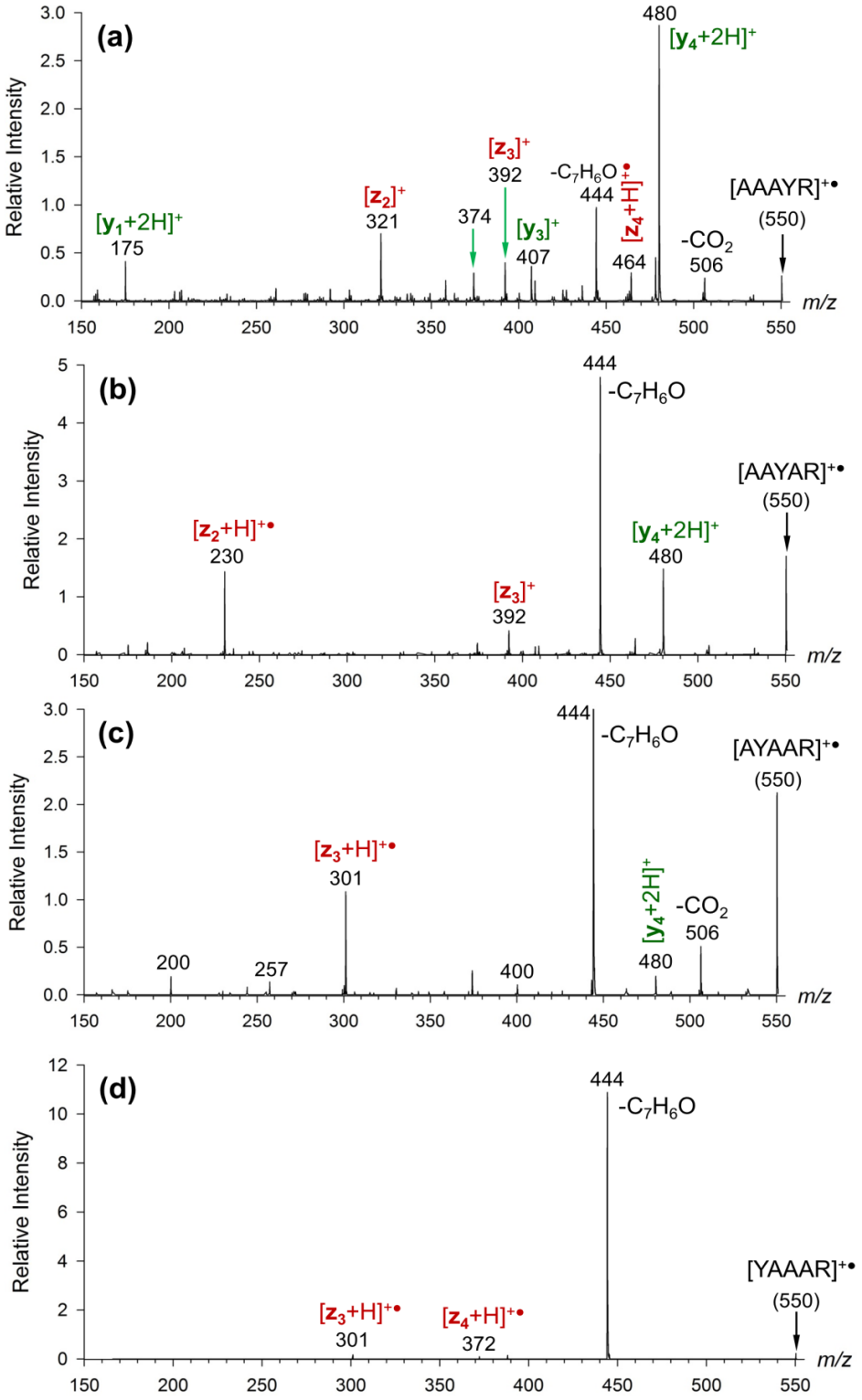


Figure 2.14. LTQ-CID-MS3 spectra of (m/z 578) (a) $[AVAYR]^{+\bullet}$, (b) $[AAYVR]^{+\bullet}$,(c) $[AYAVR]^{+\bullet}$, and (d) $[YAAVR]^{+\bullet}$.

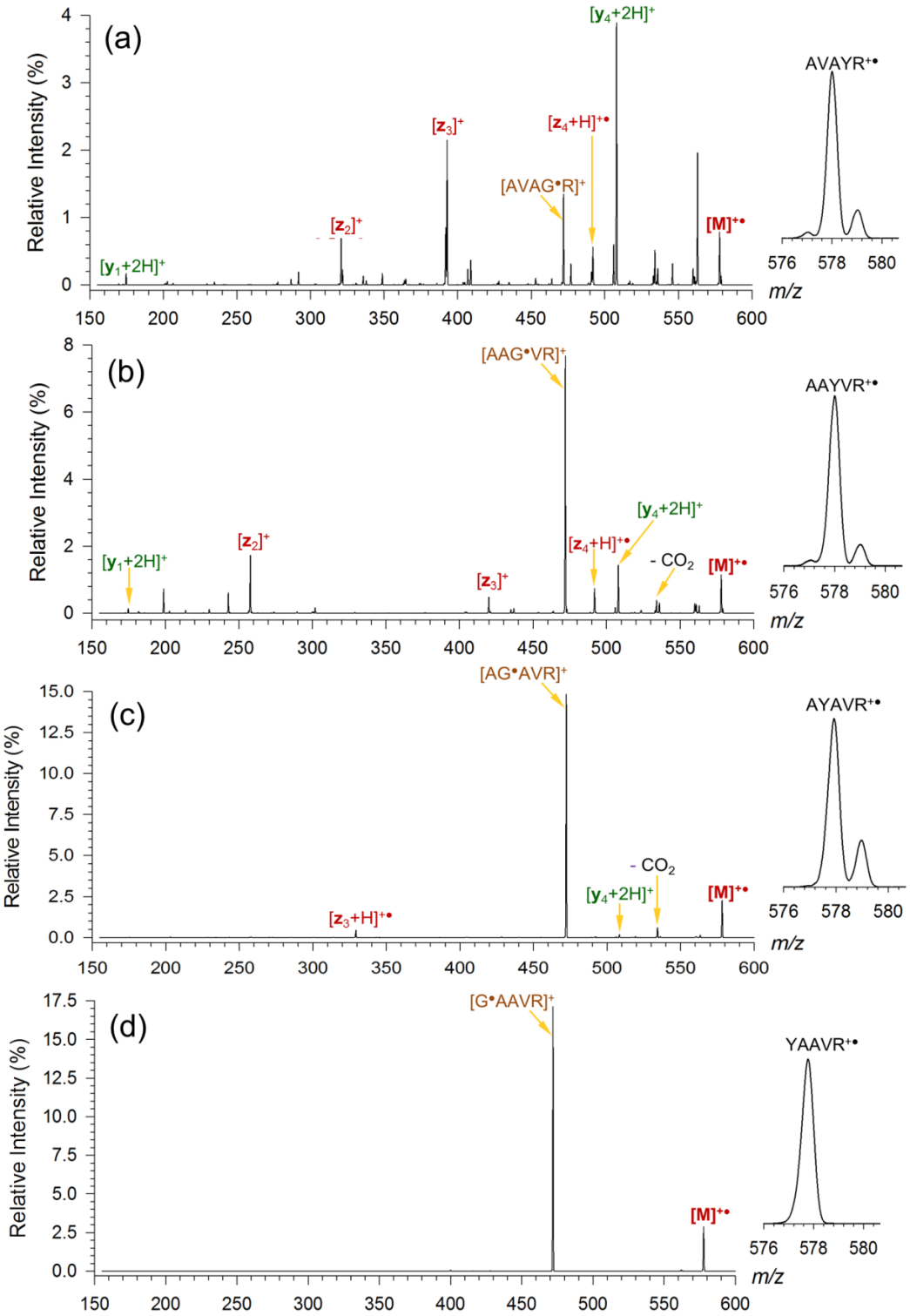


Figure 2.15. UVPD-MS3 spectra of (m/z 578) (a) $[AVAYR]^{+\bullet}$, (b) $[AAYVR]^{+\bullet}$, (c)$[AYAVR]^{+\bullet}$, and (d) $[YAAVR]^{+\bullet}$.

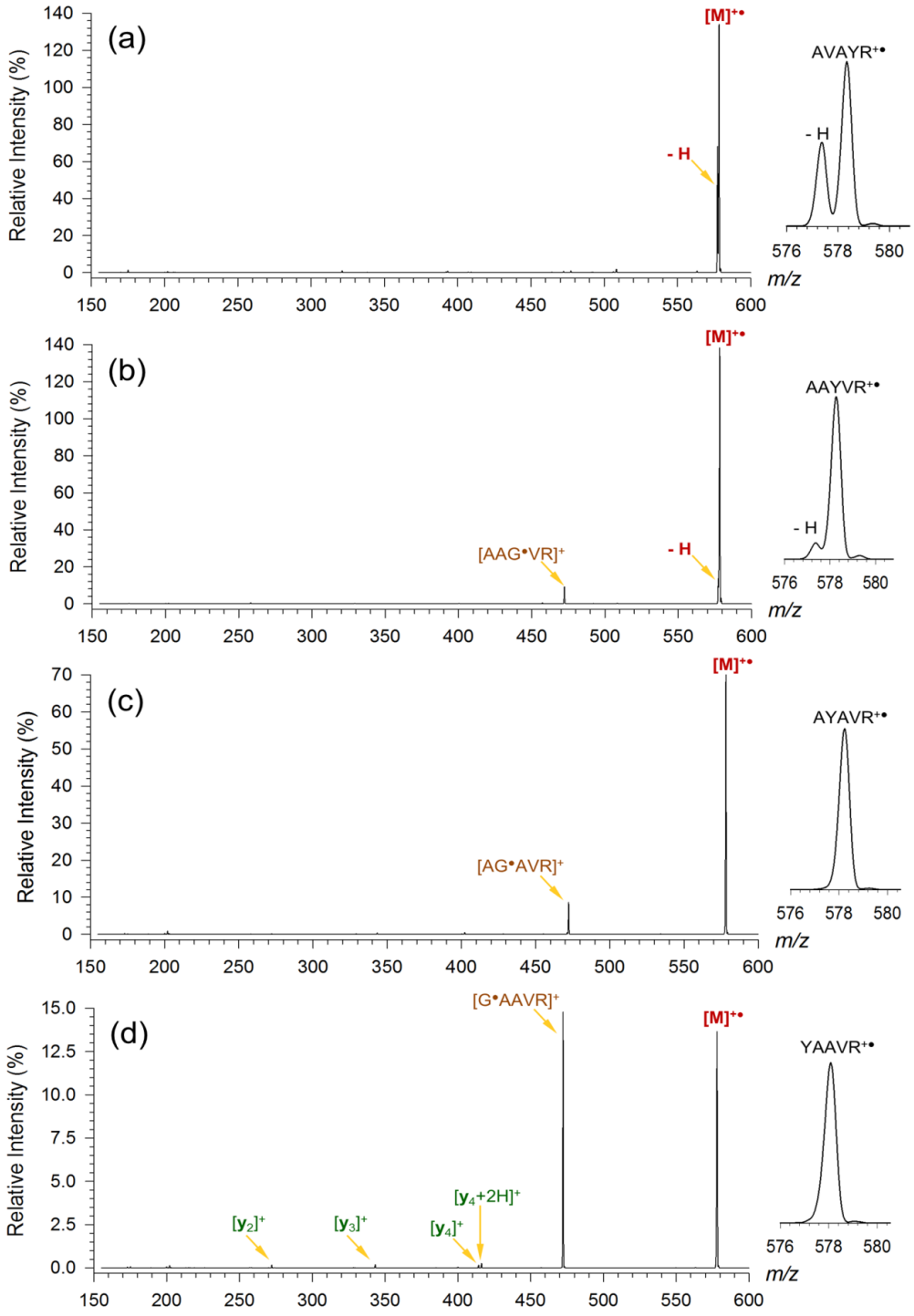


Figure 2.16. Photodepletion curves at 355 nm as a number of laser pulses of (a) $[AAAYR]^{+\bullet}$, (b) $[AAYAR]^{+\bullet}$, (c) $[AYAAR]^{+\bullet}$, (d) $[YAAAR]^{+\bullet}$.

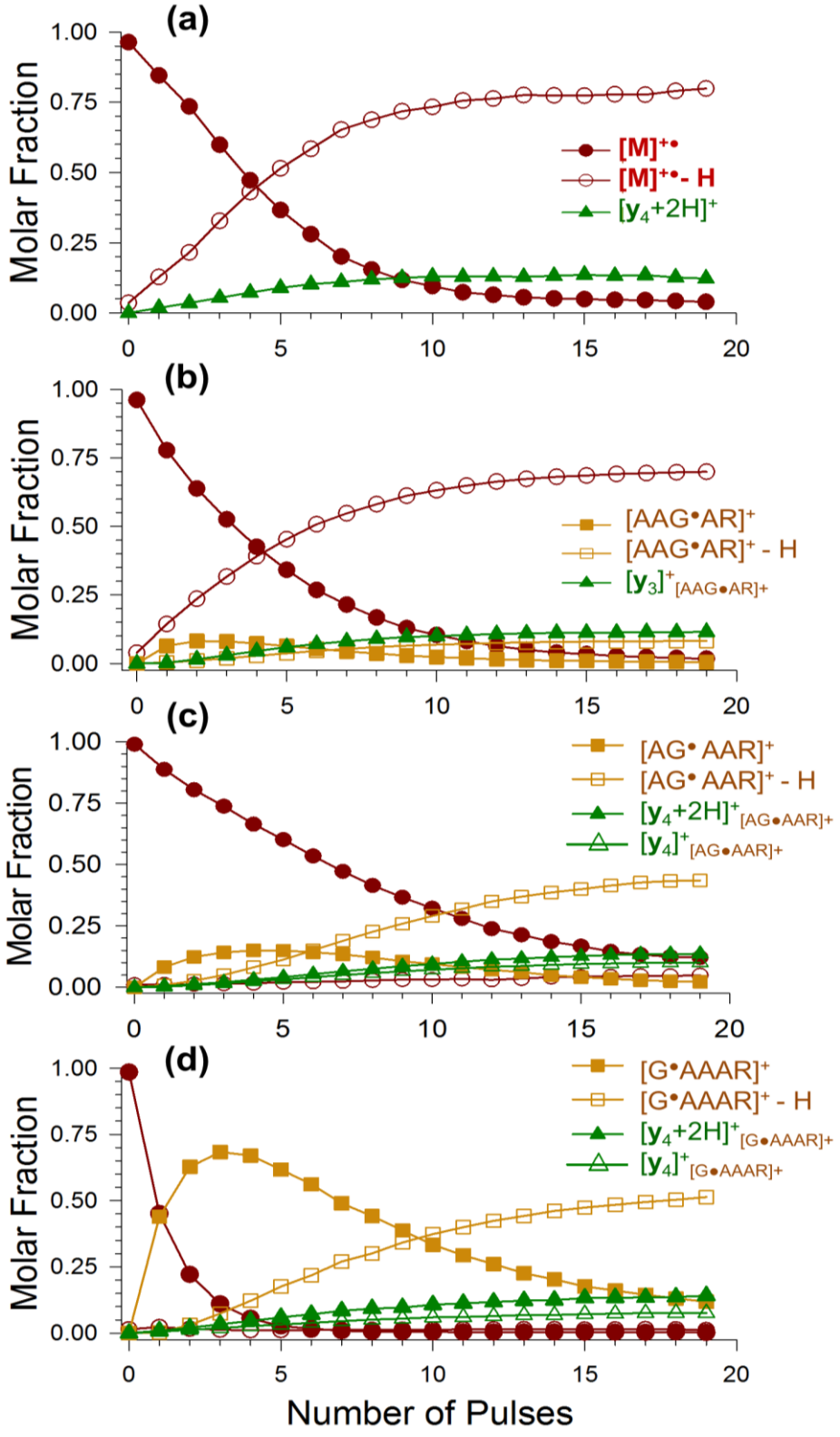


Figure 2.17. Photodepletion curves at 355 nm as a number of laser pulses of (a) $[AVAYR]^{+\bullet}$, (b) $[AAYVR]^{+\bullet}$, (c) $[AYAVR]^{+\bullet}$, (d) $[YAAVR]^{+\bullet}$.

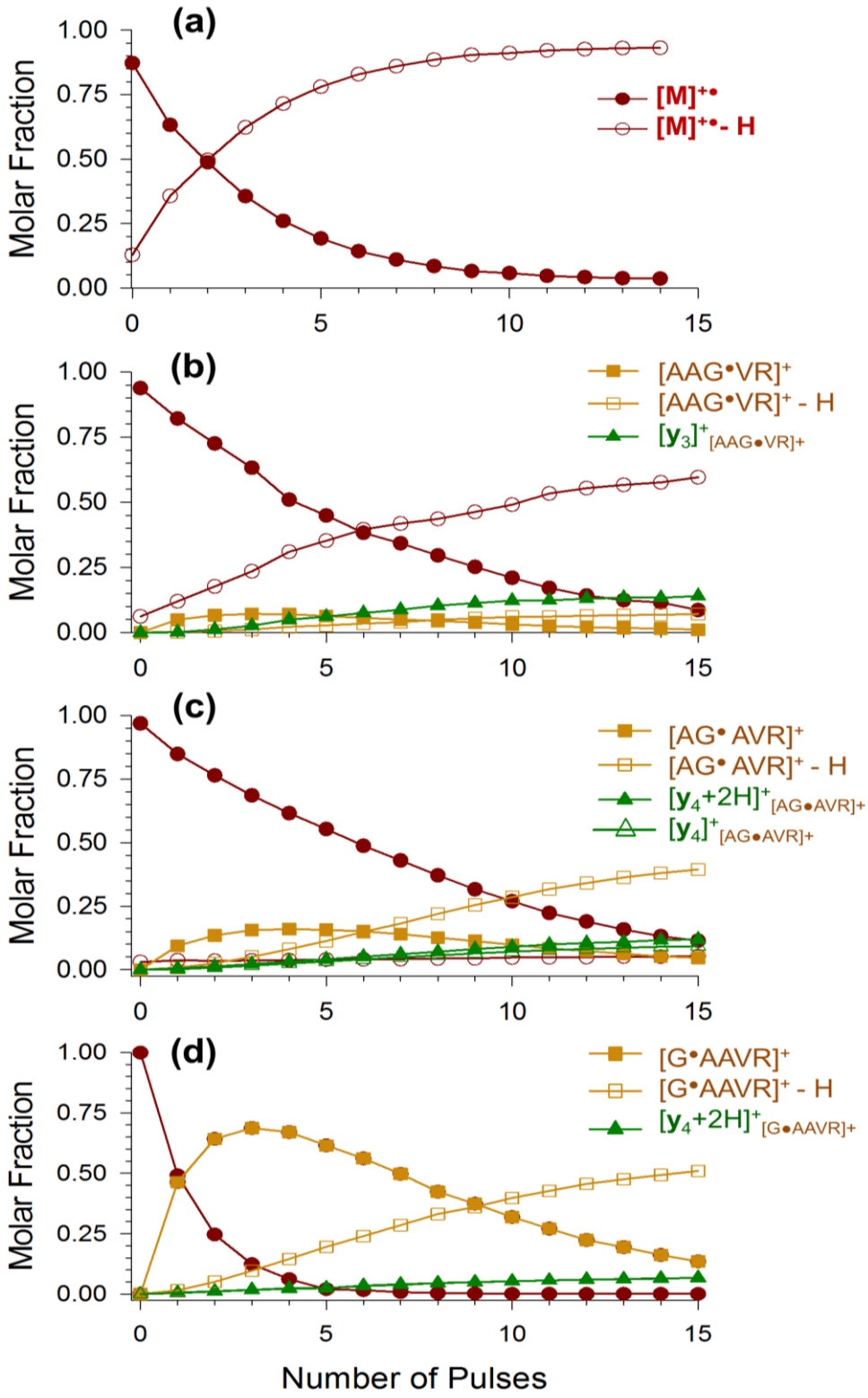


Figure 2.18. Photodepletion curves of m/z 444 ions at 355 nm as a number of laser pulses (a) $[AAAG^{\bullet}R]^{+}$, (b) $[AAG^{\bullet}AR]^{+}$, (c) $[AG^{\bullet}AAR]^{+}$, and (d) $[G^{\bullet}AAAR]^{+}$.

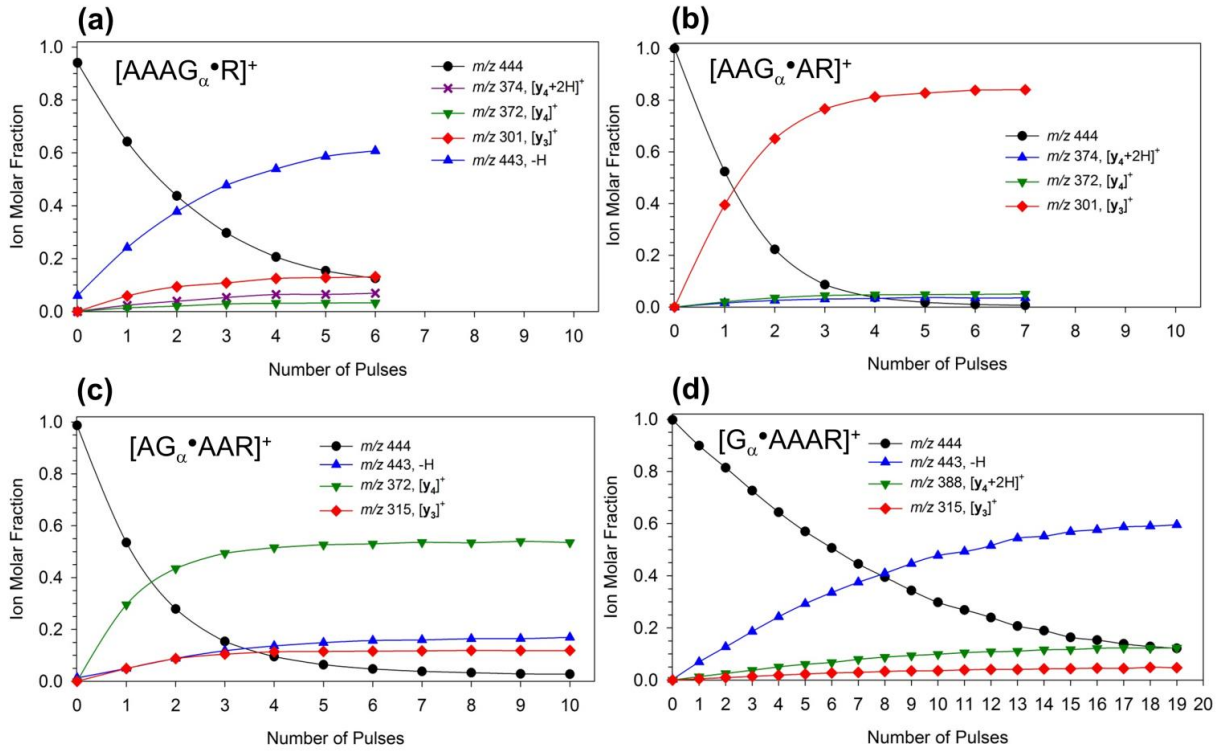


Figure 2.19. LTQ CID-MS3 spectrum of $[AAYR]^{+\bullet}$ (m/z 479).

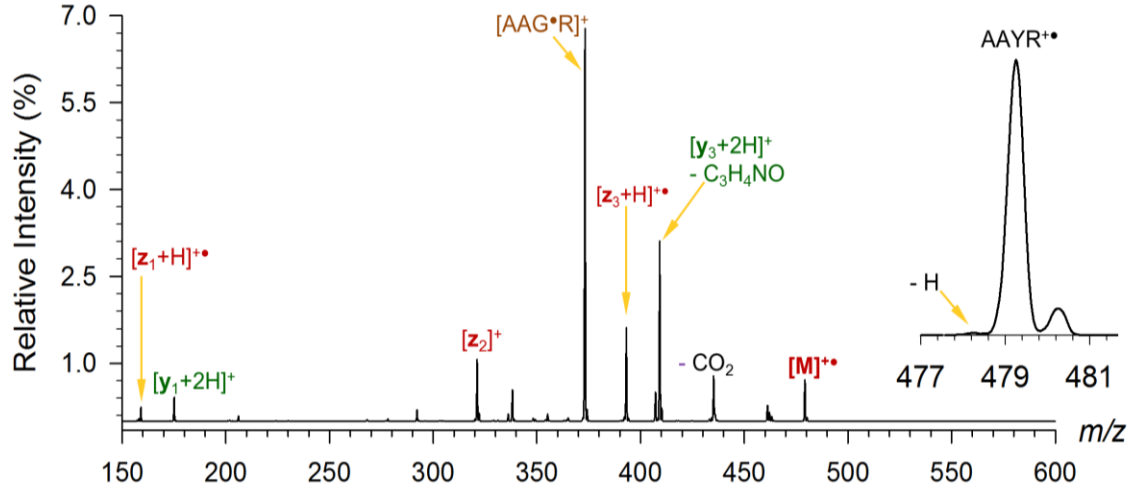


Figure 2.20. 355 nm-UVPD-MS3 spectra (10 pulses at 15 mJ/pulse) of (a) $[AAYR]^{+\bullet}$ (m/z 479), (b) $[d11\text{-}AAYR\ ]^{+\bullet}$ (m/z 490).

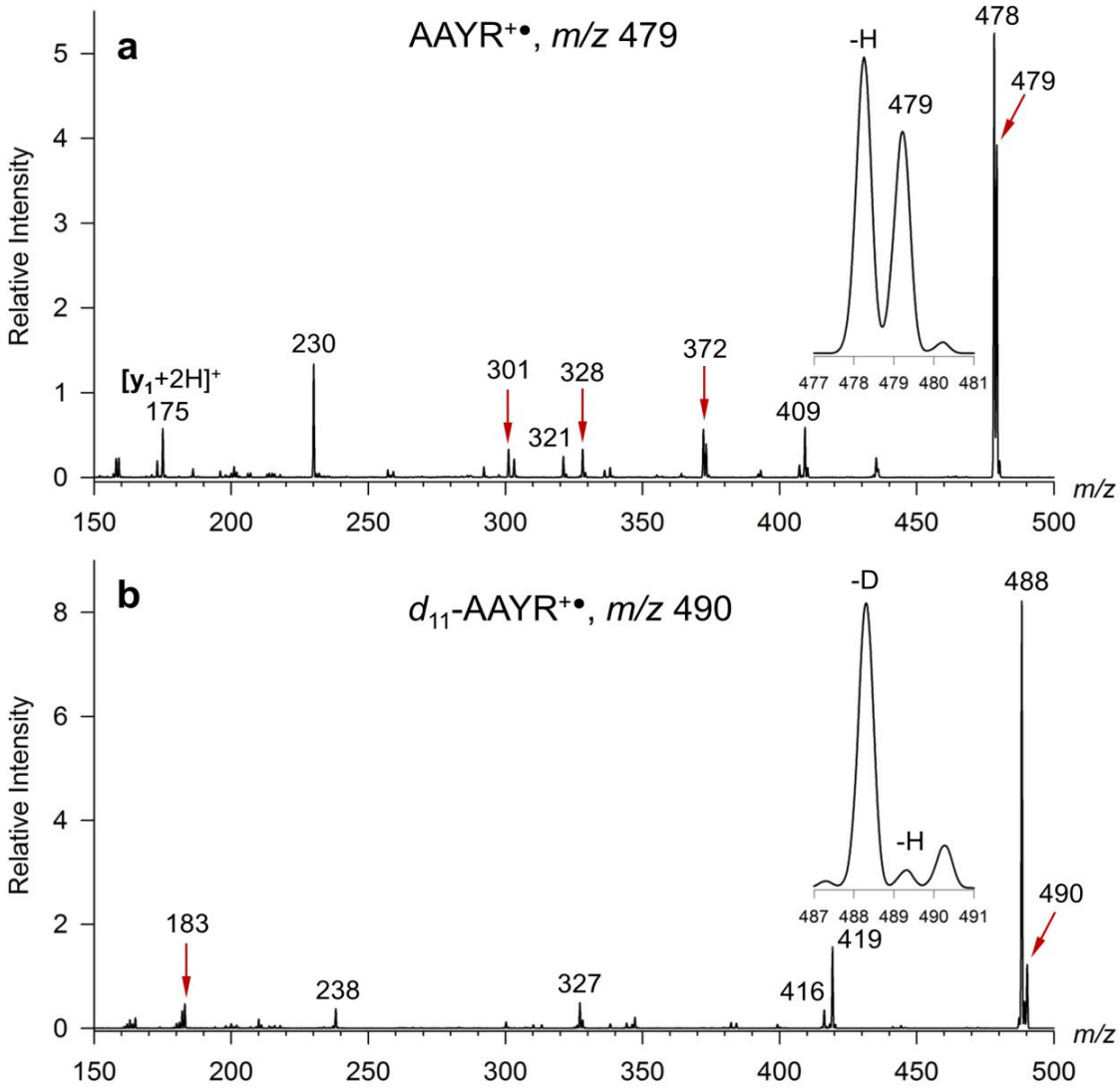


Figure 2.21. UVPD-MS3 spectrum of (a) [AAY(OCH3)R]$^{+\bullet}$ (m/z 493), (b) [d10-AAY(OCH3)R]$^{+\bullet}$ (m/z 503).

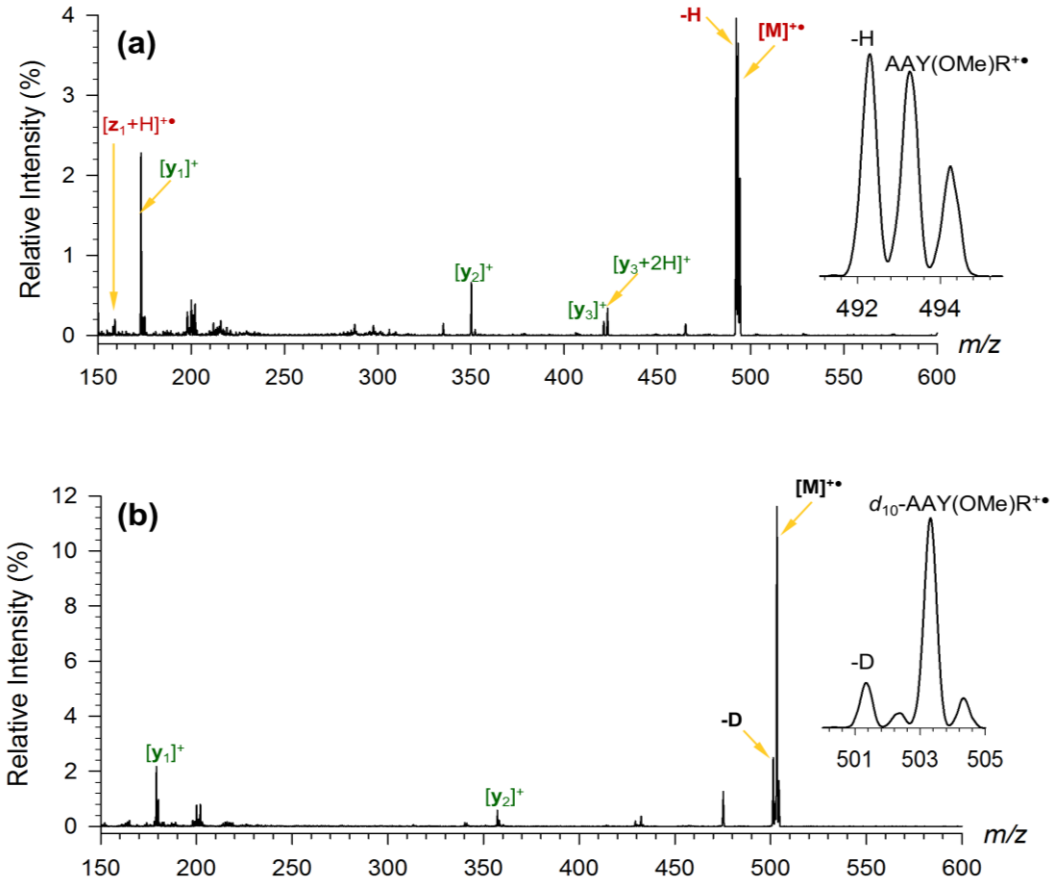


Figure 2.22. LTQ-CID-MS3 spectrum of [AAY(OCH3)R]$^{+\bullet}$ (m/z 493).

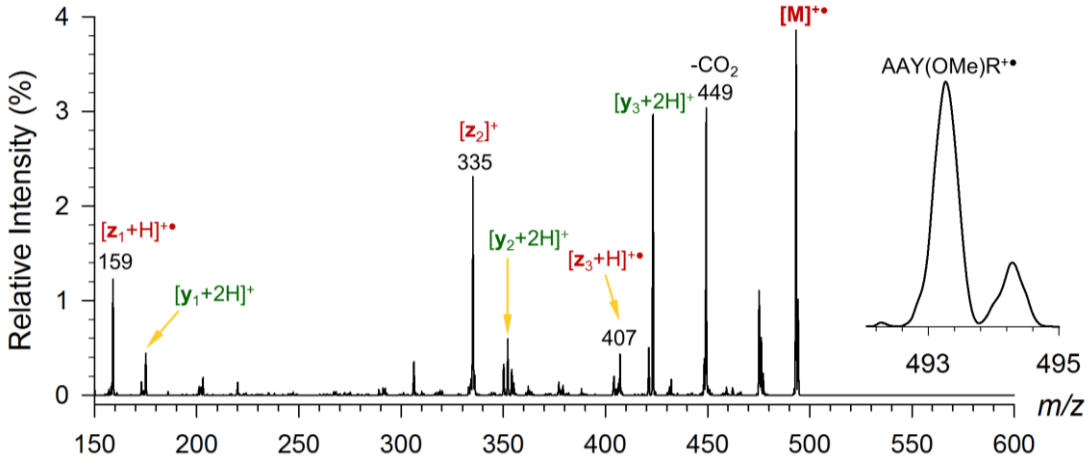
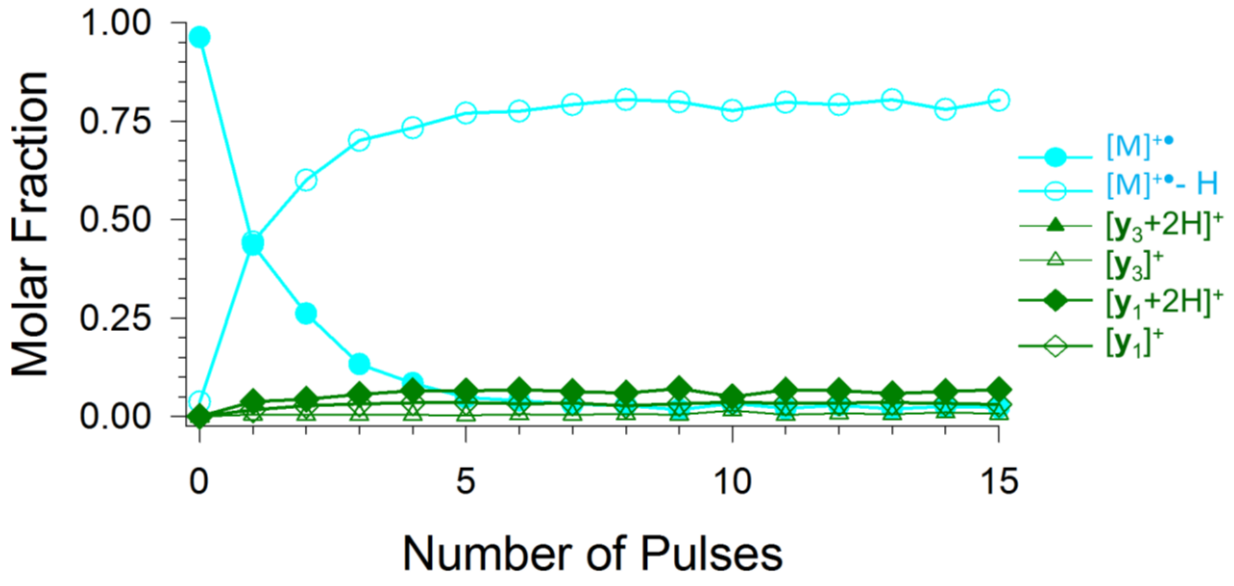


Figure 2.23. Photodepletion curve of [AAY(OCH3)R]+● at 355 nm as a number of laser pulses.



Chapter 3 UV/Vis Action Spectroscopy and Structures of Tyrosine Peptide Cation Radicals in the Gas Phase

3.1 Introduction

The tyrosine residue in proteins and peptides is susceptible to one-electron oxidation, forming aromatic cation radicals and tyrosyl O-radicals as transient reactive intermediates. These biologically important reactions play a role in the functioning of redox enzymes such as galactose oxidase,[201] ribonucleotide reductase,[202] and cytochrome C oxidase,[204,205] among others, as reviewed.[206] Tyrosyl radicals in solution and condensed phase have been primarily characterized by electron spin resonance (ESR) spectroscopy and UV-VIS spectroscopy.[207-210] Fast UV-VIS spectroscopy has also been used to characterize tyrosine radicals produced by hydrogen transfer to oxidized tryptophan residues, and a solvent dependence of the absorption maximum has been noted.[211] In contrast to solution and condensed phase studies, transient intermediates that are generated in the gas phase are immune to environmental effects and can be studied as isolated species under strictly unimolecular conditions.[212] Gas-phase tyrosine anion-radicals have been generated by electron photodetachment from peptide molecular dianions and characterized by vacuum UV spectroscopy.[213] Another efficient, and chemically more amenable, method of preparation of gas-phase peptide cation radicals relies on intramolecular one-electron oxidation of a peptide ligand by a transition metal dication in a ternary complex,[115,121,214] according to the equation $[Cu^{(II)}(Ligand)peptide]^{2+\bullet} \rightarrow [Cu^{(I)}(Ligand)]^{+} + [peptide]^{+\bullet}$. According to their stoichiometry, which is identical to that of the pertinent neutral peptide, peptide cation-radicals of this type are called hydrogen-deficient.[104] Tryptophan, a readily oxidizable amino acid, has been characterized as a cation-radical in the gas phase by infrared[125] and UV action spectroscopy.[215]

For the most part, gas-phase peptide radicals have been studied by collision induced dissociation (CID) tandem mass spectrometry,[115,214] which relies on product analysis but is related only indirectly to the peptide cation radical structure. Clearly, a more direct method of structure elucidation is needed to characterize gas-phase peptide cation radicals formed from metal complexes and resolve the issue of assigning their structure that has persisted for 16 years since the first generation of peptide cation radicals.[116] Here, we report the first study of UV photodissociation (UVPD) action spectroscopy[216] and structure characterization of two tyrosine-containing peptide cation radicals, $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$.[80] These representative peptide sequences were

selected to provide a well-defined protonation site on the arginine residue and thus to reduce the structural ambiguity to determining the position of the hydrogen-deficient radical site and resolving conformational effects. In addition, these two peptide cation-radicals were found to undergo quite different dissociations upon collisional activation and photodissociation, indicating that their radical sites may be different despite being generated by the same electron transfer reaction.

3.2 Experimental Section

3.2.1 *Materials*

All peptides were synthesized on Wang resin (Bachem Americas, Torrance, CA, USA) using commercially available Fmoc peptides (Life Technologies, Rockford, IL, USA) according to previously described procedures,[217,218] and purified by ion-exchange chromatography. 2,2':6',2''-terpyridine and $Cu(NO_3)_2$ were obtained from Sigma-Aldrich. [Cu(tpy)peptide] complexes were prepared in situ by mixing solutions of the peptide (50 μM) and the Cu(tpy) complex (600 μM) in aqueous methanol and diluting to 10 μM concentration.[118] Electrospray ionization of the complex solution with a home-built microspray source was used to generate gas-phase $[Cu(tpy)peptide]^{2+\bullet}$ ions. The ^{65}Cu or ^{63}Cu isotopologues were selected by mass and subjected to collision-induced dissociation (CID) at collision energies that were tuned to optimize peptide cation-radical formation, typically at normalized collision energies (NCE) set to 15-18 instrument units. The $[peptide]^{+\bullet}$ ions at the corresponding *m/z* were selected by mass and subjected to CID or photodissociation (UVPD). The CID and UVPD mass spectra were measured on a modified LTQ-XL linear ion trap mass (LIT) spectrometer (ThermoElectron Fisher, San Jose, CA, USA) equipped with a laser system. High-resolution mass spectra were measured on an LTQ-Orbitrap (ThermoElectron Fisher, San Jose, CA, USA). The peptide cation-radicals were prepared by CID in the LTQ and transferred to the Orbitrap for high-resolution measurements using Fourier-Transform treatment of the time-domain signal. The resolution was set to 100,000.

3.2.2 *Photodissociation*

Photodissociation of trapped ions in the LIT was performed as reported previously.[73] The irradiating light beam was produced by an Nd-YAG EKSPLA NL301G laser (Altos Photonics, Bozeman, MT, USA)

operating at 20 Hz frequency with a 3-6 ns pulse width. Photons exiting the pump laser are fed into a PG142C unit (Altos Photonics, Bozeman, MT, USA which integrates a third harmonic generator and optical parametric oscillator coupled with an optional second harmonic generator (SSH) to provide wavelength tuning between 210-409 nm at a pulse peak power ranging between 0.79-2.06 mJ. These powers are measured at each wavelength using an EnergyMax-USB J-10MB energy sensor (Coherent Inc., Santa Clara, CA, USA) to which the relative laser pulse power is later calibrated as the wavelength is changed. The laser beam is aligned by mirrors and focused by a telescopic lens to pass the small aperture drilled in the auxiliary chemical ionization source source of the LTQ-XL which can be used to produce electron donor reagent ions.[73] The laser beam diameter in the LIT is estimated at 3-4 mm to ensure overlap with the trapped ions. The beam position does change by about 3 mrad in switching between the SSH/FSH region (210-354 nm) and the ESH region (355-409 nm) which requires the realignment of mirrors. This realignment does inevitably change the cross section of the ion cloud exposed to irradiation, and different ratios of dissociation are observed in comparing the regions at normalized power and pulse frequency. Thus the ESH region (355-409 nm) has been scaled (4×) to compensate for this change. The typical experimental set up consists of selecting the ion to be photodissociated and storing it in the LIT for a chosen time period. For example, 400-ms storage time can accommodate up to 7 laser pulses spaced by 50 ms. This allows one to vary the number of pulses which are also normalized as the number of pulses used is varied depending on the degree of dissociation. The pulse-dependent UVPD measurements were performed with the 355 nm line from the laser source at 15 mJ/pulse laser power, as described previously.[102]

3.2.3 *Computation*

Conformational search for cation radicals was performed using a modification of the ConformSearch protocol.[219] Molecular dynamics calculations using NAMD[220] and the CHARMM force field[221] in a replica-exchange format[222] were first run with 8 replicas for 1 ns to generate 800,000 conformations of singly-protonated peptides, e.g. $[YAAAR + H]^+$ and $[AAAYR + H]^+$. Eight thousands of output structures were sampled at 10 ps intervals and converted to $[YAAAR]^{+\bullet}$ and $[AAAYR]^{+\bullet}$ cation-radicals that were fully optimized with PM6.[223] All calculations of open-shell species were performed for doublet spin states within

the spin-unrestricted formalism. The PM6 structures were grouped into families according to their hydrogen bonding patterns, and the lowest-energy representatives from each family were selected for DFT geometry optimization with B3LYP,[186] M06-2X,[187] and ωB97X-D[188] and the 6-31+G(d,p) basis set. The optimized structures were characterized by harmonic frequency calculations (B3LYP/6-31+G(d,p)) as local energy minima. Vertical excitation energies for the doublet spin states were obtained by time-dependent DFT calculations[224] using the ωB97X-D functional and the 6-311++G(2d,p) basis set in a spin unrestricted formalism. Typically 25 excited states were generated by TD-DFT calculations and those with spin-expectation values <1.5 were considered. Single-point energies were obtained by calculations that used the above DFT methods and Moller-Plesset perturbational treatment[189] (MP2, frozen core) with the 6-311++G(2d,p) basis set. UMP2 calculations of $YAAAR^{+\bullet}$ cations displayed large spin contamination that could not be eliminated by standard spin annihilation procedures.[225,226] All open-shell species were treated by restricted open-shell (ROMP2) single-point energy calculations[227,228] and the energies are included in **Table 3.4 (SI)**. Rice-Ramsperger-Kassel-Marcus (RRKM) calculations were carried out as described previously.[229]

3.3 Results and Discussion

The $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$ ions were generated from the respective ternary Cu complexes with 2,2':6',2''-terpyridine (tpy) which were formed as doubly charged ions by electrospray ionization.[118,230] CID or laser photodissociation at 355 nm of these doubly charged complexes produced the respective peptide cation radicals at *m/z* 550 and their identity was corroborated by accurate mass measurements for the CID-produced ions (**Table 3.1, SI**). Peptide cation-radicals produced from the ^{63}Cu and ^{65}Cu isotopologues of the $[Cu(tpy)(peptide)]^{2+\bullet}$ complexes showed identical behavior upon CID and UVPD. The photodissociation spectra[81,102] of mass-isolated $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$ obtained with a single laser pulse at 355 nm were distinctly different (**Figure 3.1a,b**).

UVPD of $YAAAR^{+\bullet}$ resulted in a dominant elimination of benzoquinone methide (C_7H_6O, 106 Da neutral fragment) from the tyrosine residue. This is a common dissociation of tyrosine-containing hydrogen-deficient peptide cation radicals[214,116] and is also prevalent in the CID spectrum of the $YAAAR^{+\bullet}$ ion (**Figure 3.4a, SI**). In contrast, UVPD of $AAAYR^{+\bullet}$ resulted in predominant loss of a hydrogen atom (*m/z* 549, **Figure 3.1b**) that was

accompanied by a minor elimination of C_3H_4NO from the *N*-terminal Ala residue (*m/z* 480, **Figure 1b**).

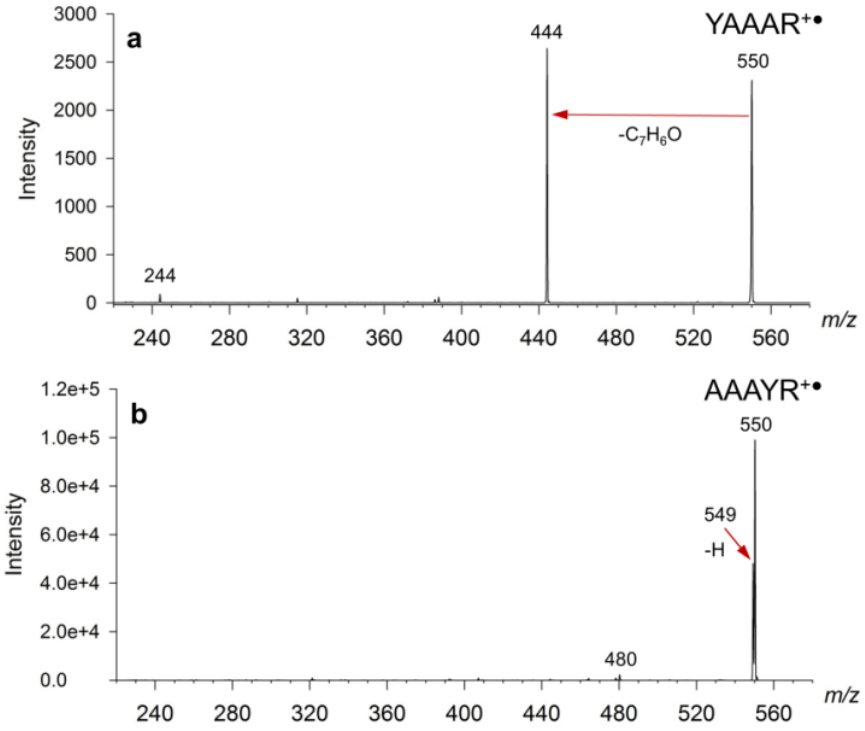


Figure 3.1. Single-pulse UVPD spectra of (a) YAAAR+● and (b) AAAYR+● at 355 nm.

The photo-induced dissociation was dramatically different from CID of $AAAYR^{+\bullet}$ which produced a number of backbone fragment ions, dominated by a major loss of C_3H_4NO from the *N*-terminus. All fragment ion assignments were corroborated by accurate mass measurements (**Table 3.1, SI**). The origin of the H atom was investigated with a lower homologue, $AAYR^{+\bullet}$, which undergoes very similar dissociations upon CID and UVPD as does $AAAYR^{+\bullet}$ (**Figure 3.5a, SI**), and in which all the labile protons can be exchanged for deuterium with a high molar conversion (90-93%)[231] to produce d_{11}-$AAYR^{+\bullet}$ (**Figure 3.5b, SI**). UVPD of d_{11}-$AAYR^{+\bullet}$ resulted in predominant (>94%) loss of D (**Figure 3.6, SI**), indicating that photodissociation selectively targeted the exchangeable hydrogen atoms in the N–H or O–H bonds.

Photodissociation at 355 nm was further studied by pulse-dependent experiments that showed exponential depletion of $[YAAAR]^{+\bullet}$ and $[AAAYR]^{+\bullet}$ ion intensities. This is illustrated by UVPD spectra after 10 laser pulses that show <2% of residual $[YAAAR]^{+\bullet}$ and $[AAAYR]^{+\bullet}$ ions (**Figure 3.7, SI**). The photodepletion

curve for $[YAAAR]^{+\bullet}$ was fitted with an exponential decay function $I(n) = I(0)e^{-(0.7367n+0.004)} + 0.0035$, where $I(0)$ is the initial $[YAAAR]^{+\bullet}$ ion intensity and n is the number of laser pulses, giving 0.6% root-mean square deviation (rmsd)(**Figure 3.8a, SI**). The photodepletion curve for $[AAAYR]^{+\bullet}$ showed a slower decay curve, $I(n) = I(0)e^{-(0.329n+0.113)} + 0.0175$, giving 4.6% rmsd against the experimental data (**Figure 3.8b, SI**).

The light absorption properties of $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$ were further studied through UV photodissociation action spectra that were measured in the 210-700 nm region. Light absorption was indicated by dissociations forming fragment ions while depleting the precursor cation-radical relative intensity. $YAAAR^{+\bullet}$ showed two major bands with maxima at 260 and 300 nm, and a broad composite band at 340-390 nm. Across this wavelength region the loss of C_7H_6O was the predominant photodissociation channel. No photodissociation was observed above 420 nm (**Figure 3.2a inset**). The action spectrum of $AAAYR^{+\bullet}$ was distinctly different, showing bands with maxima at 219, 225 and 246 nm, a broad composite band at 250-310 nm tailing to 380 nm, and another broad absorption band covering the 400-660 nm region (**Figure 3**). Loss of H was the predominant photodissociation channel across the entire wavelength region where absorption was observed. The action spectra further indicated that the photodepletion curves at 355 nm were due to only weakly absorbing parts of the spectra, consistent with the slow photodepletion at this wavelength (**Figure 3.8b, SI**).

The very different photodissociation and action spectra of $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$ strongly indicated that these isomeric peptide cation radicals had different radical chromophore groups. To interpret the action spectra, we performed extensive density functional theory (DFT) calculations of ion structures that were combined with time-dependent DFT (TD-DFT) calculations of electronic excitation energies and oscillator strengths in the cation-radicals. Structures of $YAAAR^{+\bullet}$ were generated in several steps starting with an exhaustive search of the conformational space of arginine-protonated YAAAR ions. Several lowest-energy ions obtained by this procedure were converted to $YAAAR^{+\bullet}$ radicals and fully optimized with B3LYP and ωB97X-D[188] DFT methods, as described in the Supplement (**Table 3.2, Figure 3.9, SI**). Previous benchmarking of excitation energies and transition intensities had identified ωB97X-D as the most reliable TD-DFT method[78] for peptide radical chromophores, which was therefore used to calculate absorption spectra of $YAAAR^{+\bullet}$, as shown for four lowest-energy conformers **1**-**4** (**Figure 3.2b**). The calculated spectra show an excellent overall agreement with the experimental action spectrum.

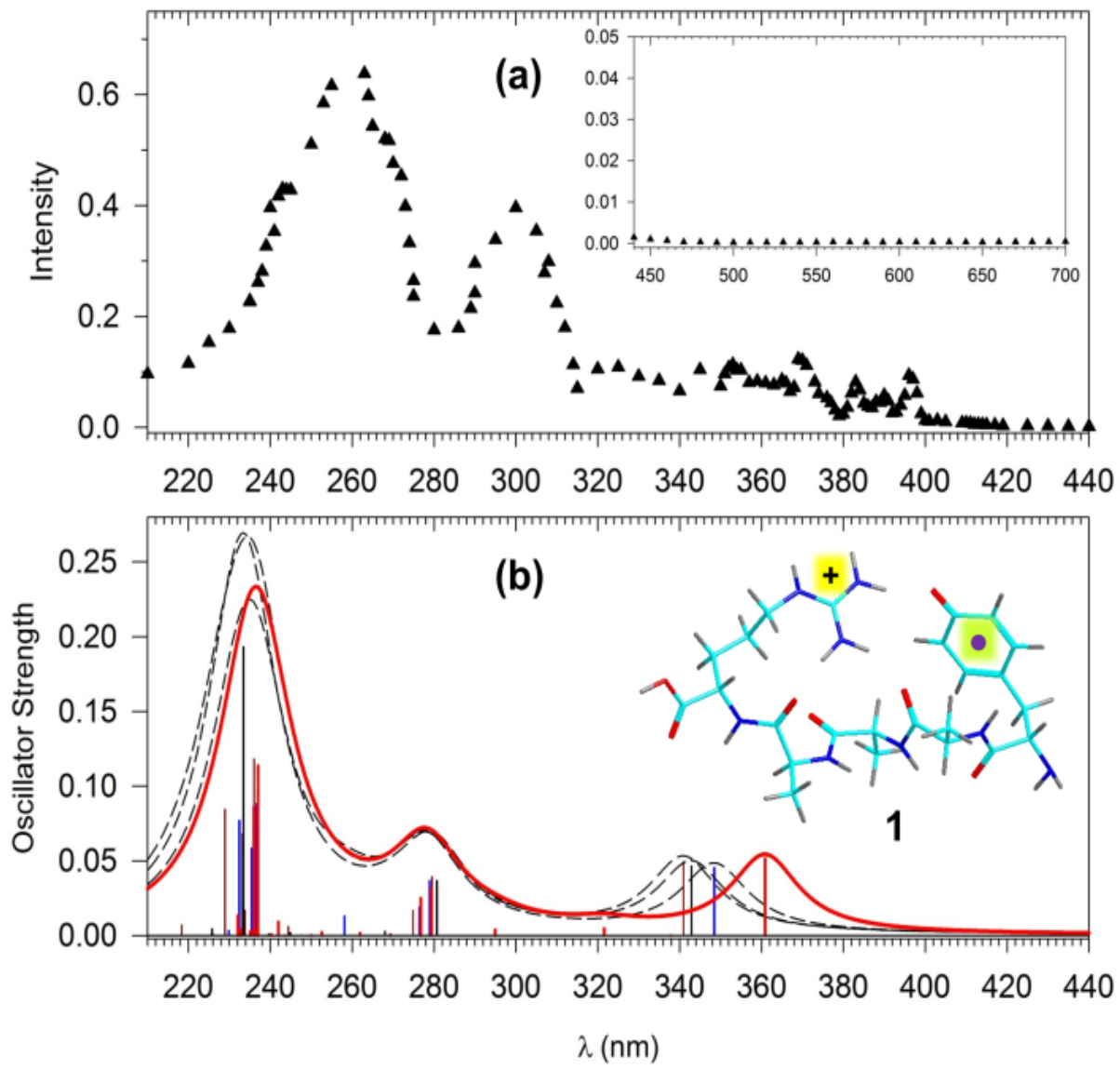


Figure 3.2. (a) Action spectrum of YAAAR$^{+\bullet}$. The trace (▲) shows the wavelength-dependent relative intensity of the major photofragment ion at *m/z* 444 (loss of C_7H_6O) after single-pulse photodissociation. Inset shows the absorption-free 440-700 nm region. (b) Calculated UV-VIS absorption spectra of YAAAR$^{+\bullet}$ conformers **1**(red line) and **2**, **3**, and **4** (dashed lines). The calculated wavelengths (bars) were convoluted with Lorentzian functions at 12 nm fwhm. Inset shows the ωB97X-D optimized structure of **1**. For other conformers see **Figure 3.10** in the Supplemental Information (**SI**).

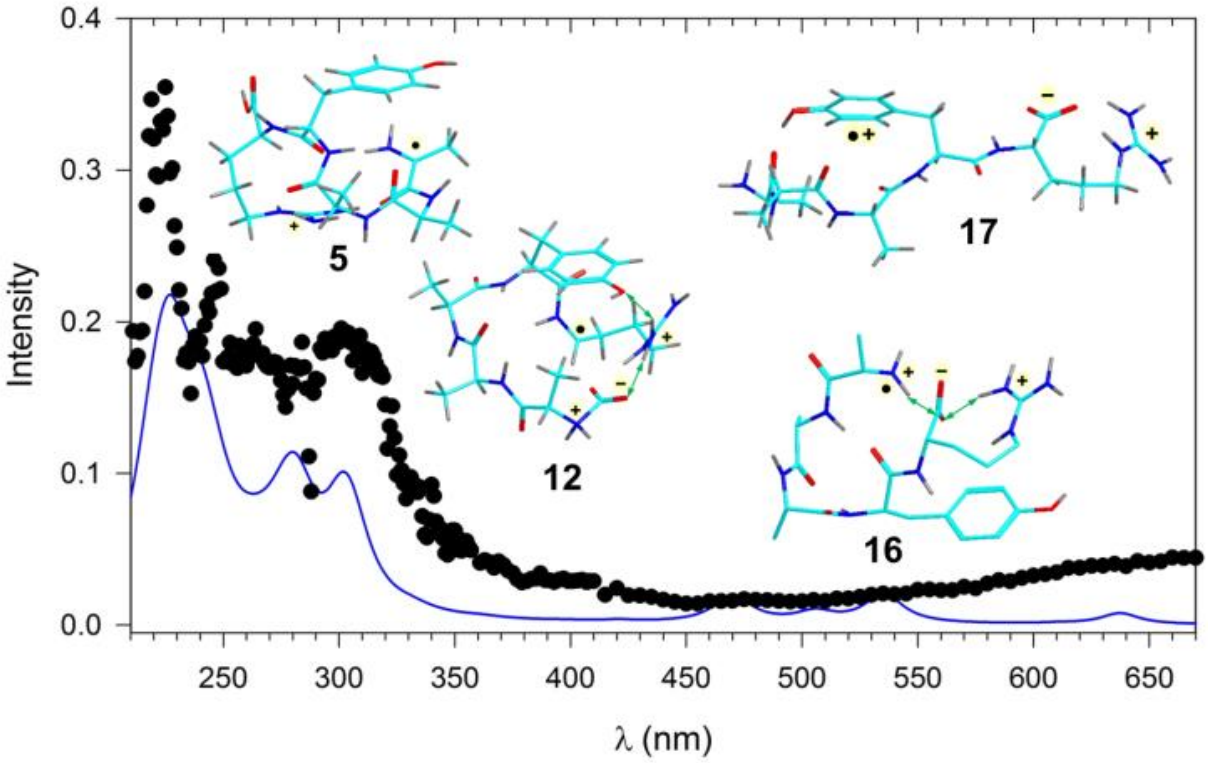


Figure 3.3. Action spectrum of AAAYR+● (black circles) and combined absorption spectra of 5, 6, 12, 16, and 17 ion isomers (blue line).

The red shift of the major bands in the action spectrum can be in part explained by v'' ≥ 1 to v' = 0 vibronic transitions in thermal ions at the ion trap temperature (310 K), whereas the calculated spectra refer to vertical excitations at 0 K. According to the calculated harmonic frequencies, there are ca. 50 (22%) normal modes in $YAAAR^{+\bullet}$ with frequencies <400 cm^{-1} for which statistical thermodynamics analysis[232] predicts >20% population of v'' ≥ 1 states at 310 K.

The two short wavelength UV bands in Figure 2b are remarkably insensitive to the ion conformation. The 360 nm band for **1** shows a larger spread of excitation energies among **1**-**4** depending on the ion conformation. This can be explained by the electronic nature of this excitation (**Figure 3.10, SI**). The major transitions resulting in the 360 nm band involve electron excitation to the semi-occupied aromatic p-orbital at the Tyr residue (MO147). Most of these transitions occur from doubly occupied p-orbitals at backbone amide groups (MO133, 144, 145, 146, **Figure 3.10, SI**) and are therefore expected to be sensitive to their position with respect to the Tyr ring, which in turn depends on the peptide ion conformation. Considering the dominant

loss of the tyrosineC_7H_6O side chain upon both thermal activation and photoexcitation, as well as the match of the UV-action spectrum with the theoretical electron excitations, we can unequivocally conclude that the $YAAAR^{+\bullet}$ ions formed from the Cu complexes have the structure of Tyr O-radicals.

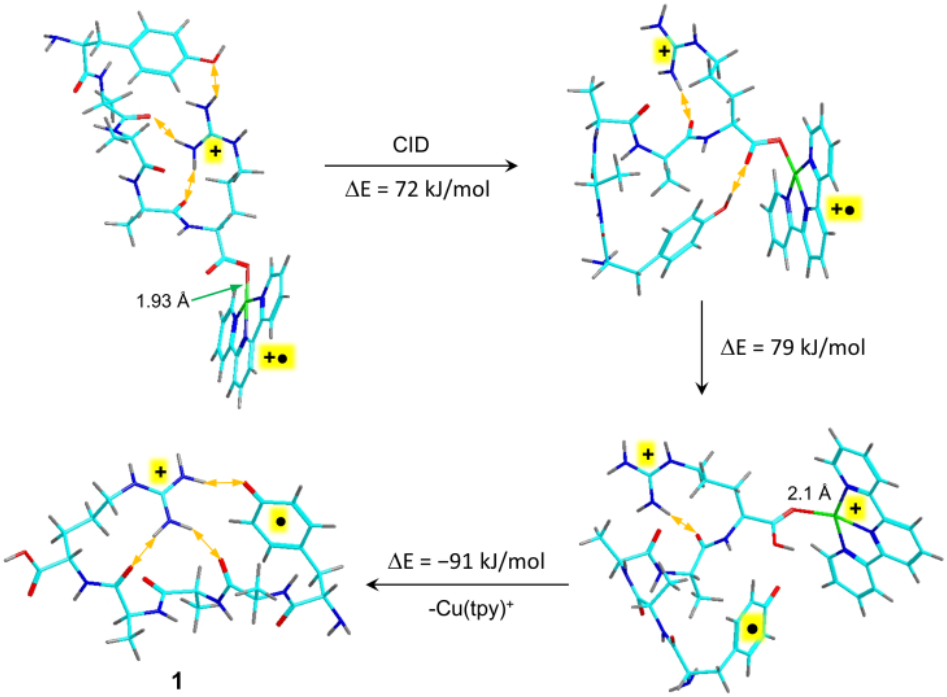
In contrast to $YAAAR^{+\bullet}$, analysis of the $AAAYR^{+\bullet}$ action spectrum indicated that the ion population consisted of a mixture of isomers. Multiple $AAAYR^{+\bullet}$ structures were generated by exhaustive conformational search followed by DFT geometry optimization. The calculations indicated that one-electron oxidation of the Tyr residue was accompanied by extensive rearrangements and dissociations, yielding only a minority of stable Tyr cation radicals with an electron defect in the aromatic ring. These Tyr cation radicals represented high-energy local potential energy minima that were metastable with respect to exothermic rearrangements or dissociations. **Figure 3.11 (SI)** shows the representative structures of several types of $AAAYR^{+\bullet}$ isomers (**5**-**18**) which span a broad range of relative energies (**Table 3.3. SI**). TD-DFT calculations indicated that none of the several $AAAYR^{+\bullet}$ ion structures alone gave an acceptable match of the theoretical absorption spectrum with the action spectrum. Because of the lack of C_7H_6O elimination from $AAAYR^{+\bullet}$ (Figure 1b), structures **9**-**11** having Tyr O-radical groups (**Figure 3.11, SI**) were excluded. Likewise, structures **14** and **15** (**Figure 3.11, SI**) having an electron defect in the *N*-terminal amino group were excluded on the basis of their calculated absorption spectra that showed strong absorption bands in the 380-450 nm region which were absent in the action spectrum. In a large number of optimized structures, one-electron oxidation of the carboxyl group resulted in cleavage of the Arg C_α–$COO^\bullet$ bond in the intermittent carboxyl radical, forming ion-molecule complexes of Arg C_α-radicals with CO_2 (e.g., **7** and **8**, **Figure 3.11, SI**). Although these complexes had low relative energies, they were only weakly bound (ΔH_{diss} < 40 kJ mol^{-1}) with respect to CO_2 elimination. Since the formation of $AAAYR^{+\bullet}$ by CID of the Cu complexes is not accompanied by CO_2 loss (**Figure 3.12, SI**), nor it occurs on UVPD, the CO_2 complexes are unlikely to be present in the population of stable $AAAYR^{+\bullet}$ ions and thus can be excluded. Considering both the spectroscopic data and observed photodissociations, we conclude that the $AAAYR^{+\bullet}$ ions are represented by a mixture of isomers whose combined absorption spectra provide an acceptable agreement with the action spectrum of the ion (**Figure 3**). The individual spectra of several isomers are shown in **Figure 3.13, SI**. Ions **5**, **6**, and **12** (**Figure 2**) can provide the absorption bands in the 210-350 nm region, but do not have transitions in the visible region. The chromophore in ions **5** and **6** is an Ala-C_α radical group that shows electronic transitions in the 220-380 nm region. Ion **12** and its

conformers are Arg-C_α radicals for which we calculate a strong absorption band at 300 nm which is prominent in the action spectrum. We found the *N*-terminal amine cation radicals (**16** and its conformers) and the Tyr aromatic cation-radicals (**17** and **18**) as the only chemically compatible structures that had absorption bands in the 450-700 nm region to account for the long-wavelength part of the action spectrum. Ions **5** and **6** belong to the more stable structures (**Table 3.3, SI**) in which one-electron oxidation was associated with H_α migration onto the COO group. Ion **12** is also a product of COO^- oxidation, followed by CO_2 migration onto the proximate *N*-terminal amine. Isomer **12** is substantially less stable than **5** and its presence in the oxidized ion population can be only explained by kinetic trapping. Ions **16** and **17**, **18** which result from oxidation of the *N*-terminal amine group and Tyr aromatic ring, respectively, also represent high-energy isomers. In summary, the experimental action spectrum of $AAAYR^{+\bullet}$ is best interpreted as arising from photodissociation of a mixture of cation-radical isomers produced by one-electron oxidation of the peptide ligand in the $[Cu(tpy)(AAAYR)]^{2+\bullet}$ complex.

The nature of the observed photodissociations was investigated by analyzing the kinetics of the major reactions of $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$. Loss of C_7H_6O from **1** was calculated to be 156 kJ mol^{-1} endothermic and required 154 kJ mol^{-1} in the transition state for breaking the Tyr C_α–C_β bond, forming the $[^\bullet GAAAR^+ + C_7H_6O]$ complex at 84 kJ mol^{-1} relative to **1** (ωB97X-D/6-311++G(2d,p) + zero-point energies, **Table 3.4, SI**). RRKM calculations using these energies and assuming vibronic redistribution of internal energy indicate <7% dissociation within the experimental time of 50 ms for **1** having up to 466 kJ mol^{-1} internal energy. Considering the mean rovibrational enthalpy of thermalized **1** (100 kJ mol^{-1} at 310 K) and excitation energy from single photon absorption (337 kJ mol^{-1} at 355 nm), the dissociation is predicted to be slow, leading to <2% depletion of **1** (**Figure 3.14a, SI**). This contrasts the photodepletion curve of $YAAAR^{+\bullet}$ which shows 55% dissociation by loss of C_7H_6O in 50 ms following one 355-nm laser pulse (Figure 1a, **Figure 3.8a, SI**). This leads to the conclusion that the photodissociation occurs locally from an excited electronic state involving the Tyr-O radical. The dissociation kinetics under slow-heating conditions of resonant collisional activation can be described by transition-state theory, which gives 90% dissociation within 50 ms at 700 K (**Figure 3.14b, SI**). Such an effective temperature is readily accessible in CID in the ion trap.[17,233,234]

Similar conclusions can be made regarding the photodissociative loss of an exchangeable H atom from $AAAYR^{+\bullet}$. Calculations of loss of amine and amide H atoms in **5** gave TS energies of >210 kJ mol^{-1} which were

prohibitively high to drive dissociation of an isolated ion in 50 ms. Again, a plausible explanation of the kinetics must consider dissociation proceeding from an excited electronic state but preceding vibronic relaxation of the



Scheme 3.1. Conformational transformation in $[Cu(tpy)(YAAAR)]^{2+\bullet}$ followed by intramolecular electron transfer oxidation of the peptide ligand, proton transfer, and dissociation to $YAAAR^{+\bullet}$ (1) and $[Cu(tpy)]^{+}$. The yellow double-ended arrow indicate major hydrogen bonds.

excitation energy, as described recently for photodissociation of a peptide anion-radical.[77] This conclusion is also consistent with the absence of H loss from $AAAYR^{+\bullet}$ upon vibrational activation by collisions in the slow-heating excitation regime (**Figure 3.4b, SI**).

The different structures and hence chemistry of $YAAAR^{+\bullet}$ and $AAAYR^{+\bullet}$ are the result of different and sequence-dependent courses of intramolecular electron transfer in the Cu(tpy)(peptide) complexes. The peptide ligand in $[Cu(tpy)(YAAAR)]^{2+\bullet}$ can assume a conformation in which the Tyr side chain is in the vicinity of the Cu-binding COO^{-} group, forming a hydrogen bond (see the ωB97X-D optimized structures in **Scheme 3.1**). Following electron-transfer oxidation, the Tyr cation radical can intramolecularly protonate the COO^{-} group, forming a Tyr-O radical while weakening the peptide ion bonding to $Cu(tpy)^{+}$. In contrast, the AAAYR ligand likely adopts several different conformations in the Cu(tpy) complex. Those where Tyr participates in electron transfer presumably lead to the formation of aromatic cation-radicals **17** and **18** (**Figure 3.11, SI**)

whereas oxidized intermediates from other complex conformers are either amine cation radicals (**16**) or carboxyl radicals undergoing rapid stabilization by hydrogen transfer forming the most stable Ala-C_α (**5**) and Arg-C_α (**12**) radical isomers.

3.4 Conclusion

In conclusion, UV-VIS photodissociation action spectroscopy allowed us for the first time to establish structures of tyrosine peptide cation radicals formed by intramolecular electron transfer in Cu complexes. The structures point to different modes of peptide ligand oxidation depending on the position of the tyrosine residue in the peptide sequence.

3.5 Supplemental Information

Table 3.1. Accurate Mass Measurements.[a]

Ion *m/z*	Exp.	Calc.	Elemental composition	Assignment
YAAAR				
423	423.6534	423.6585	$^{63}CuC_{39}H_{50}N_{11}O_6$	$[Cu(tpy)YAAAR]^{2+}$
550	550.2853	550.2858	$C_{24}H_{38}N_8O_6$	$YAAAR^+$
444	444.2436	444.2439	$C_{17}H_{32}N_8O_5$	$[YAAAR - C_7H_6O]^+$
AAAYR				
423	423.6559	423.6585	$^{63}CuC_{39}H_{50}N_{11}O_6$	$[Cu(tpy)AAAYR]^{2+}$
550	550.2846	550.2858	$C_{24}H_{38}N_8O_7$	$AAAYR^+$
480	480.2554	480.2565	$C_{21}H_{34}N_7O_5$	$[AAAYR - C_3H_4NO]^+$
464	464.2368	464.2378	$C_{21}H_{32}N_5O_5$	$[AAAYR - C_3H_6N_2O]^+$
444	444.2430	444.2439	$C_{17}H_{32}N_8O_5$	$[AAAYR - C_7H_6O]^+$
392	392.1922	392.1927	$C_{18}H_{26}N_5O_5$	$[AAAYR - C_6H_{12}N_3O_2]^+$
321	321.1553	321.1557	$C_{15}H_{21}N_4O$	

[a]Measured on a Thermo-Electron Fisher LTQ-Orbitrap mass spectrometer at mass resolving power of 120,000.

Table 3.2. Relative Energies of YAAAR Cation Radicals.

Ion	Relative Energy[a,b]
1	0
2	19
3	24
4	35
5	20
6	20

[a]In units of kJ mol^{-1}. [b]From ωB9X-D/6-311++G(2d,p) single-point energy calculations on ωB9X-D/6-31+G(d,p) optimized structures.

Table 3.3. Relative Energies of AAAYR Cation Radicals.

	Relative Energy[a,b]			
Ion	B3LYP[c]	M06-2X[d]	ωB97XD[e]	Structure Type
5	0	0	0	*N*-term. Ala-$C_\alpha^\bullet$, COOH
6	9	51	52	*N*-term. Ala-$C_\alpha^\bullet$, COOH
7	-5	63	18	CO_2 complex, Arg-$C_\alpha^\bullet$
8	10	48	49	CO_2 complex, Arg-$C_\alpha^\bullet$
9	45	40	57	Tyr-$O^\bullet$, COOH
10	71	99	80	Tyr-$O^\bullet$….*N*-term NH_3^+, COO^-
11	68	95	76	Tyr-$O^\bullet$….*N*-term NH_3^+,COO^-
12	86	125	115	NH_2–CO_2 adduct, Arg-$C_\alpha^\bullet$
13	95	130	127	NH_2–CO_2 adduct, Arg-$C_\alpha^\bullet$
14	107	142	135	*N*-term. $NH_2^{\bullet+}$, COO^-
15	121	146	135	*N*-term. $NH_2^{\bullet+}$, COO^-
16	141	179	178	*N*-term. $NH_2^{\bullet+}$, COO^-
17	108		158	Tyr aromatic cation-radical
18	141			Tyr aromatic cation-radical

[a]In units of kJ mol^{-1}; [b]Including B3LYP/6-31+G(d,p) zero-point energies; [c]Single-point energies with the 6-311++G(2d,p) basis set on B3LYP/6-31+G(d,p) optimized geometries. [d]Single-point energies with the 6-311++G(2d,p) basis set on M06-2X/6-31+G(d,p) optimized geometries. [e]Single-point energies with the 6-311++G(2d,p) basis set on ωB97X-D/6-31+G(d,p) optimized geometries.

Table 3.4. Energies of Select Reactions.

Reaction	Relative Energy[a,b] B3LYP[c]	ωB97X-D[d]	PMP2[c]
$Cu(terpy)^+ \rightarrow Cu(terpy)^{2+}$	987 (10.23)[e]	987 (10.23)[e]	-
Neutral **1** → Cation-radical **1**	554 (5.74)[e]	557 (5.78)[e]	-
$Cu(terpy)(YAAAR)^{2+} \rightarrow Cu(terpy)^+ + YAAAR+$	-6	-	
1 → $^{\bullet}GAAAR^+ + C_7H_6O$	82	156	161 (181)[f]
1 → TS(loss of C_7H_6O)	137	154	129 (181)[f]
1 → [$^{\bullet}GAAAR^+ + C_7H_6O$] complex	41	84	87 (107)[f]
Loss of CO_2 from **8**	12	33	-
Loss of H-2 from **5**	207	217	180 (179)[f]
TS for loss of H-2 from **5**	219	239	208 (198)[f]

[a]In units of kJ mol^{-1}; [b]Including B3LYP/6-31+G(d,p) zero-point energies; [c]Single-point energies with the 6-311++G(2d,p) basis set on B3LYP/6-31+G(d,p) optimized geometries. [d]Single-point energies with the 6-311++G(2d,p) basis set on ωB97X-D/6-31+G(d,p) optimized geometries. [e]Ionization and recombination energies in electron volts. [f]Values in parentheses are from restricted open-shell (ROMP2) single point energy calculations.

Figure 3.4. CID spectra of m/z 550 cation-radicals (a) $YAAAR^{+\bullet}$ and (b) $AAAYR^{+\bullet}$.

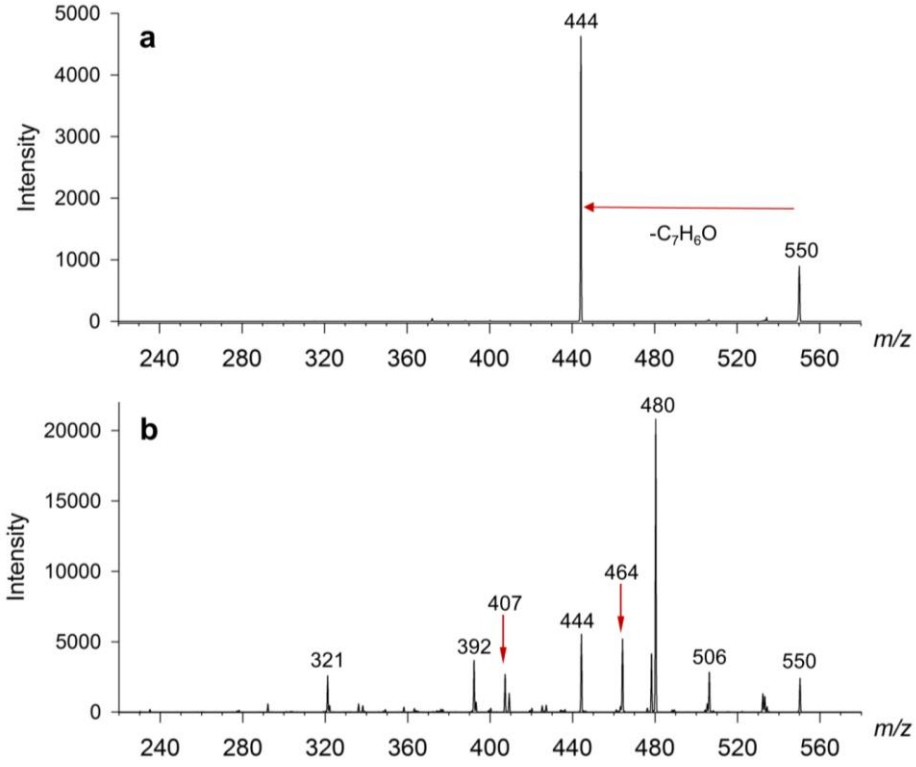


Figure 3.5. CID spectra of cation-radicals (a) AAYR$^{+\bullet}$ and (b) d11-AAYR$^{+\bullet}$.

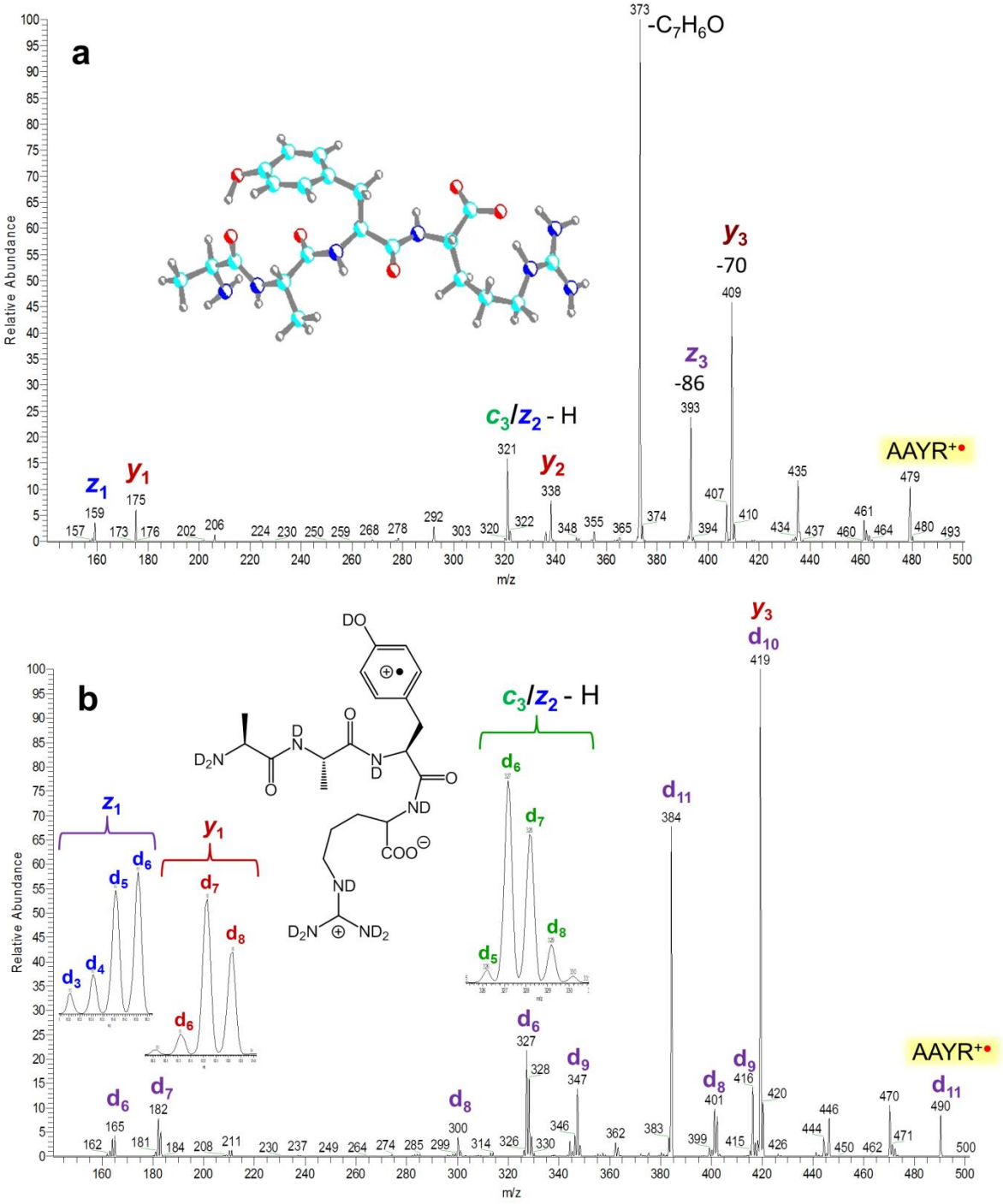


Figure 3.6. UVPD spectrum (5 laser pulses) at 355 nm of cation-radical d11-AAYR$^{+\bullet}$. For fragment ion assignment see **Figure 3.5b, SI**.

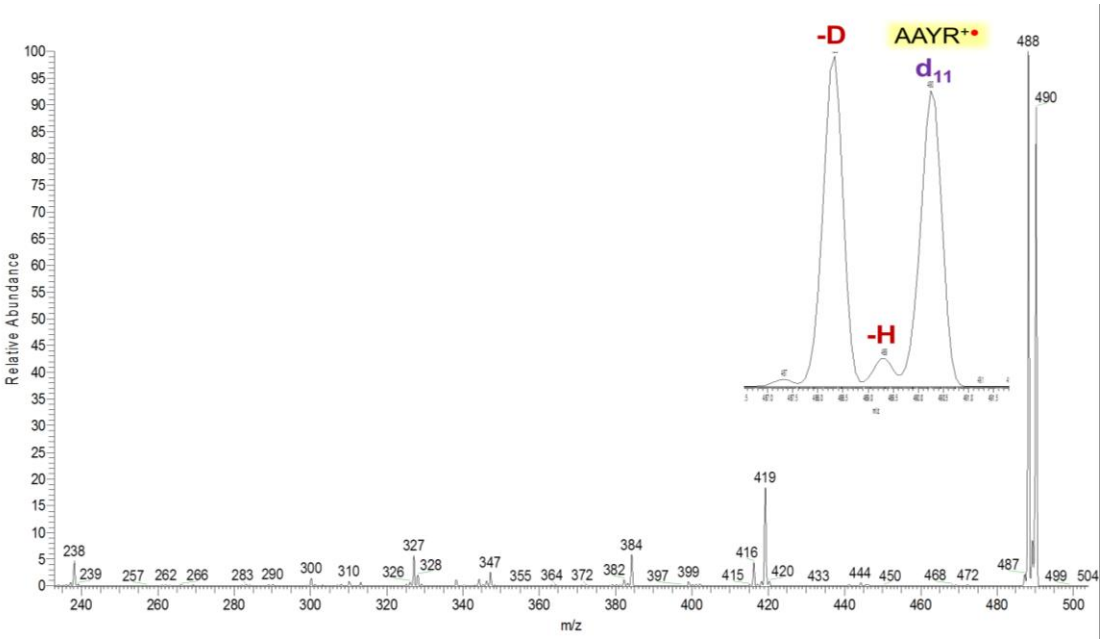


Figure 3.7. Photodissociation mass spectra of (a) YAAAR+● and (b) AAAYR+● at 355 nm following 10 laser pulses at 15 mJ/pulse.

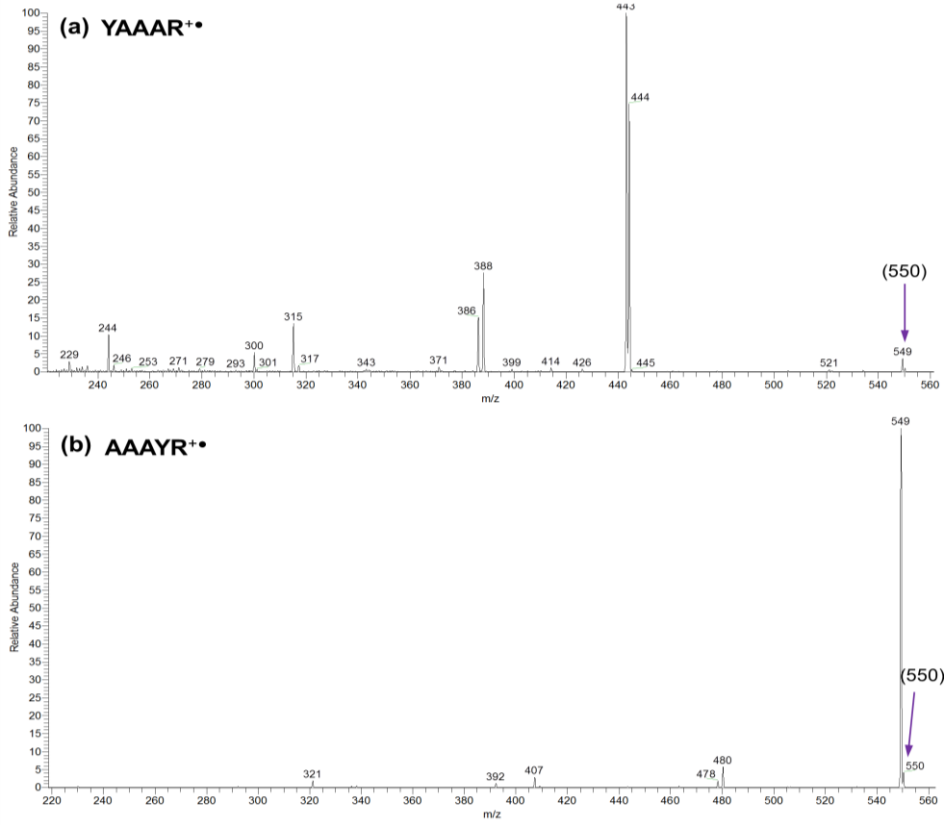


Figure 3.8. Laser pulse dependence of parent and fragment ion intensities in photoddisociation of mass-selected ions at 355 nm and 15 mJ/pulse laser power. (a) YAAAR+●, (b) AAAYR+●.
The dashed lines show the best exponential fits, $I(n) = I(0)e^{-(0.7367n+0.004)} + 0.0035$, and $I(n) = I(0)e^{-(0.329n+0.113)} + 0.0175$, for YAAAR$^{+\bullet}$ and AAAYR$^{+\bullet}$, respectively.

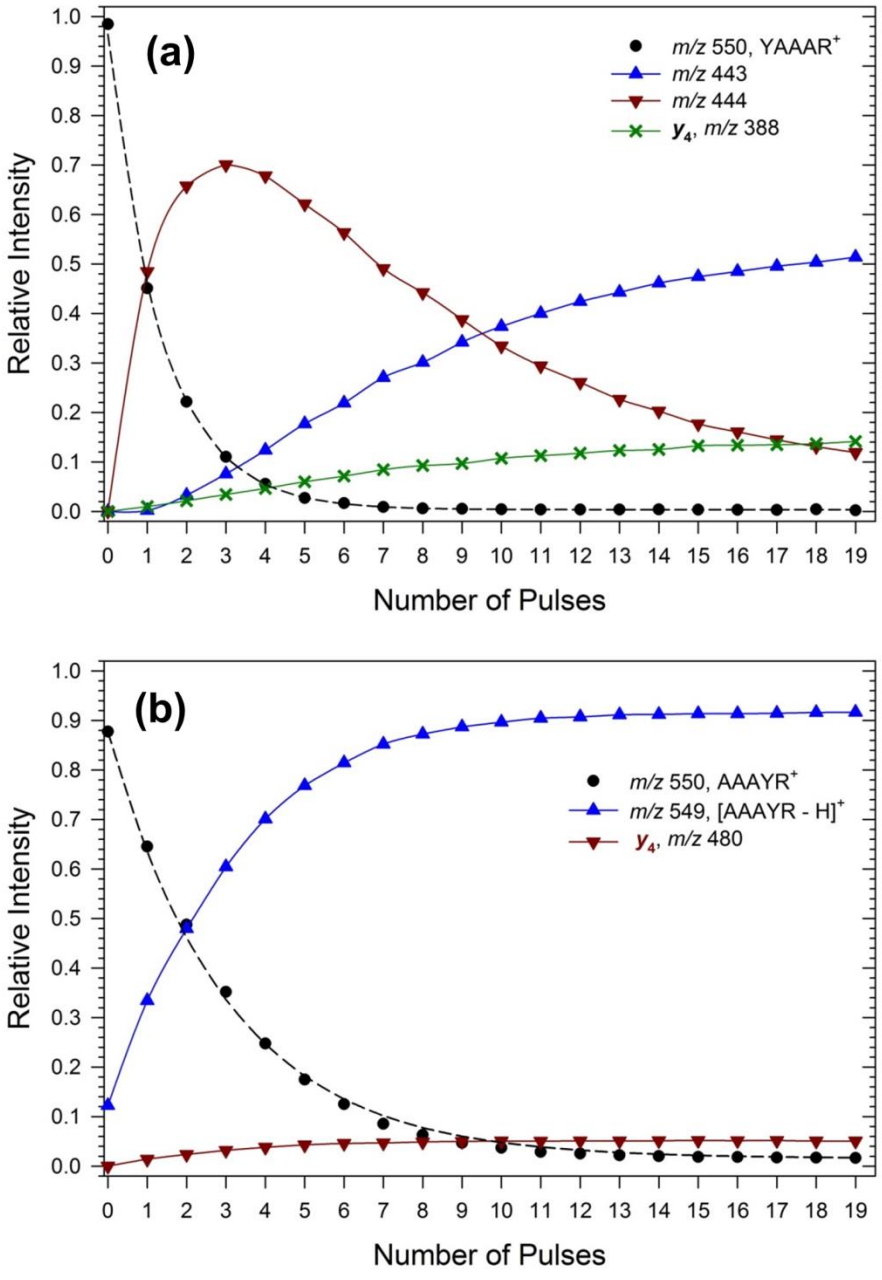


Figure 3.9. ωB97X-D/6-31+G(d,p) optimized structures of YAAAR+● conformers.
Only exchangeable (O-H, N-H) hydrogens are shown. Atom color-coding: cyan=C, red=O, blue=N, gray=H.

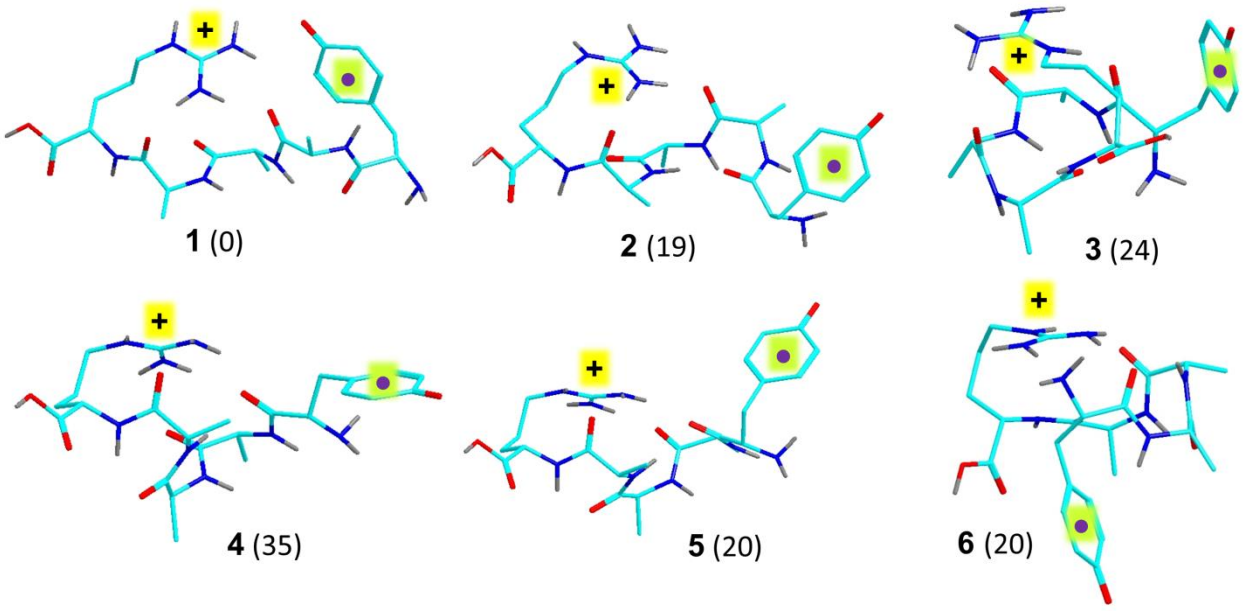


Figure 3.10. Molecular orbitals (ωB97X-D/6-311++G(2d,p)) involved in electron excitation and photon absorption of YAAAR+● conformer 1 at 360 nm. The numbers show transition vector amplitudes.

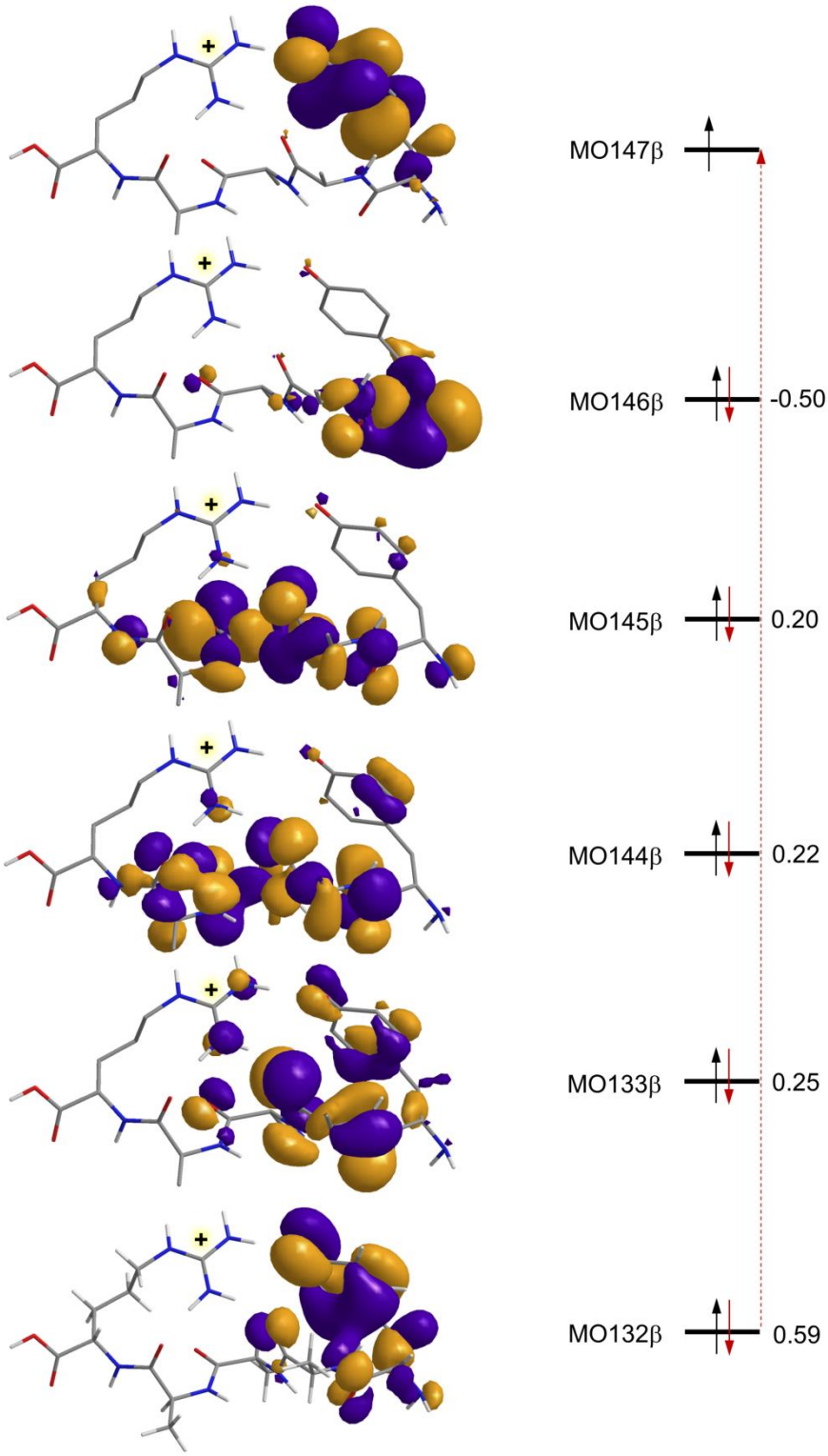


Figure 3.11. ωB97X-D/6-31+G(d,p) optimized structures of AAAYR+● cation radicals **5-18**. Italic numbers show the major spin densities.

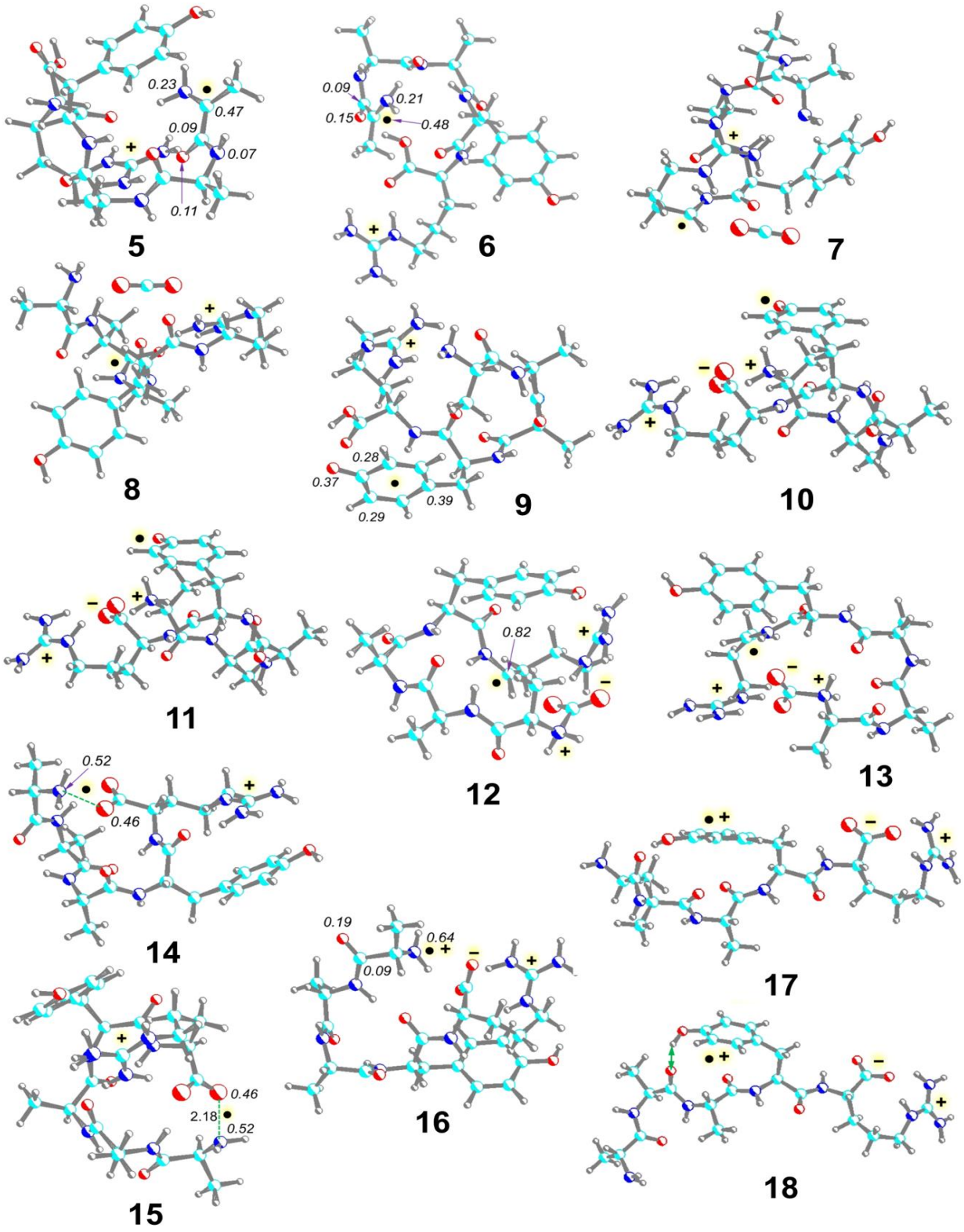


Figure 3.12. CID spectrum of [Cu(tpy)AAAYR]2+● (m/z 423) showing the formation of the AAAYR cation radical at m/z 550. Inset shows the minor fragment ions at m/z 444-506.

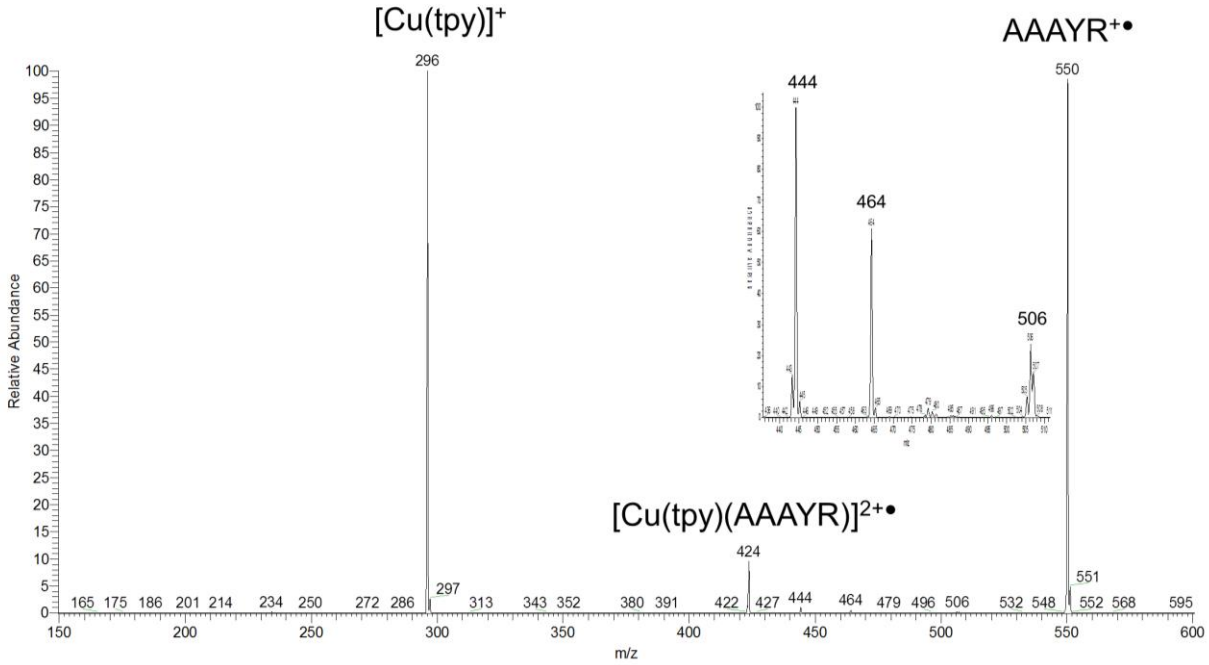


Figure 3.13. ωB97X-D/6-311++G(2d,p) calculated absorption spectra of AAAYR+● isomers 5, 6, 8, 9, 12, 14, 16, and 17.

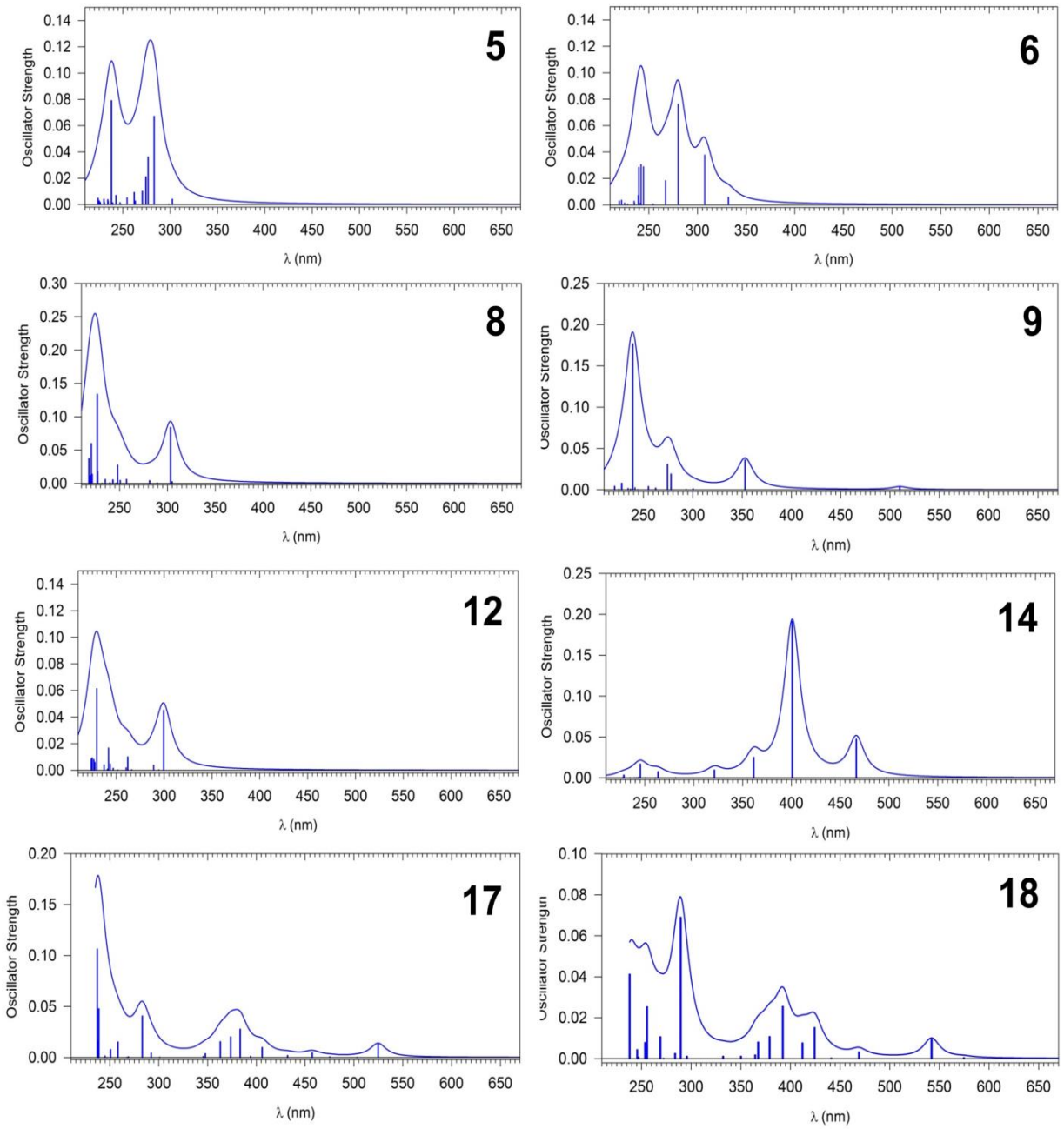
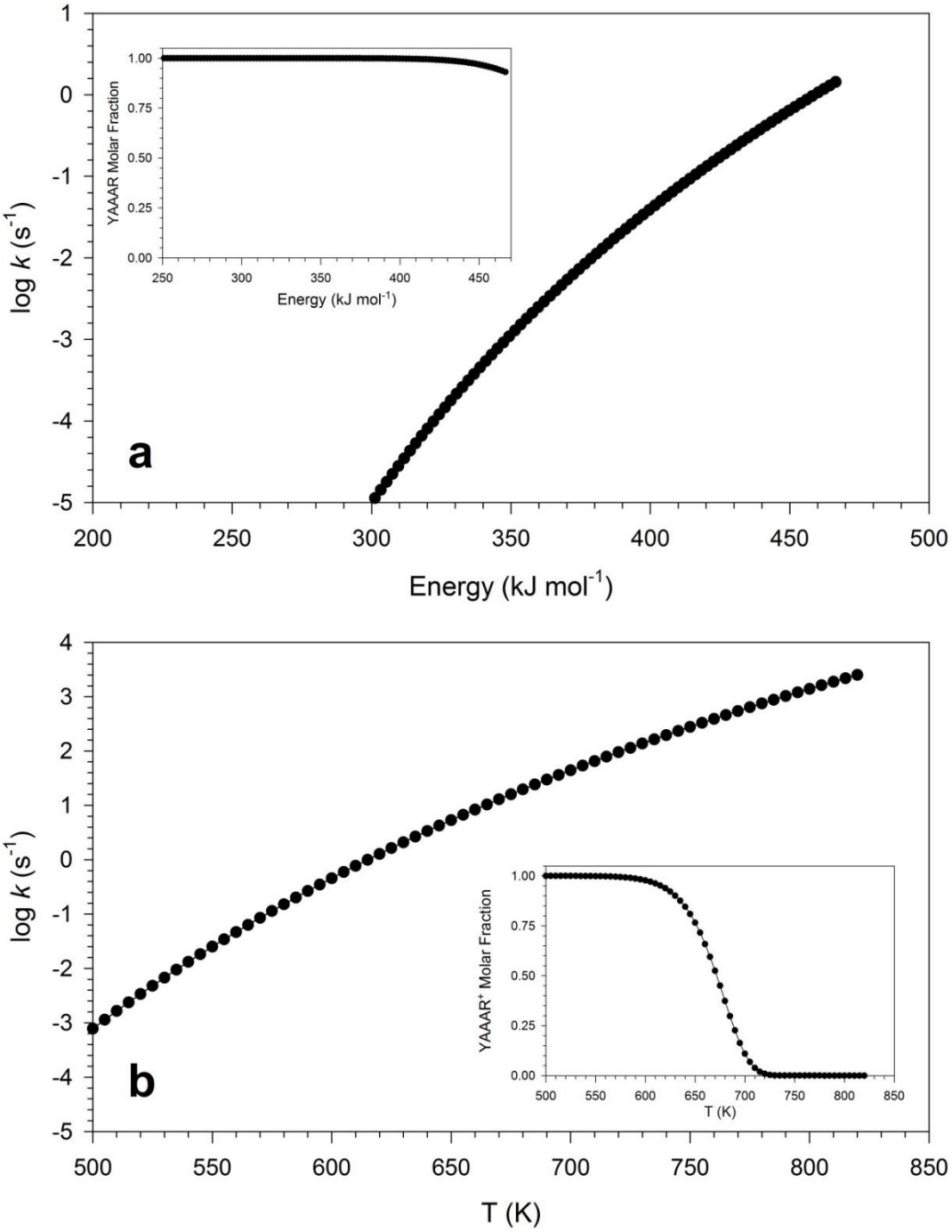
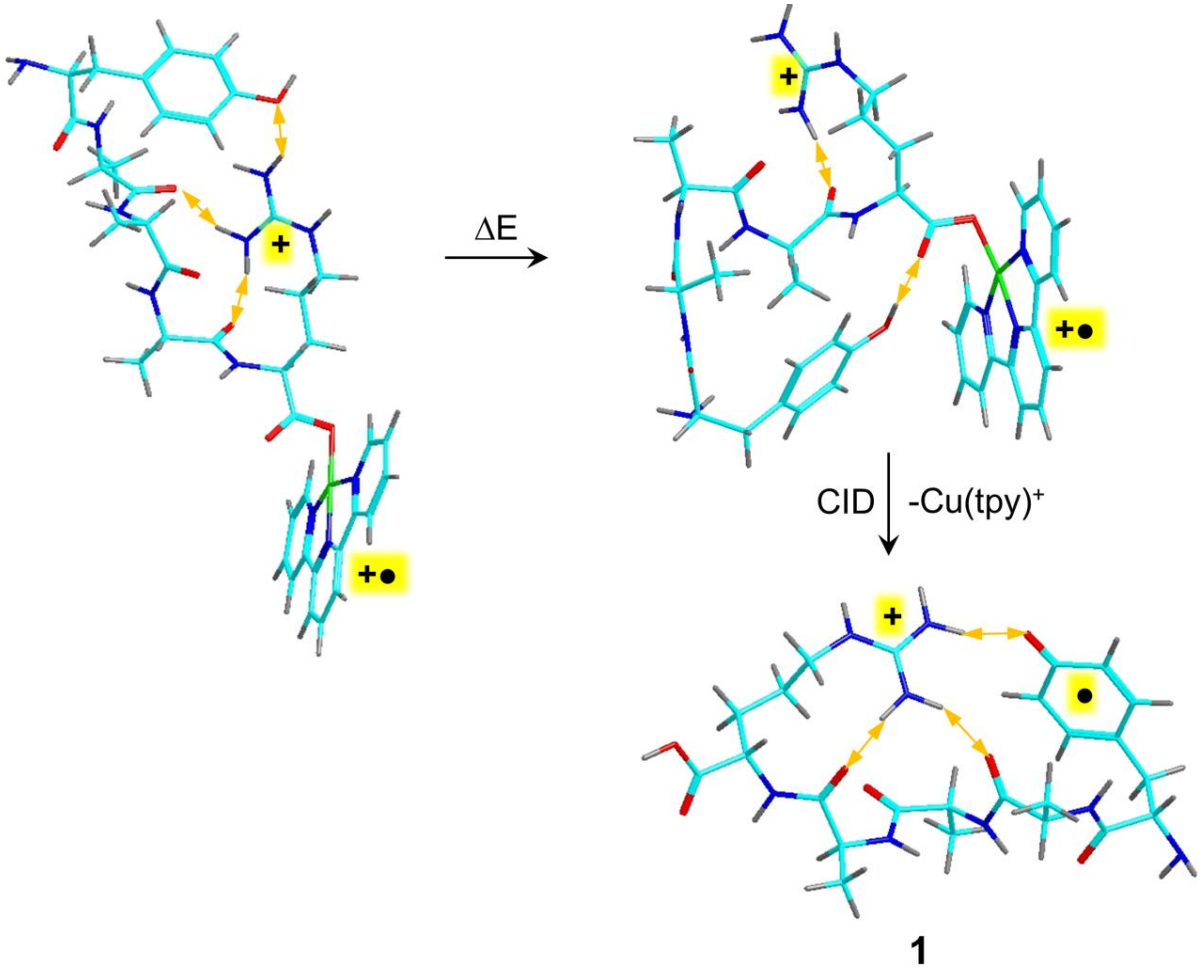


Figure 3.14. (a) RRKM calculated rate constants for C7H6O loss from 1. (b) TST calculated rate constants for the same reaction. Insets show the molar fractions of non-dissociating $YAAAR^{+\bullet}$ at 50 ms reaction time. Both calculations used the ωB97X-D/6-311++G(2d,p) + ZPVE transition state energy of 154 kJ mol^{-1} from **Table 3.4 (SI)**.



Scheme 3.2. Conformational transformation in $[Cu(tpy)YAAAR]^{2+\bullet}$ followed by intramolecular electron transfer oxidation of the peptide and dissociation to $YAAAR^{+\bullet}$ and $[Cu(tpy)]^{+}$.
Structures from ωB97X-D/6-31+G(d,p) geometry optimizations. The yellow double-ended arrow indicate major hydrogen bonds.



Chapter 4 Exploiting the Large Mass-Defect of Iodine in the Development of a Charge Tag for Enhanced Sequencing of Synthetic Peptides

4.1 Introduction

The *de novo* approach to sequencing has many advantages over bottom-up sequencing, such as a higher reliability of peptide identification,[235] however, its usefulness is impaired by the frequent incomplete backbone cleavage and resulting ion series.[160,236] In order to address this shortcoming, various sequence tags have been developed to aid in sequence identification[237-240] The exploration of a suitable tag for the direct enhancement of peptide sequence coverage rests on the need for *de novo* sequencing. Unlike bottom-up proteomics, which rely on spectral database searching and matching, *de novo* sequencing relies on the analytical process of deriving a peptide's amino acid sequence from its experimentally obtained mass spectrum (ms^n) with no reference to sequence databases.[241] As such, the main advantage of *de novo* sequencing lies in its ability to lead to the identification of both known and novel peptides.[163] The use of tandem mass spectrometry (whereby the product ion from a fragmented precursor ion can itself be isolated and further fragmented) is crucial for *de novo* analysis. A peptide is isolated out of a mixture (ms^2) after UPLC separation and dissociated via one out of several ion activation methods; commonly, collision-induced dissociation (CID) that yields b- and y-ions and electron transfer dissociation (ETD), which yields c- and z-ions are used for the study of peptide sequences. The latter, ETD, is of particular interest in the sequencing of post-translational modifications of proteins and peptides as the method can help locate the PTMs on a particular sequence and it leaves these modifications unaltered, unlike CID. [129, 242] Regardless of the method of choice to cleave the backbone, *de novo* sequencing makes use of the mass difference between two fragment ions to match it to the mass of one amino acid residue of the peptide backbone.[161] *De Novo* sequencing can not only be useful in the analysis of a novel peptide, for which no spectral comparison via database is available, but it can also be used when a database is available so as to validate the peptide identification obtained via spectral comparison.[161,243] Moreover, it can reveal the presence of novel peptide mutations for example.[150,240,244,245] Additionally, though the manual approach is thought of as a slow method of peptide determination, commercially available software can assist in *de novo* analysis and help speed the analytical process. [162,163,246]

4.2 Experimental Section

4.2.1 *Materials*

All synthetic pentapeptides (AAXAK, where X = F, K, D, N, H) were purchased from GeneScript (Piscataway, NJ, USA) and each diluted before use to make 10 mg/mL solutions in MeOH/H_2O (1:1). Bovine Serum Albumin was purchased from Sigma-Aldrich (Milwaukee, WI, USA). Tyramine hydrochloride and potassium iodide were purchased from Alfa Aesar (Heysham, UK). Iodine, H-PAL ChemMatrix® and di-(2-propyl) thionocarbonate were purchased from Sigma-Aldrich (Milwaukee, WI, USA). Common reagent grade chemicals such as N,N-dimethylformamide and trifluoroacetic acid were purchased from Sigma-Aldrich (Milwaukee, WI, USA). Sodium acetate, sodium periodate, sodium chloride, sodium thiosulfate, acetic acid and benzylamine were all purchased from Fisher-Scientific (Fair Lawn, NJ, USA). Benzylamine was flash distilled before use. The disposable reaction vessels (3-mL fritted syringes) were purchased from CSPS Pharmaceuticals (San Diego, CA, USA).

4.2.2 *Instrumentation*

4.2.2.1 Development of Tag Synthesis Using Benzylamine (MS)

The identity of the products obtained during the synthetic scheme development (coupling of tag to benzylamine) was screened and verified using both low and high resolution Thermo-Fisher instruments (San Jose, CA). Low resolution measurements were made on a commercial Thermo LTQ-XL with ETD capability and modified with a Nd:YAG Ekspla laser system, as described in the previous chapters. Capillary nanospray made of fused silica capillary tubing with polyimide coating (Fisher-Scientific, 360 μm OD, 75 μm ID, cut to 8 - 10 cm) and tapered at the tip using a laser puller was used to introduce the samples in the mass spectrometer via a custom-built source. The flow rate was set at 2 μL/min, High resolution measurements were made on an LTQ-XL-Orbitrap instrument at 100,000 resolution. Samples were introduced in a 12 μL loop connected to a divert/inject valve, to which an injection port and LC pump are also attached. Once the sample was loaded onto the loop, the valve was switched to the Inject position and the ACN/H_2O (1:1, with 1% FA) solvent flow carried the sample to the ion source in ESI mode. A 5 μL/min flow rate was used across experiments.

4.2.2.2 Application of Developed Tag Using Synthetic Peptide Mixture (MS and LC)

The LC -MS^2 (ETD) measurements of the tagged synthetic peptides were performed on a Thermo Fisher Orbitrap-Fusion mass spectrometer coupled to a Waters nanoAcquity LC system. Set-up for the chromatography and mass spectrometry are described below.

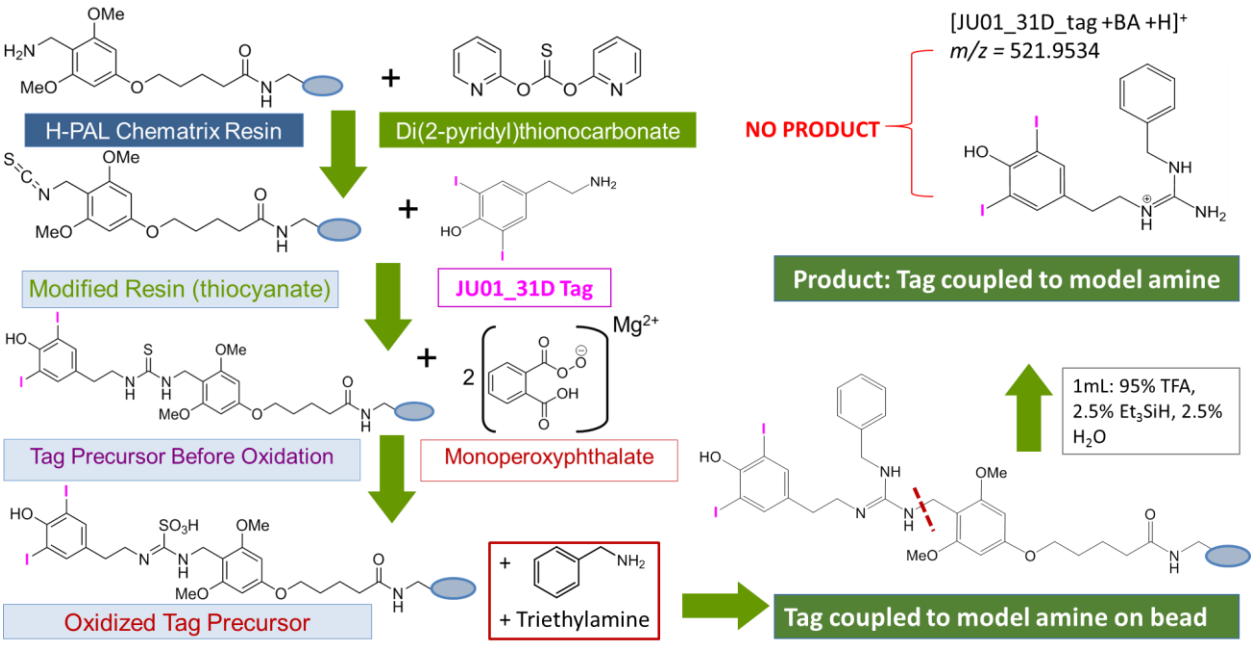
Mass spectrometry. The ETD chemical ion source utilizes fluoranthene as the radical anion donor necessary for the transfer of one electron to a multiply charged ion. Capillary nanospray was used to introduce the samples in the mass spectrometers. Fused silica capillary tubing with polyimide coating (Fisher-Scientific, 360 µm OD, 75 µm ID, cut to 8 -10 cm) was tapered at the tip using a laser puller. The spectra were collected in positive ion mode, the temperature of the ion transfer tube was set at 350 °C, and the mass range was set to normal (*m/z* = 200-1600).

Liquid chromatography. The tagged products for each peptide in the mixture were separated using C-18 reversed-phase chromatography. A pre-column (IntegraFrit™ CAPILLARY, 360 µm OD, 100 µm ID, 50 cm, with frit) and column (Fisher-Scientific fused silica capillary tubing with polyimide coating, 360 µm OD, 75 µm ID, cut to 20 cm) were packed in-house using packing material from Bruker-Michrom (Auburn CA, USA). The pre-column was packed using Magic 5µm 200 Å C18 AQ beads and the column was packed using Magic 5µm 100 Å C18 AQ beads; both bead types were prepared as a slurry of isopropanol/methanol/water (2:2:1). The total time for the complete elution of one sample (3 µL) injection was 90 min for the gradient elution described as follows. The liquid chromatography (LC) was run with a binary solvent: aqueous solvent A is H_2O/formic acid (99.9/0.1) and organic solvent B is ACN/formic acid (99.9/0.1). The sample was trapped at 2 µL/min with 98% A and 2% B for 10 minutes and then separated at 0.3 µL/min according to the following gradient: 1 - 60 min 5% B - 30% B, 61 – 70 min, 80% B – 80% B, 71 – 90 min, 80% B - 2%B.

4.2.3 *Overall Synthetic Scheme Development*

The original synthetic scheme (**Scheme 4.1**) for the iodinated charge tag was modeled after that of former group member, Dr. Chang Xue.[247] The guanidination method described here is identical to that used for her alkylated guanidine tags. In a fritted 4-mL syringe the H-PAL Chematrix® substrate is modified into a

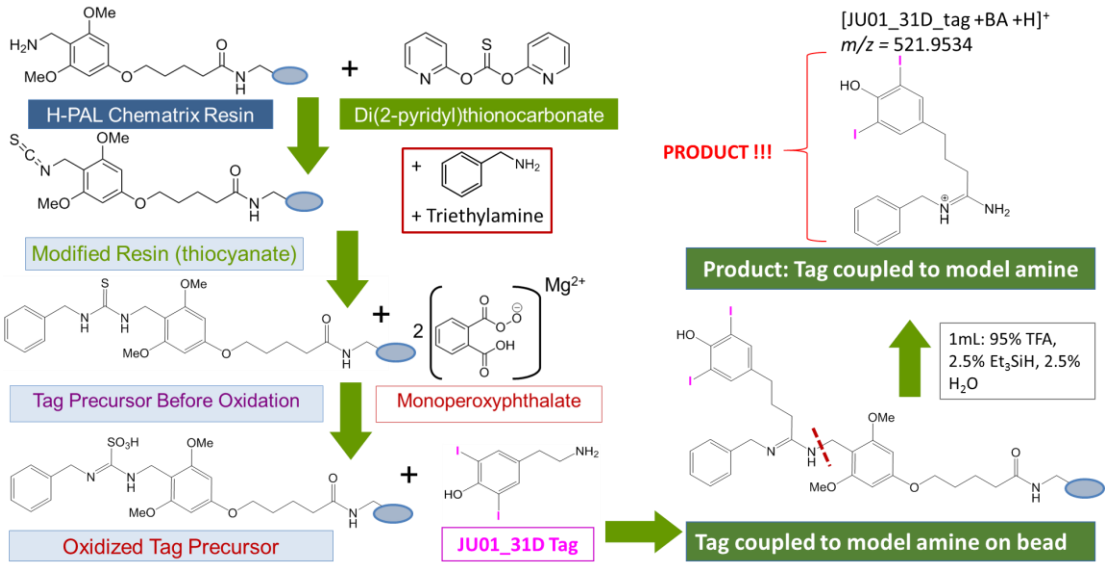
thiocyanate group at the free primary amine using an excess of di(2-pyridyl) thionocarbonate. At this step of the reaction, was originally added an excess of tyramine diiodinated at the ortho positions. This 2,6-diiodinated tyramine was synthesized in-house by one of our collaborators, Jan Urban, and its primary amine was reacted with the previously prepared thiocyanate moiety of the substrate to form a thiourea analog.



Scheme 4.1. Original Synthetic Scheme

The subsequent steps of the reaction involved oxidizing the thiourea analog, or "precursor tag" as it is referred to below, adding in an incoming amine in the presence of a base (benzylamine in triethylamine here for coupling testing) and thus completing the synthesis of the tag before cleaving it from the resin bead. When following this reaction scheme, an expected singly charged product at *m/z* 521.9534 was, however, not observed. Since tag formation and amine coupling was previously observed by Dr. Xue when developing her alkylated charge-tags,[247] it was hypothesized that the iodinated tyramine may not survive one of the synthetic steps. Careful reviewing of the scheme helped identified two particularly harsh steps in the overall scheme, which might cause deterioration of the tag. The first potentially detrimental step to the tag formation was surmised to be the oxidation step, where an excess of monoperoxyphthalate is added to the precursor tag. The second potentially harmful reaction was thought to be at the cleaving step, where a solution of

trifluoroacetic acid (TFA) is reacted with the tag fully built on the resin bead. In order to diagnose which of the two steps may be interfering with the preservation of the tag, if either, the addition order for 2,6-diiodotyramine and benzylamine were swapped, as described by the diagnostic scheme below (**Scheme 4.2**).



Scheme 4.2. Diagnostic Synthetic Scheme

Since the addition of 2,6-diiodotyramine then took place after the oxidation step, two possible outcomes would result: 1) a singly charged product formation at *m/z* 521.9534 would be observed, thereby indicating that the oxidation step was detrimental to the tag integrity when iodinated earlier, or 2) the product would not form, thereby indicating that the cleaving step was solely involved in the degradation of the tag. As seen from the ESI spectrum resulting from the diagnostic reaction synthesis, the expected productwas indeed made and a clean, singly charged peak at *m/z* 521.9435 corresponding to the iodinated tag coupled to benzylamine was observed. This suggested that the iodinated precursor tag did not survive the oxidation step and that the iodination of the precursor tag should occur after that step in order to ensure its integrity. Since the ultimate goal is for this tag to provide a labeling tool for tryptic peptides, it is crucial to this synthesis that the amines from a given digest be introduced in the last step of the synthesis for coupling to the tag.

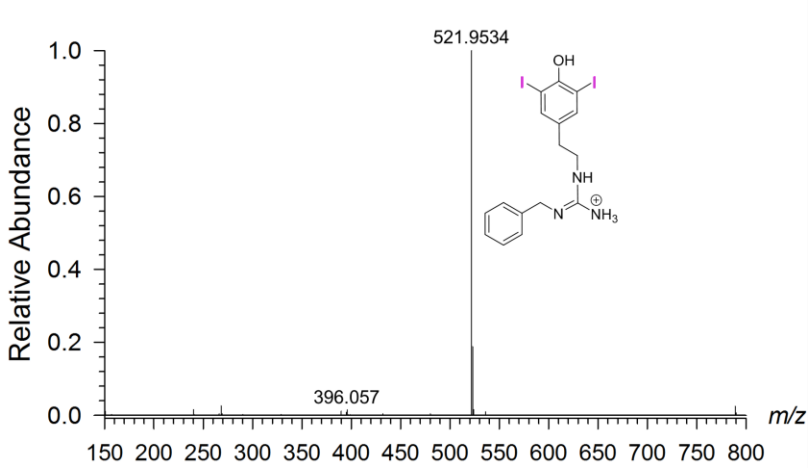
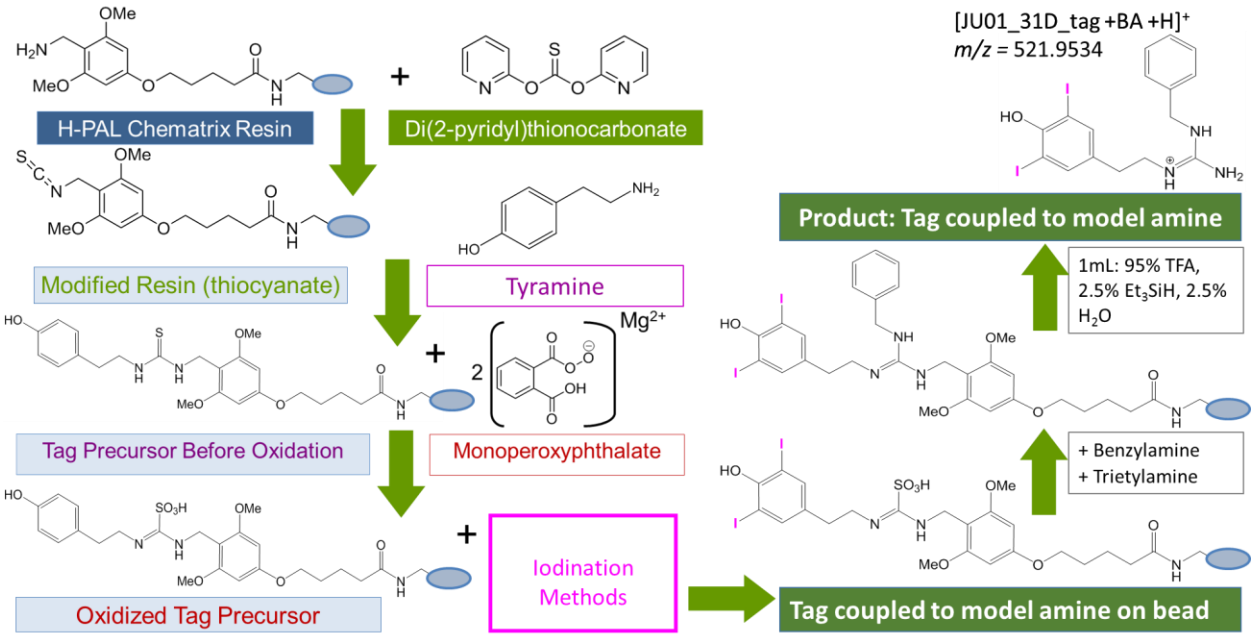


Figure 4.1. Tag Product obtained from the diagnostic synthetic scheme.

The need for this criterion to be satisfied rests on the undesirability for these tryptic peptides to undergo any chemical change other than their mere coupling to the precursor tag. This also implied that although the diagnostic scheme was helpful in identifying the source of the problem in successfully synthesizing the tag, it would not be useable as it stood for the actual coupling to tryptic peptides since the "model peptide" (benzylamine in this case) was added just before the oxidation step. Whilst redesigning the synthetic scheme for the tag, it was decided that tyramine would be used as the backbone for the tag and that the iodination would be done in-house in an additional step following the oxidation of the precursor tag (**Scheme 4.3**). The hydroxyl group on the benzene ring of tyramine would predictably direct the electrophilic aromatic substitution with iodine at the ortho positions,[248,249] and the iodination reaction would be optimized to obtain a high yield of the diiodinated product (>90-95%) compared to the monoiodinated product (<5-10%). This extra step in the reaction synthesis was developed as described below.



Scheme 4.3. Updated and final synthetic scheme for iodinated charge-tag synthesis.

4.2.4 *Iodination Method Development*

A wealth of methods for the mono-, di-, and triiodination of aromatics, and more specifically phenols, have been published over the last four decades and have been motivated by the biological importance of these compounds, as evidenced by thyroid hormones for instance.[250] All these methods were described for the bulk liquid synthesis of iodinated aromatic compounds. During the development of an adequate iodination method, scalable to the micromolar quantities required for our solid phase synthesis, several of these methods were deemed suitable to our needs based on the ease of the chemical reaction, cost and availability of reagents. Seven different reaction methods[251-260] were originally scaled down and tested for their potential to diiodinate tyramine (12 µmol) before its coupling to benzylamine. Out of all the above-mentioned methods, three showed to be particularly promising; the first method used a combination of $NaClO_2$/NaI/HCl reagents as an efficient, inexpensive and mild approach,[251] while the second method involved a mixture of $NaIO_4$/KI/NaCl meant to generate iodine monochloride (ICl) *in situ,* as a source of electrophilic I^+ ions.[252] As illustrated in the electrospray spectra obtained for the products of these two reactions, the precursor tag was iodinated, but a significant portion was converted into its monoiodinated

form relative to its diiodinated form (**Figure 4.2 a-c**). The first method, which used sodium chlorite, showed a slightly higher conversion into the monoiodinated product than the diiodinated one (1:0.85), while the second method, which uses sodium periodate, showed a lower conversion into the monoiodinated product and a slight preference towards the diiodinated product (0.9:1). Last but not least, the third method involved a direct iodination of the tag using excess Iodine (I_2) dissolved in acetic acid. While this direct iodination of the precursor tag yielded no product when iodine was solely dissolved in acetic acid, the reaction benefited tremendously from dissolving iodine in a buffered solution of acetic acid and sodium acetate (**Figure 4c**). Based on those results, efforts were concentrated on improving the second and third method. It is worth noting that several mechanisms for the Iodination of phenols have been proposed in the past, and many agree that an important step in the iodination of a phenol likely involves the temporary deprotonation of the hydroxyl group.[261] Seeing that the reaction may benefit from the presence of a base, syringes containing the precursor tag were first rinsed with a solution of sodium acetate (0.4 M; 0.5 mL, twice).

Attempts to drive these individual reactions towards a higher yield of diiodinated product by varying the concentration of sodium acetate in the buffer system were unsuccessful. Nevertheless, a final endeavor proved fruitful when the two iodination methods were used in series; a direct tag iodination in acetic acid/sodium acetate buffer immediately preceding a modified $NaIO_4$/KI/NaCl method (in sodium acetate buffer rather than acetic acid alone) proved to be the key recipe –described in details below. This serendipitous finding provided a high yield of diiodinated product (>95%) and was successfully reproduced thrice. It should be noted that while the method worked reliably, it was never fully optimized.

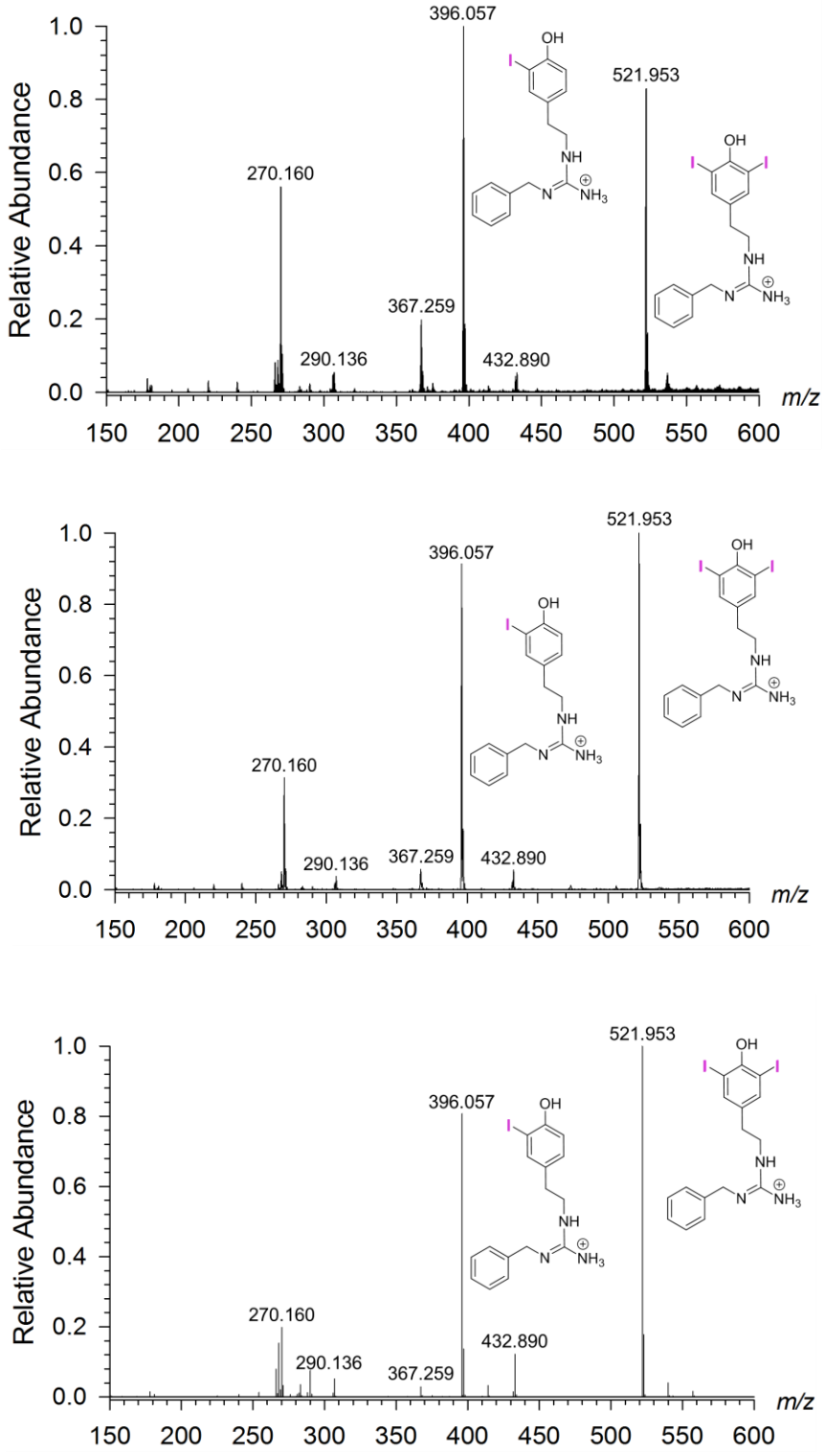


Figure 4.2. Product of the iodinated charge-tag synthesis using a) NaClO2/NaI/HCl method, b) NaIO4/KI/NaCl method and c) I2 in NaOAc buffer method.

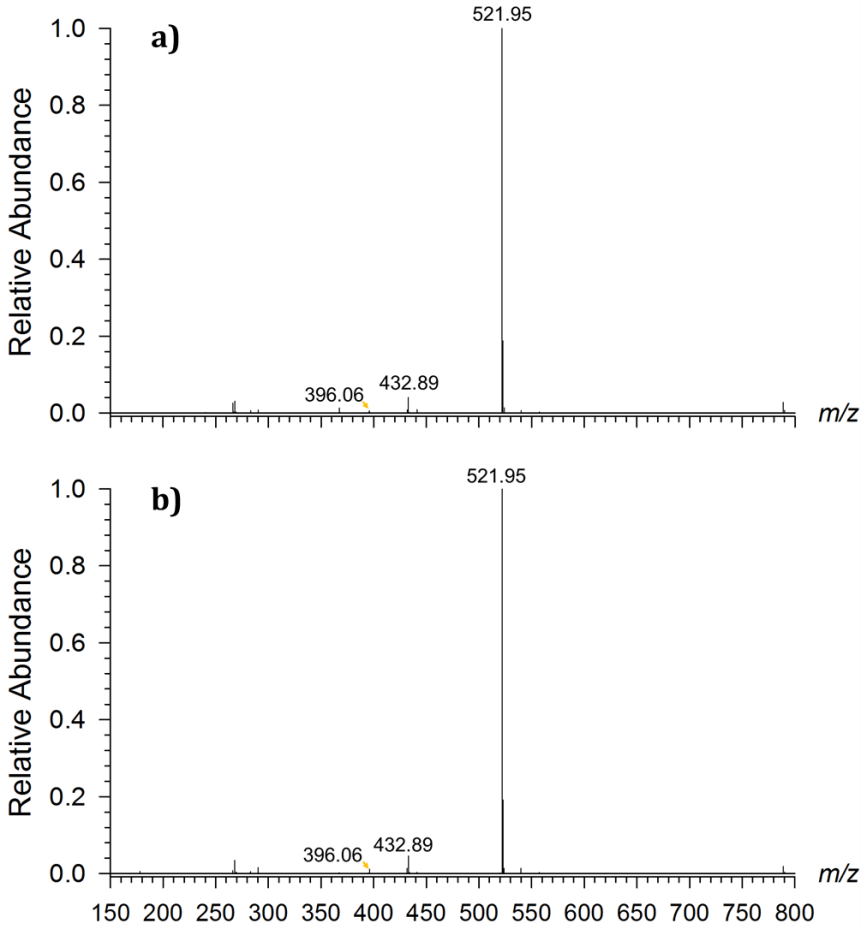


Figure 4.3. Double iodination products: (a) product obtained using a combination of I2 dissolved in acetic acid and sodium acetate buffer followed by the NaIO4/KI/NaCl/NaOAc method; (b) the reproducibility of the method was validated on a different day.

4.2.5 *General Procedure for the Preparation of the Diiodotyramine Tag Precursor*

H-PAL ChemMatrix® resin (25 mg, 12 µmol) was weighed out directly in a tared frit-equipped syringe and allowed to swell in 1 mL of dimethylformamide (DMF) for 1h. A needle was placed on the tip of the syringe and the solvent was discarded through it into a waste container, while the resin beads remained in the fritted syringe. Next, di-(2-propyl) thionocarbonate (DTP) (11 mg) was diluted in 0.5 mL DMF and aspirated into the syringe. The syringe was then allowed to react for 1h while mixing by rotation at 100 rpm. The DMF solvent and excess DTP reagent were then discarded through the needle and further washed off

with DMF in 0.5 mL portions, five times. Tyramine hydrochloride (26.4 mg, 120 μmol) was dissolved in 0.5 mL DMF and triethylamine was added (9.6 μL). The solution was aspirated into the syringe and allowed to react while mixing for 2h. After discarding the excess reagent and solvent through the needle, the beads were washed with DMF (0.5 mL portions, five times). Magnesium monoperoxyphthalate hexahydrate (21.4 mg, 43.2 μmol) was dissolved in 0.25 mL of DMF before the solution was aspirated in the syringe and allowed to react while mixing for 2h. Note that following each of these reaction steps and subsequent washes, the beads in the syringe can be kept in DMF in a -4°C fridge or a -20°C freezer for up to a few days without affecting the stability of the reagents. The syringe was then washed with DMF (0.5 mL portions, five times) and with 0.4 M solution of sodium acetate (0.5 mL portions, four times). At this stage of the reaction, the tag precursor is ready to be iodinated. Iodine (50 mg, 197 μmol) was dissolved in 4 mL glass vials in glacial acetic acid (900 μL) at 35° C overnight and in the dark before adding deionized water (100 μL). Once brought to room temperature, the iodine solution was introduced into the syringe and allowed to react while mixing for 2 h. It is critical that the syringe be covered with aluminum foil so as to prevent photodegradation. Following this first iodination step, a second iodination step immediately followed with no washing between the two. This important point must not be disregarded as it was shown to affect the yield of the di-iodinated product relative to the mono-iodinated product. The second iodination solution consisted of a mixture of 15 mg/mL solution of sodium periodate (307 μL), 300 mg/mL solution of sodium chloride (140 μL), 0.2 M sodium acetate solution (123 μL), all of which using methanol/water (1:1) as the solvent, 50 μL of acetic acid in water (9:1) and a 10 mg/mL solution of potassium iodide in water (200 μL). Potassium iodide was added last, before the solution was quickly aspirated in the syringe and allowed to react while mixing for 3 h. Dark clumping may be observed in the syringe and is normal. Subsequently, the reaction solution was discarded through the needle and any excess was washed off with a 1.0 M solution of sodium thiosulfate in 0.5 mL portions, discarded into a clear beaker, until any remnant of yellow tint disappeared from the waste, suggesting complete neutralization of the excess iodide. Four more washes ensued, this time with a solution of methanol/water (8:2) in 0.5 mL portions (five minutes rotation each). At this stage of the synthetic scheme, the tag synthesis is complete and ready for coupling with an incoming primary amine in the presence of triethylamine. Detailed description of the reaction for various amines is found below. After the reaction between the tag and the chosen amine reached 64h, the syringe was thoroughly washed with five portions of

methanol/water (8:2) as described above and subsequently washed another five times with diethylether. The syringe, now containing the bead-bound synthesized tag coupled to the primary amine, was dried down using a mechanical pump for 30 min-2 h, until dryness was reached. Lastly, the newly synthesized product was cleaved from the resin beads with a trifluoroacetic acid (TFA) cleaving solution (1 mL) consisting of TFA/water/triethylsilane (95:2.5:2.5), which was introduced in the syringe through the needle, quickly sealed with a rubber stopper at the needle tip and allowed to react for 1 h while mixing. Following the release of the product from the bead, the TFA/water/triethylsilane solution containing the product was collected in a glass vial and dried down under nitrogen stream and reconstituted in an appropriate solvent before analysis via LC-MS/MS (ETD).

4.2.6 *Use of Model Compound Benzylamine for Peptide Coupling Test*

A model amine compound, benzylamine, was used to test the successful coupling reaction between an amine and the newly synthesized guanidine tag precursor during the optimization of the synthetic scheme. Benzylamine was found to reproducibly attach to the diiodo-guanidine tag, and was therefore chosen as a method control to be used alongside other amine-to-tag coupling reactions. Benzylamine (2.6 µL, 24 µmol) was first purified by flash-distillation and diluted with methanol (0.68 mL), mixed with triethylamine (6.7 µL, 48 µmol,), and aspirated into the syringe containing the diiodinated tag on beads at the amine introduction step of the general procedure. The syringe was allowed to react at room temperature for 64 h. Next, the syringe was washed, mechanically dried and cleaved as described above.

4.2.7 *Use of Pentapeptides as Model Tryptic Peptides*

A mixture of five peptides solutions (AAKAK, AAFAK, AANAK, AADAK and AAHAK) with triethylamine (6.7 µL, 48 µmol) was then introduced in the syringe. The peptide solutions consisted of 100 µL of each 10 mg /mL peptide solution (500 µL total). This last coupling step was allowed to react in the foil-wrapped syringe for 64 h at room temperature while taped on a rotary evaporator for mixing. After completion of this last reaction step, the beads were washed with a solution of methanol/water (8:2) in 0.5

mL portions and washed five more times with diethylether to ensure complete water removal and dryness. The syringe was then dried using a mechanical pump, as previously described for the general procedure and benzylamine to tag coupling.

4.3 Results and Discussion

4.3.1 *Coupling Reaction to Mixture of Synthetic Pentapeptides*

Following the successful synthesis of the diiodotyramine tag and its coupling to benzylamine, as described above (Figure 4.1), the next step was to evaluate the potential for this tag to couple to a model peptide. A mixture of five pentapeptides was chosen as the model system mimicking a tryptic digest peptide mixture. Accordingly, the pentapeptides chosen were all lysine-terminated (C-terminus) and were made up of four alanine residues (alanine is one of the simplest amino acids structurally and it chemistry is highly predictable) and one center amino acid, which differed from peptide to peptide. All five peptides in the mixture (AAFAK, AAKAK, AADAK, AANAK and AAHAK) were successfully tagged at once and the resulting spectra for their fragmentation by ETD are displayed in Figure 4.4a-e. Untagged peptide standards were used for comparison of the ETD efficiency and sequence coverage relative to the tagged peptides, Figure 4.4f-j. The increase in number of z-ions, to include the z_1 ion missing from the underivatized peptides, and across all five peptides can readily be seen and imply a higher sequence coverage of the peptides than when untagged. Though the increase in sequence coverage was expected to result from the improved efficiency of the ETD backbone fragmentation, it should be noted that in this particular case, the efficiency of the backbone fragmentation by electron transfer occurred even though the charge state of each peptide was not increased during the tagging process (+2 untagged → +2 tagged). A distinct fragmentation pattern is observed for all five peptides, with the complete absence of the c-ion series, a signature loss of 127 Da corresponding to the loss of one iodine atom from the molecule and its associated z-ions, and finally a peak corresponding to the loss of the entire tag from the peptide.

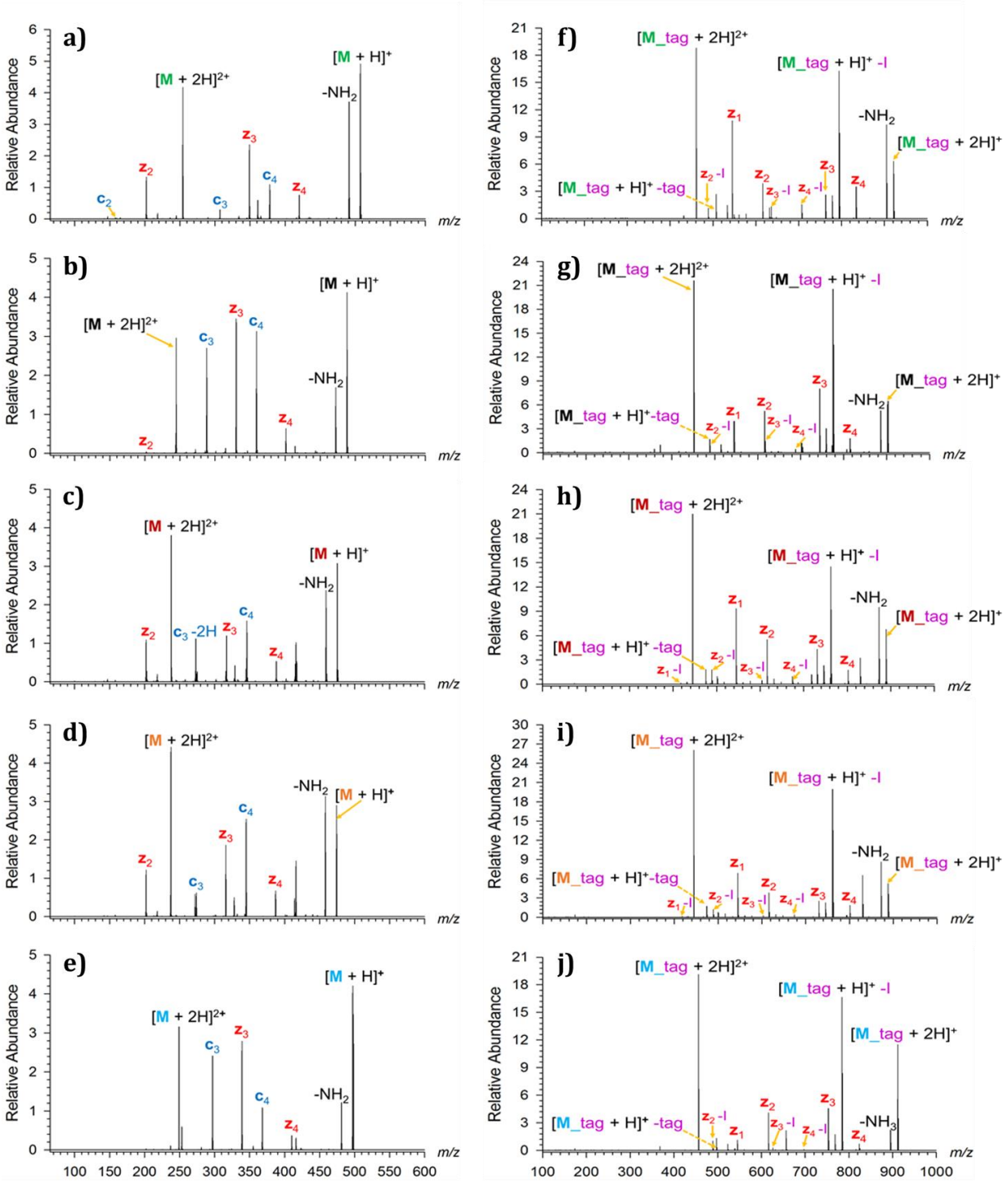


Figure 4.4. Side-by-side comparison of MS2-ETD spectra for a-e) the mixture of untagged peptides M = AAXAK (where X is F, K, D, N and H) and f-j) the mixture of peptides modified with the iodinated charge-tag. This observation may partially be due to the fact that these are short peptides and can therefore only accommodate two charges at most.

4.3.2 *Efficiency of the Method*

Though the efficiency of the tagging process was not yet evaluated for the coupling to peptides, future work should assess it as follows. All the diiodinated-tagged synthetic pentapetides were analyzed and identified by LC-MS. Since separation from the unreacted (untagged) peptides from the modified (tagged) peptides following the tagging reaction is crucial for future potential quantitation studies, it is important to test the efficiency of the tagging method. An aliquot of the five-peptide mixture should be saved as a standard before tagging and another aliquot of the tagged peptides saved after cleavage from the beads. Additionally, all washes following the peptide coupling were collected and should be analyzed for unreacted peptide content. In theory, if the solid phase synthesis method allows the complete removal of the untagged (i.e. unreacted) peptides, they should not be detected along with the derivatized peptides. Both aliquots should therefore loaded on an LC column for comparative LC-MS. The percentage of unreacted peptides remaining in the tagged sample after washing is then reported and the conversion ratio of each peptide in the mixture into a tagged peptide is calculated from the results. Future use of the tagging method should include an efficiency evaluation.

4.4 Conclusion

The development of a method for the diiodination of tyramine proved fruitful and useful as a mass defect moiety introduced to a guanidinated synthetic peptide mixture. A significant enhancement of the z-ion sequence coverage was observed, including the z_1 ion, often undetected due to low mass cut-off inherent to ion traps. These encouraging findings led to the subsequent experiments, described in **Chapter 5**, aimed at tagging peptides from the bovine serum albumin (BSA) protein for future use in *de novo* sequencing. The ultimate goal is to use the tag as a tool for efficient and straightforward sequencing of either unknown or known peptides without a need to rely on the less than ideal database search approach. If successful, this tag could be applied to standard methods in the peptide mapping of a variety of proteins.

Chapter 5 Heavyweight Champion: Iodine's Large Mass Defect Weighs in on the Enhancement of the Sequence Coverage of Digested Peptides via ETD

5.1 Introduction

Following the successful coupling of the novel diiodinated charge tag to synthetic peptides, as described in **Chapter 4**, the potential for the tag to be coupled to a mixture of peptides from tryptic digests was assessed. While several studies have been published on the enhancement of electron transfer dissociation efficiency using peptide tagging,[98,267-270] there is, to date, no existing report of a mass defect-containing charge tag that selectively targets the C-terminal lysine side-chain of digested proteins. The value of this tag lies in its potential for distinguishing isobars at both the MS^1 and MS^2 levels using high mass accuracy measurements. The concept is illustrated for MS^1 in **Figure 5.1**, where compounds of the same nominal mass (i.e. *m/z* 438 here), can be easily distinguished at high resolution thanks to the mass shift provided by the mass defect, as contributed by the two iodine atoms in each tag. In the figure, the mass spectral position of an actual peptide from a Lys-C digest of bovine serum albumin, LCVLHEK -alkylated at the cysteine residue and tagged at the C-terminal lysine residue, is compared to the mass distribution of two hypothetical peptides containing naturally occurring mass defects from phosphorylation and from the presence of one sulfur atom. Despite their intrinsic mass defects, which naturally shifts their masses compared to peptides made exclusively of C,H,O and N atoms, it is evident that no overlapping between the untagged peptides and the iodinated charge-tagged peptides would occur This also makes the derivatized peptides more easily detectable and identifiable during peptide mapping. Additionally, each of the peptide's z-ions generated from ETD fragmentation contains the mass defect and are also shifted in spectral areas where no native peptide fragment ion exists, allowing for easier sequencing and thus, peptide identification. It should be noted that the z_1 ion of each tagged peptide, which corresponds to the C-terminal lysine residue derivatized with the diiodinated charge tag, is the same for all tagged peptide and has a *m/z* = 544.9667. The implication of having the smallest z-ion be identical for all peptides during sequencing is that even in cases where this ion does not get detected during the data collection process, it has no negative impact on the ability to fully sequence a given peptide. Indeed, the subsequent fragment ion in any sequence, the z_2 ion, will be a combination of the z_1 mass (544.9667 Da) plus the mass of any one of the twenty naturally occurring

amino acids. Consequently, one can readily build a spreadsheet or write a script that enables the prediction of the next possible amino acid and associated fragment ion mass in a sequence for simplified *de novo* sequencing.

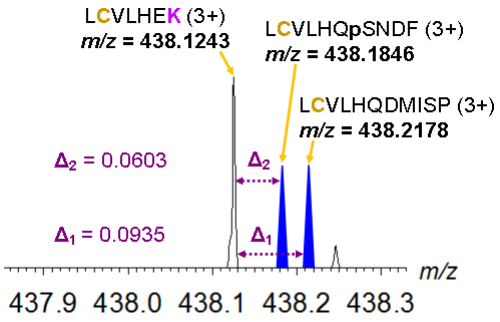


Figure 5.1. Spectral Separation of Isobars Due to Mass Defect Labeling.

Furthermore, due to the uniqueness of the z_1 ion as the only possible common ETD fragment ion to all tagged peptides in a digest and as having a large mass defect (*m/z* 544.9667) a script was developed in order to extract all scans that contain this ion in a given LC-MS/MS sample run (See **Scheme 5.1, SI**). The data extraction output is organized into three columns, namely, scan level (MS^1, MS^2), scan number, and intensity of the z_1 ion. This allows one to manually inspect the spectra individually and to confirm the presence of a tagged peptide containing the z_1 ion, as well as to make a rapid comparison to the compiled peptides identified by the database. Using a *de novo* approach in combination with the assistance of a spectral database permits the identification of tagged peptides present in the digest that may not be detected by database searching and matching. It also equips one with an effortless way to gather all the spectra of interest for *de novo* analysis. Since bovine serum albumin (BSA) is a known protein, its peptide mapping is readily available using database searching. As such, using Proteome Discoverer, all LC-MS/MS sample runs from the digests were analyzed for the presence of tagged peptides. In order to do so, the peptide tag modification (413.8726 Da) with a mass tolerance for incorporation of one or two hydrogen atoms (+/- 1.0078 and 2.0156 Da) from protonation was included in the search engine and run for BSA proteolysis by both trypsin and LysC enzymes. All peptides identified by database searching were visually inspected and only the ones with a high confidence interval matching (>95%) were recorded. The script was then applied to the same RAW file from

each LC-MS/MS sample run and the peptides that contained the z_1 ion were compiled in **Tables 5.2** and **5.4** for the trypsin and Lys-C digest respectively.

Moreover, there is a general interest in the differentiation between leucine and isoleucine during peptide sequencing. Several methods of distinction between the two isomeric amino acids have been developed.[271-276] In this work, we attempted to differentiate between isoleucine and leucine residues in a synthetic peptide, KQTALVELLK, where one Leu at a time is substituted for one Ile following a method reported by Lebedev et al.[277] The method uses an ETD-MS^3 approach of isolation and fragmentation of z-radical (z•) ions, in which the radical initiation triggers two different side-chain cleavages typical of either isomer. Briefly, the doubly or triply charged peptide precursor ion is dissociated by ETD in an MS^2 experiment and each z-ion radical produced that holds either a Leu or Ile residue at the radical site in the peptide is isolated in an MS^3 experiment and further fragmented by CID. Depending on which amino acid residue the radical is initiated, either a loss of 43 Da ($-C_3H_7$) or 29 Da ($-C_2H_5$) will be observed that corresponds either the Leu or Ile isomer, respectively.

5.2 Experimental Section

5.2.1 *Materials*

The protease rLysC was purchased from Promega and reconstituted with the buffer provided by the company. Trypsin was also purchased from Promega and reconstituted according to the company's protocol (in 50mM acetic acid). All other materials used for the coupling of the diiodinated charge tag to synthetic peptides are the same as listed in the materials **section 4.2.1**. Synthetic KQTALVELLK, KQTAIVELLK, KQTALVEILK and KQTALVELIK were synthetized on pre-loaded Wang resin (Sigma-Aldrich) using a microwave-assisted Liberty Blue peptide synthesizer.

5.2.2 *Trypsin Digest and peptide tagging*

Using ammonium bicarbonate in water (0.1 M) as the solvent, three separate stock solutions of 10 mg/mL BSA, tris(2-carboxyethyl) phosphine (TCEP, 0.5 M) and iodoacetamide (IAA, 1 M) were prepared. Acetic acid in water (50 mM) was used to dissolve trypsin (100 μg) according to the Trypsin Gold Protocol

provided from Promega. Four Eppendorf tubes were prepared for BSA tryptic peptides by the following method: In each tube, 91.3 µL of 10 mg/mL BSA was mixed with 821.3 µL 100 mM ammonium bicarbonate and then 9.1 µL of 0.5 M TCEP (final concentration: 5 mM) was added into the tube before incubation at 37 °C for 30 minutes in order to break disulfide bonds within and between proteins. Then, 9.1 µL of 1 M IAA (final concentration: 10 mM) was added to each tube and all the tubes were vortexed at RT for 10 minutes to enable covalent binding of IAA to cysteine in the protein. After that, 22.8 µL of 0.2 µg/µL trypsin (final trypsin: protein ratio of 1:200) was added to each tube and all four tubes were incubated at 37 °C overnight. Following this digestion, each tube was acidified with adding formic acid (~ 7 µL) to pH = 3 and tested with pH paper. Trypsin-digested peptides from two tubes were combined and desalted with a single Sep-Pak C18 cartridge. Two Sep-Pak C18 cartridges were used according to the protocol provided from Waters and described as follows. First, the cartridge was conditioned with 1 mL of Acetonitrile (ACN) at a 5-10 mL/min flow rate. Second, the cartridge was equilibrated with 2 mL of ACN/H_2O/TFA (2/97.9/0.1) mixture at a 5-10 mL/min flow rate. Third, the sample was loaded from two Eppendorf tubes into the cartridge, reloading each sample twice at 1 mL/min flow rate. Fourth, the cartridge was washed with 3 mL of an ACN/H_2O/TFA (2/97.9/0.1) mixture at a 5-10 mL/min flow rate. Finally, BSA digested peptides in the cartridge were first eluted with 1 mL of ACN/H_2O/acidic acid (30/69.5/0.5) mixture then eluted with 1 mL of ACN/H_2O/acidic acid (69.5/30/0.5) mixture at a 1 mL/min flow rate. The 4 mL of eluates from the two cartridges were combined and dried down with a stream of nitrogen, and were redissolved in 0.8 mL of a MeOH/H_2O (50/50) solvent mixture. Twenty microliters (20 µL) of the peptide solution was taken out and saved as a standard for future LC-MS/MS analysis of BSA tryptic peptides standard. Triethylamine (6.7 µL, 48 µmoles) was mixed with the remaining BSA tryptic peptide solution, which was then introduced into the syringe with modified beads (diiodinated charge-tag precursor synthesis described in detail in **Chapter 4**). The syringe was stirred at RT for 64 h. Following binding of the peptides to the beads, the unreacted peptides left in the solution of the fritted syringe were collected for later analysis. Next, the beads in the syringe were washed with MeOH/H_2O (50/50) five times (five minutes stirring each time) and then washed with dichloromethane another three times (five minutes rotation each time). The syringe was then dried under a mechanical vacuum pump for 2 h. A mixture (1.0 mL) of TFA/ H_2O/triethylsilane (95/2.5/2.5) was introduced into the syringe to cleave the modified peptides from the beads. After rotating for 1 h, the solution in the syringe was

collected in a glass vial. The beads in the syringe were washed with TFA (0.5 mL × 2) another two times (five minutes stirring each time) to enable the cleavage of the modified peptides from the beads into the solution. The washes were combined into the previous glass vial, dried under nitrogen and redissolved in 0.25 mL methanol. The modified peptides dissolved in methanol were dried and then redissolved in 100 µL (estimated to give roughly 100 µM for modified digested peptides) of H_2O/acetic acid (99/1) for further LC-MS/MS analysis with ETD fragmentation. The BSA tryptic peptide standard collected before was dried and then redissolved in 12.5 µL (estimated to give 100 µM for digested peptides) of H_2O/acetic acid (99/1).

5.2.3 *Lys-C Digest and Peptide Tagging*

Bovine Serum Albumin (BSA) proteins (10 mg/mL) were dissolved in an 8 M Urea, 1mM EDTA and 25 mM Tris-HCl lysis buffer (pH 8.5). The Tris-HCl solution was prepared by mixing Tris base and a few drops of concentrated HCl until the desired pH 8.5 was obtained. Four separate Eppendorf tubes were prepared; into each, 75 µL of the 10 mg/mL BSA solution was dispensed and 125.6 µL of the lysis buffer was added. A reducing agent, TCEP (500 mM), was then added to the protein-buffer mixture in each tube so as to break the disulfide bonds of the proteins (2 µL, final concentration 5 mM) and the test tubes were incubated at 37°C for 30 minutes. After incubation, a freshly made solution of Iodoacetamide (IAA, 500 mM) was added to each test tube in order to prevent the reformation of the disulfide bonds and let react in the dark for 15 min at RT (2 µL, final concentration 10 mM). Afterwards, 1.1 mL of the buffer (25 mM Tris-HCl and 1 mM EDTA only, pH 8.5) was used to dilute the concentration of urea in each test tube to less than 1 M. The rLysC protease (4 X 15 µg) was subsequently resuspended in its vial using the reconstitution buffer (4 X 75 µL) provided by Promega to make four 0.2 µg/µL r-LysC solutions (final ratio of rLysC : protein by weight is 1:50). To each tube was added the full 75 µL of protease contained in each vial and the tubes were then incubated at 37°C overnight (18 hours). After digestion, each tube was acidified to pH = 3 with formic acid (~5-7 µL) before desalting via C18 Sep-Pak. The next steps for the coupling were the same as described previously in **section 5.2.1**.

5.3 Results and Discussion

5.3.1 *Trypsin Digest*

Not accounting for miscleavages, a typical BSA trypsin digest produces 74 different peptides (greater or equal to 2 amino acids), out of which 52 are lysine-terminated while the rest are arginine-terminated. Previous work[247] has shown the ease of separation between Arg-terminated and Lys-terminated peptides through washing steps. The Proteome Discoverer search results showed that the sequence coverage for the tagged samples were 79.57% and 84.68% for the two repeats, implying incomplete digestion of the protein compared to 80.2 and 88.9% for the untagged peptides. The compiled data suggests that a total of fifteen peptides were identified as tagged peptides using the database (**Figures 5.2-5.5, SI Figures 5.11-5.33**) and that enhancement of the sequence coverage was observed for a total of six tagged peptides in the digest, namely, KQTALVELLK, SHCIAEVEK, LSQKFPK, LCVLHEK, AEFVEVTK and HLVDEPQNLIK (**Figures 5.2-5.5**, **SI Figures 5.10-5.18**). Displayed in **Figure 5.2** is the ETD-MS^2 spectrum for the standard (untagged) KQTALVELLK detected and identified from the standard sample. It is worth noting that the z-ion sequence coverage is poor and that the peptide is doubly charged (+2). While no triply charged untagged species (+3) was detected in the standard, it was the only species detected in the tagged peptide (+3), **Figure 5.3**. The tagged peptide exhibits a complete coverage of both the c- and z-ions, including the z_1 ion. It must be mentioned that the miscleaved KQTALVELLK peptide tagged at the N-terminal Lys was also detected, **Figure 5.10, SI**. While the diiodinated charge tag did increase the charge state of this N-terminal tagged peptide, no benefit in the sequence coverage was observed and both the z_1 and z_7 ions are missing from the spectrum. Similar observations about the increased charge state and sequence coverage of the C-terminal tagged KQTALVELLK can be made for the cysteine-alkylated LCVLHEK sequence detected by database comparison. Shown in **Figure 5.4** is the ETD-MS^2 spectrum for the underivatized doubly charged LCVELHEK identified by database search. Out of the six total z-ions for this peptide, four were observed and only two out of the six total c-ions were present. This stands in sharp contrast with the spectrum obtained for the derivatized analog, for which a comprehensive sequence coverage was observed. Indeed, all six z-ions were present in the spectrum and a total of five c-ions were detected for the peptide.

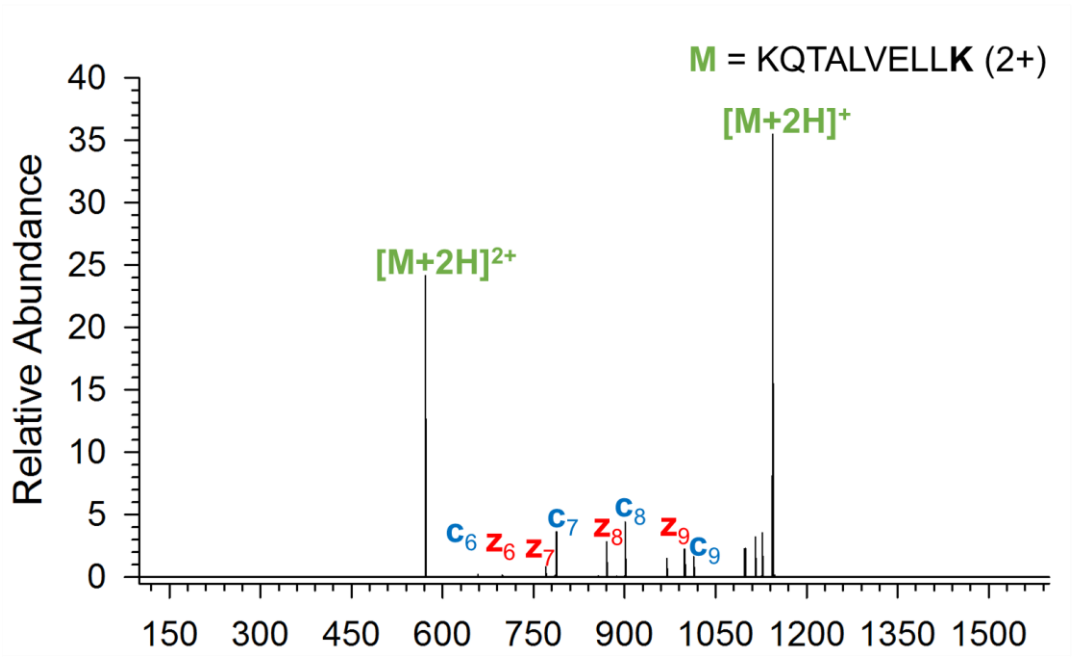


Figure 5.2. ETD-MS^2 of Underivatized Trypsin-Digested KQTALVELLK Detected in Standard.

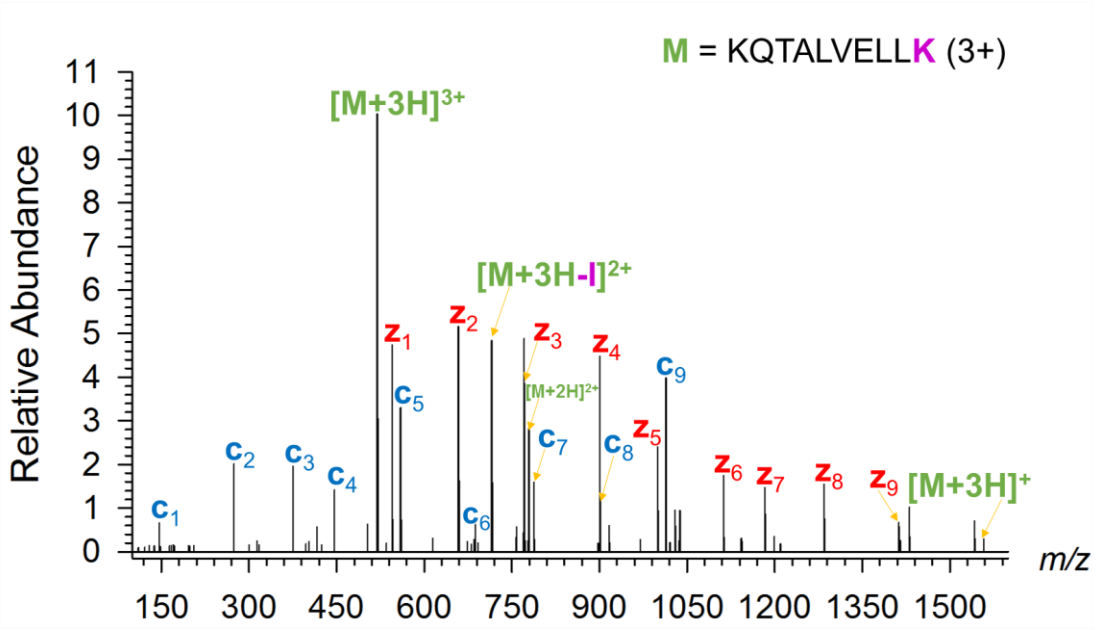


Figure 5.3. ETD-MS2 of Derivatized Trypsin-Digested KQTALVELLK Detected in Tagged Sample.

Worth noting is the increase in the charge state of that peptide from untagged to tagged, with the doubly charged species (+2) being the only form of the peptide obtained for the untagged LCVLHEK compared to the triply charged state (+3) of the tagged analog as the only species present in the digest. This emphasizes the power of the guanidination of a peptide in increasing the charge state of that peptide and thus, in improving the electron transfer dissociation efficiency. In addition, the diiodinated charge tag aided in the detection of smaller peptides (i.e. CASIQK, LVTDLTK, NYQEAK, and GACLLPK, **SI Figures 5.30-5.33**) as summarized in **Table 5.2**. These five-to-seven amino acids long peptides escaped detection in the untagged standard but were identified in the tagged digest mixture with the database searching.

Lastly, it must be mentioned that out of the total fifteen identified tagged peptides by database search, a few showed either no enhancement of the sequence coverage (i.e. FKDLGEEHFK, **SI Figure 5.28**) or even a decrease in the sequence coverage compared to their untagged counterparts (i.e. ATEEQLK, LKPDPNTLCDEFK and RPCFSALTPDETYVPK, **SI Figures 5.19-5.24**). Though two of these peptides appear to be on the larger side compared to the rest of the identified peptides (13 and 17 amino acids long), no specific pattern, such as peptide size, can be associated with this decrease in sequence coverage since the decrease was also seen for a shorter, 7 amino acids long peptide.

The need for a *de novo* approach in the determination of the sequence of a tagged peptide digest is greatly underlined by the results showed in **Table 5.3**. The table is a compilation of all the peptides containing the unique z_1 ion at *m/z* 544.9667 that were extracted by the abovementioned script (See **Scheme 5.1, SI**). One can see that a total of twenty-three peptides failed detection by the Proteome Discoverer database (**SI Figures 34-56)**. This suggests that this is the minimum number of peptide sequences that were not detected by the database and that potentially more peptides -those not containing the z_1 ion for which the script was developed, failed detection by the database. For those peptides, a *de novo* approach of identification will need to be conducted, as illustrated for two of the undetected peptides in the Lys-C digest mentioned in the next section (**SI Figures 5.82,5.83**). In order to do so, the spectra are inspected for z-ion fragments and the difference of mass between two z-ions, starting at Δz_1-z_2 is assigned to an amino acid until the entire sequence is identified. The identified sequence can be confirmed with the aid of a peptide fragmentation calculator and the exact masses checked.

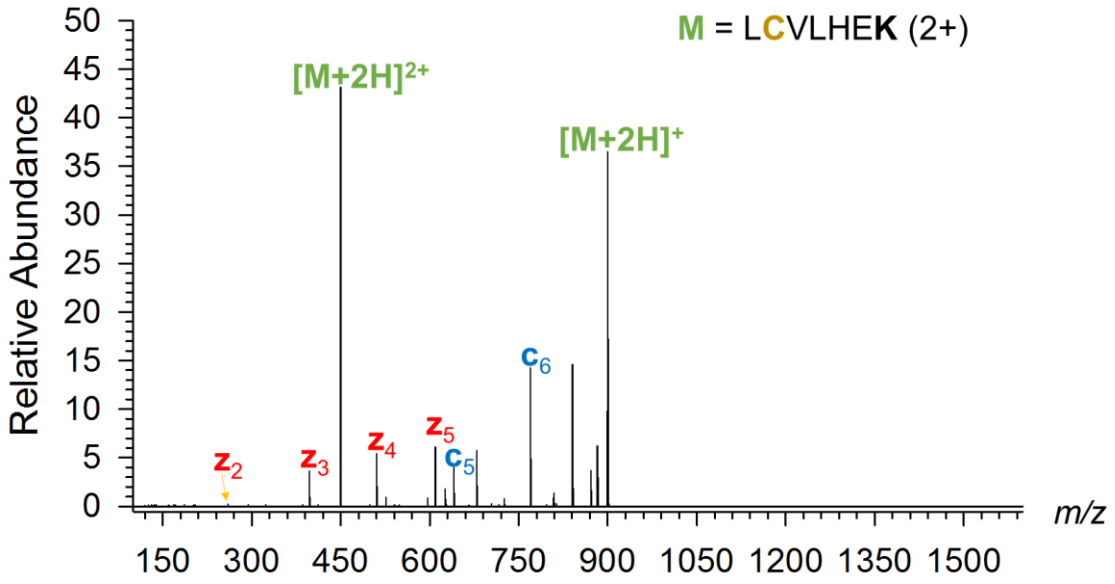


Figure 5.4. ETD-MS2 of Underivatized Trypsin-Digested LCVLHEK Detected in Standard.

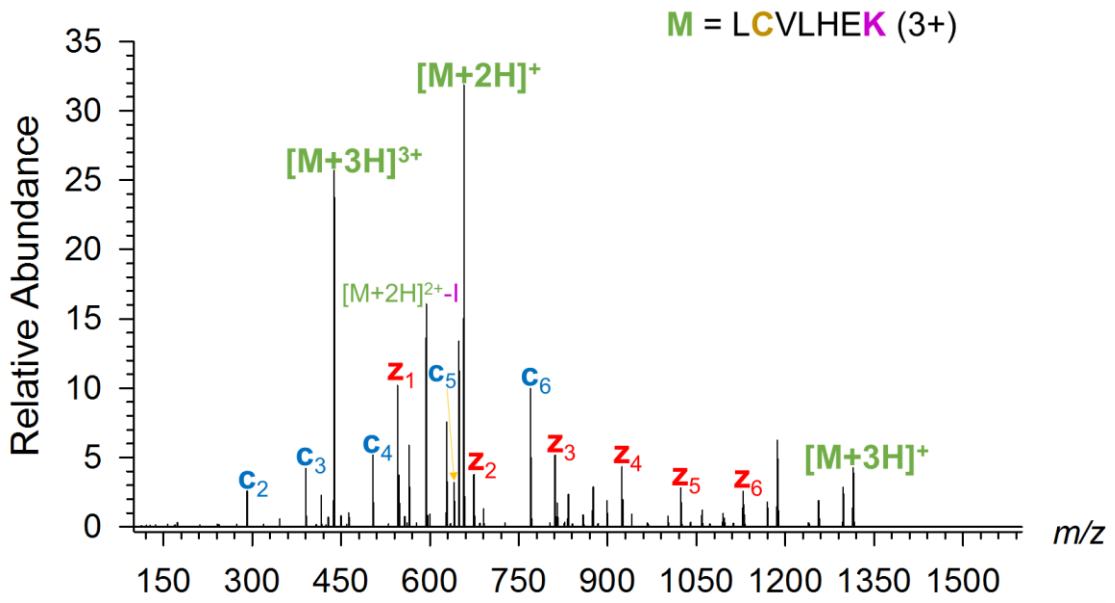


Figure 5.5. ETD-MS2 of Derivatized Trypsin-Digested LCVLHEK Detected in Tagged Sample.

	Charge State				z-ions Sequence Coverage		Comment
Peptide	(+2) Standard	(+2) Tagged	(+3) Standarc	(+3) Tagged	Untagged	Tagged	
KQTALVELLK	-	-	-	519.1393	4/9 = 44%	9/9 = 100%	Both tagged & not tagged at C-terminal
SHCIAEVEK	536.7588	-	-	496.1321	4/8 = 50%	7/8 = 87.5%	Not found in standard for one repeat
LSQKFPK	424.2563	-	-	421.1311	2/6 = 33%	5/6 = 83%	
LCVLHEK	449.7451	-	-	438.1229	3/6 = 50%	5/6 = 83%	
AEFVEVTK	461.7486	668.685	-	-	4/7 = 57%	5/7 = 71%	
RPCFSALTPDETYV	-	-	627.6465	765.6038	10/15 = 67%	4/15 = 27%	
KVPQVSTPTLVEVS	-	-	547.3185	685.2755	8/14 = 57%	12/14 = 86%	
LKPDPNTLCDEFK	-	-	526.2621	664.2194	9/12 =75%	7/12 = 58%	Not tagged at C-terminal
HLVDEPQNLIK	-	-	435.9114	573.868	7/10 = 70%	8/10 = 80%	
FKDLGEEHFK	-	-	417.5475	555.1697	8/9 = 89%	8/9 = 89%	Both tagged & not tagged at C-terminal
CASIQK	-	561.1097	-	-	-	4/5 = 80%	Not found in standard
ATEEQLK	409.7166	616.6541	-	-	4/6 = 67%	3/6 = 50%	
NYQEAK	-	583.6199	-	-	-	3/5 = 60%	Not found in standard
GACLLPK	-	586.6523	-	-	-	3/6 = 50%	Not found in standard
LVTDLTK	-	602.1769	-	-	-	6/6 = 100%	Not found in standard

Table 5.1. Summary of z-ion Count Comparison for Untagged and Tagged Peptides from Trypsin Digest.

Peptides containing z1 ion at m/z 544.9667			
	Charge State		
	(+1)	(+2)	(+3)
1	831.1609	416.0819	-
2	842.4963	421.0683	-
3	902.1572	451.5811	-
4	888.1402	445.1202	-
5	913.1486	456.5735	-
6	921.166	461.0865	-
7	935.1855	468.186	-
8	1383.851	692.4292	462.147
9	1075.236	537.6183	-
10	931.1718	466.0895	-
11	1024.1661	512.0836	-
12	806.113	403.057	-
13	1193.7198	598.3606	-
14	1373.3148	926.59	463.1476
15	1324.825	883.514	441.945
16	1471.9021	979.5654	490.9711
17	1104.2083	552.1045	-
18	1721.6041	860.8017	574.2026
19	1116.2805	558.1416	-
20	1428.8589	952.1605	476.0818
21	1456.8529	-	485.2819
22	833.4723	417.2394	-
23	886.5609	443.2818	-

Table 5.2. Summary of peptides from trypsin digest that contain the tagged z1 (m/z 544.9667) as detected by the in-house extraction script.

5.3.2 *Lys-C Digest*

Not accounting for miscleavages, a typical BSA Lys-C digest produces 57 different Lys-terminated peptides (=/> 2 amino acids). Out of this total, Proteome Discoverer was able to identify sixteen tagged peptides, **Table 5.3**; of those, four showed an improved z-ion sequence coverage relative to the untagged peptides (ie. FGERALK, IETMREK, LGEYGFQNALIVRYTRK, and AWSVARLSQK, **Figures 5.6-5.9**, **SI Figures 5.57-5.60**), four showed the same level of coverage (i.e. VPQVSTPTLVEVSRSLGK, DLGEEHFK, AEFVEVTK and ATEEQLK, **SI Figures 5.61-5.68**), two showed a decrease in the sequence coverage (i.e. DDSPDLPK and SLHTLFGDELCK, **SI Figures 5.69-5.72**), and six were identified when tagged but not detected when untagged (SEIAHRFK, TPVSEK, AFDEK, SHCIAEVEK, LVTDLTK and FWGK, **SI Figures 5.73-5.78**), which affirms one of the values of the diiodinated charge tag. Two examples of the successfully improved peptide sequence coverage for FGERALK and IETMREK are illustrated in **Figures 5.7** and **5.9** relative to their respective untagged analogs, **Figures 5.6** and **5.8.** The untagged FGERALK was identified and found to be present in the standard (untagged sample) only as a doubly charged species (+2), while it was found to be present in the tagged sample in a triply charged state (+3) only. The untagged FGERALK peptide displayed poor sequence coverage for both c- and z-ions while the tagged counterpart displayed a complete coverage of the z-ion series, including the z_1, ion and an improved c-ion coverage. The same observations are true for the other listed example, IETMREK for which the incorporation of the diiodinated charged tag dramatically improved the z-ions sequence coverage, from 67% when untagged compared to 100% when tagged. Furthermore, as described for the trypsin digest data, the database search missed the detection of several tagged peptides that the in-house script was able to extract, as illustrated in Table 5.4. A total of 52 peptides that contained the tagged z_1 ion at *m/z* 544.9667 were extracted by the script, **Figures 5.79-5.130**. These unidentified peptides require a *de novo* approach to sequencing for identification. Two examples of the de novo identification of these peptides are illustrated (**SI Figures 5.180,5.181**) for the Lys-C digested unidentified peptides 4 and 5 (**SI Figures 5.82, 5.83**). Unidentified peptide 5 was confidently assigned to MTERALK and unidentified peptide 4 was identified as either FAERALK or AFERALK since its fragmentation pattern is missing the z_6 ion, thereby making the distinction between the two impossible. Another point worth noting is the complementarity of the identified peptide sequences in the trypsin and Lys-C digest. Only a few of the digest

peptides were found to be present in both the Lys-C and trypsin digests, namely, SHCIAEVEK and AEFVEVTK, while the other identified Lys-C peptide sequences were absent in the trypsin digest.

	Charge State						z-ions Sequence Coverage	
Peptide	(+2) Standard	(+2) Tagged	(+3) Standard	(+3) Tagged	(+4) Standard	(+4) Tagged	Untagged	Tagged
VPQVSTPTLVEVSRSLGK	-	1157.4753	-	771.6503	475.2723	578.9902	13/17 = 77%	13/17 = 77%
LGEYGFQNALIVRYTRK	-	1223.0007	672.0363	815.3337	504.0269	611.5004	11/16 = 69%	15/16 = 94%
DLGEEHFK	487.7332	695.1716	-	463.7827	-	-	7/7 = 100%	7/7 = 100%
SEIAHRFK	494	701.7117	-	467.8089	-	-	Not found in standard	7/7 = 100%
AEFVEVTK	461.748	668.6856	-		-	-	6/7 = 86%	6/7 = 86%
ATEEQLK	409.7166	616.6539	-	616.6539	-	-	4/6 = 67%	4/6 = 67%
TPVSEK	-	537.6193	-		-	-	Not found in standard	4/6 = 67%
FGERALK	410.7372	617.6741	-	412.1184	-	-	4/6 = 67%	6/6 = 100%
IETMREK	453.7401	660.6755		440.7868	-	-	4/6 = 67%	6/6 = 100%
AFDEK	-	512.0846	-		-	-	Not found in standard	4/4 = 100%
SHCIAEVEK	-	744.1976	-	496.4677	-	-	Not found in standard	7/8 = 87.5%
LVTDLTK	-	602.1771	-	-	-	-	Not found in standard	6/6 = 100%
AWSVARLSQK	573.325	781.7723	-	520.9417	-	-	4/9 = 44%	8/9 = 89%
DDSPDLPK	443.713	650.6479	-	-	-	-	4/7 = 57%	3/7 = 43%
SLHTLFGDELCK	710.8533	918.2884	474.2369	611.8625	-	-	10/11 = 91%	6/11 = 54.5 %
FWGK	-	512.0846	-	-	-	-	Not found in standard	2/3 = 67%

Table 5.3. Summary of z-ion Count Comparison for Untagged and Tagged Peptides from LysC Digest.

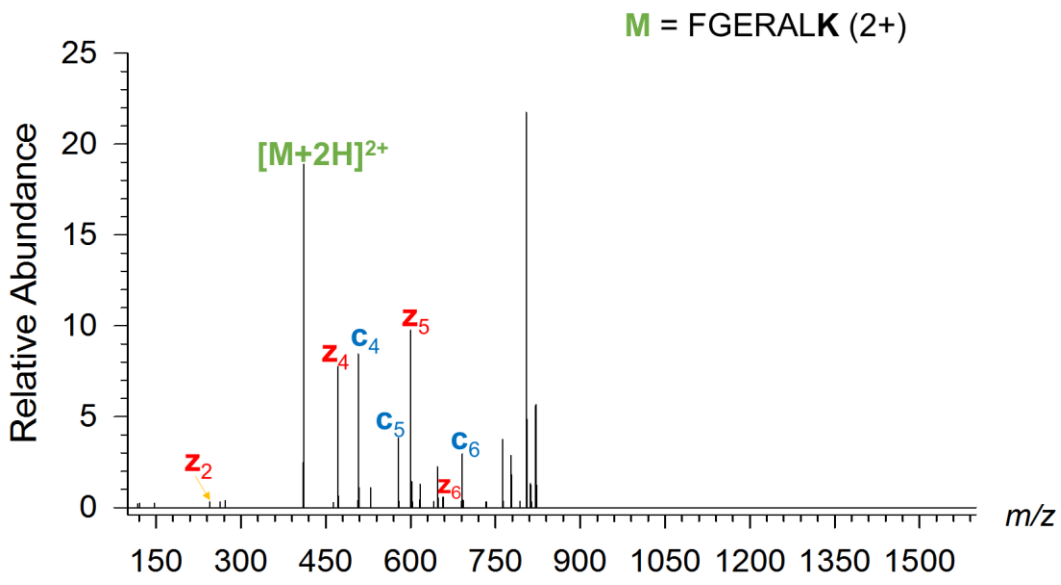


Figure 5.6. ETD-MS2 of Underivatized LysC-Digested FGERALK Detected in Standard.

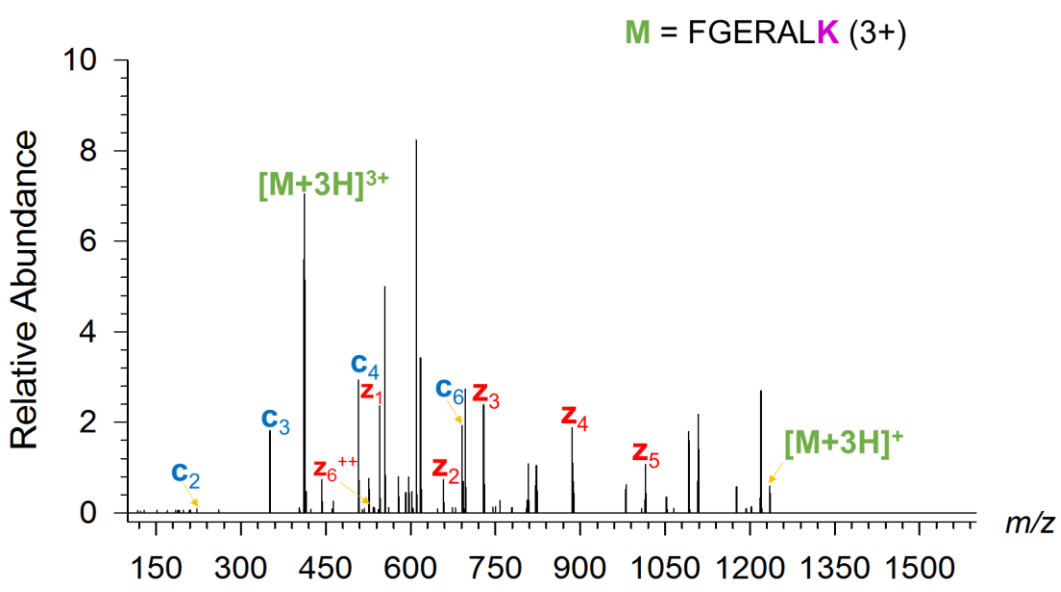


Figure 5.7. ETD-MS2 of Derivatized LysC-Digested FGERALK Detected in Tagged Sample.

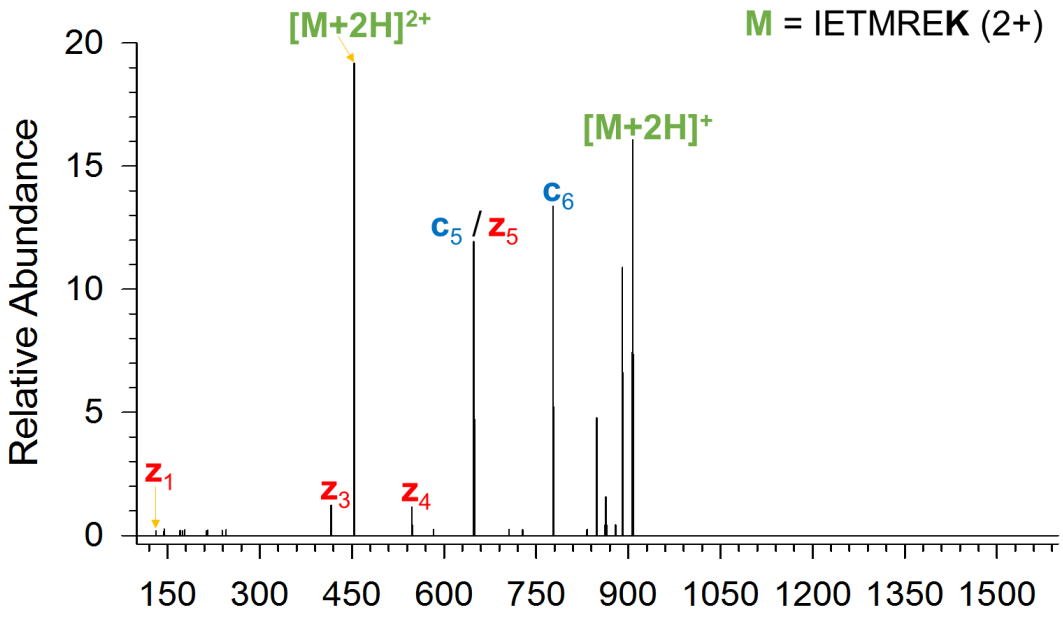


Figure 5.8. ETD-MS2 of Underivatized LysC-Digested IETMREK Detected in Standard.

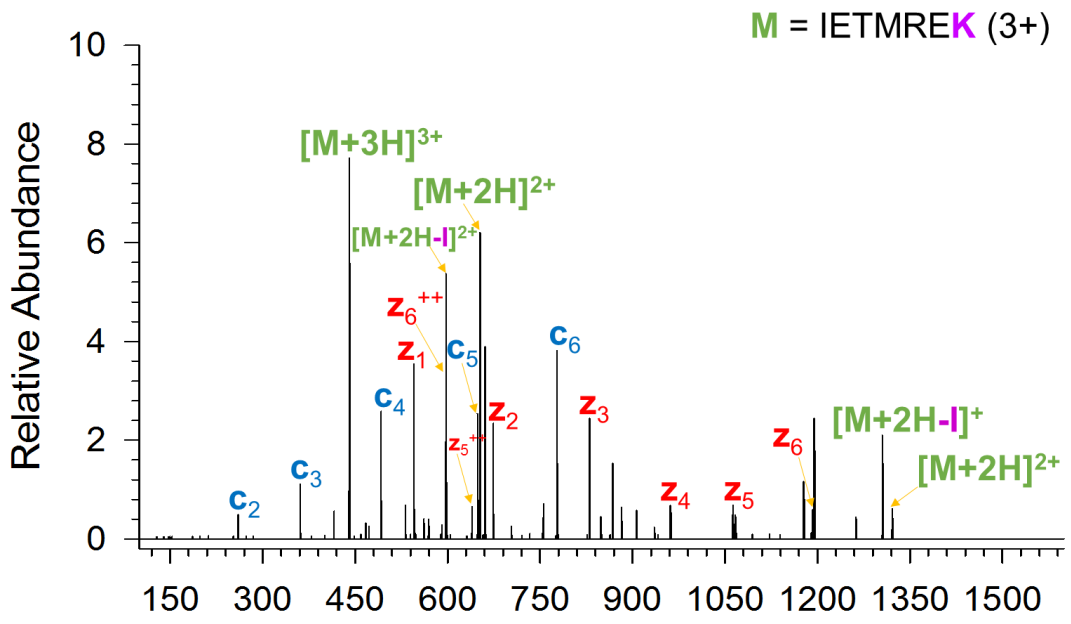


Figure 5.9. ETD-MS2 of Derivatized LysC-Digested IETMREK Detected in Tagged Sample.

5.3.4 *Leucine / Isoleucine Differentiation Results*

An attempt at evaluating the effect of the diiodinated charge tag on the differentiation between the leucine (Leu) and the isoleucine (Ile) isomers in a given peptide sequence was undertaken. The method followed the one utilized by Lebedev et al[277] and required the ETD generation of the z-ion radical ions, MS^2, located at either a Leu or Ile residue, and their subsequent isolation and fragmentation by CID-MS^3, **SI Figure 5.131**. The MS^3 spectra were analyzed for resulting fragments, with a distinctive loss of a 43 Da propyl radical ($-C_3H_7$) being indicative of the presence of a Leu residue at the radical site initiation, and the loss of a 29 Da ethyl radical ($-C_2H_5$) being indicative of the presence of an Ile residue. For this purpose, four peptide sequences were made synthetically using a microwave-assisted peptide synthesizer, namely KQTALVELLK, KQTAIVELLK, KQTALVEILK and KQTALVELIK, before being tagged using the described procedure in section 4.3.1. These specific sequences were chosen because KQTALVELLK was identified as one of the tagged peptide present in the trypsin digest of this study. In order to generate the z-ion radicals, both the triply (**SI Figures 5.132-5.134**) and doubly (**SI Figures 5.135-5.138**) charged peptide were individually isolated and

fragmented. Note that the triply charged KQTALVELIK was could not be isolated due to the low signal to noise ratio. For each peptide, the electron transfer dissociation activation time was varied between 50 ms and 200 ms so as to select the best ETD fragmentation efficiency and produce •L/IK (*m/z* 545), •L/ILK (*m/z* 787) and •L/IVELLK (*m/z* 1128) before further fragmentation of the radicals by CID. Unfortunately, as evidenced by both the triply and doubly charged spectra, **Figures 5.132-5.138**, the poor electron transfer dissociation efficiency yielded an insufficient z-ions signal across all four peptides to allow CID-MS^3. This low z-ions conversion did not allow their selection and further evaluation for the potential of leucine/isoleucine differentiation of the tagged peptides. This sort of discrimination is, however, important in sequencing and should be reevaluated in future work.

5.4 Conclusion

The work described herein presents the usefulness of a diiodinated charge tag for enhanced sequence coverage of tryptic peptides in the context of *de novo* sequencing. Indeed, successful tagging of the digested protein combined with a script that extracts the z_1 reporter ion allowed for detection of tagged peptides that were both identified and undetected by the database search. Undoubtedly, the unreliability of spectral database comparison is emphasized by the number of unidentified peptides from the database search. Conversely, these peptides that escaped detection via database search can be extracted from a script aimed at looking for the reporter ion. In addition, this script can easily be modified so as to screen LC-MS/MS runs for any given reporter ion. Since the sequence coverage enhancement was not observed for all peptides, it may be crucial to always perform the peptide mapping of a given peptide or protein on both tagged and untagged peptides for comparison. Nonetheless, the power of this tool for *de novo* sequencing and for enhanced sequence coverage by ETD implies a potential for a variety of applications. One such application is the tagging of histone fragments for enhanced ETD sequence coverage and location of their commonly seen post-translational modifications at lysine and arginine residues.[278] The methylation and acetylation of histones play a crucial role on the regulation of gene expression and aberrant modifications can lead to disruption in the gene expression and lead cause the onset of disease.[279] The quantification of these histone modifications is often achieved by bottom-up mass spectrometry and relies on the prior derivatization of the

histone fragments.[281-283] The ETD approach to studying histone fragments is of high value since the method does not alter PTMs and the mass shift may allow for ease of detection at both MS^1 and MS^2 levels. The diiodinated charge-tagging of such fragments may dramatically help in both the quantification and the correct assignment of modification sites, which pose challenges.[280]

Peptides containing z1 ion at m/z 544.9667			
	Charge State		
	(+1)	(+2)	(+3)
1	798.4686	400.1975	-
2	805.1074	403.0571	-
3	818.1308	409.5643	-
4	833.0744	416.79	-
5	1263.3478	631.671	421.45
6	850.0226	424.571	-
7	1277.3466	639.1816	426.1229
8	858.1194	429.0612	-
9	1287.3867	643.6937	435.1393
10	1316.3917	658.196	438.7974
11	1326.7844	663.8958	442.9335
12	899.5421	449.6147	-
13	907.4865	453.7439	-
14	1374.3967	687.1988	458.4668
15	1417.4397	708.7195	472.4802
16	949.1815	475.0941	-
17	953.1709	477.0903	-
18	957.5466	480.0076	-
19	961.5976	481.3035	-
20	968.1635	484.0794	-
21	1464.5493	732.2747	488.5176
22	981.1273	491.0871	-
23	1488.4045	744.1973	496.468
24	996.1915	498.593	-
25	1008.1884	504.0943	-
26	1015.6299	508.2113	-
27	1019.2581	510.1338	-
28	1535.5889	767.7936	511.8635
29	1023.1636	512.085	-
30	1549.6011	774.8014	516.5342
31	1561.5913	780.8013	520.3122
32	-	786.1619	524.11
33	1579.5958	789.2964	526.1984
34	1606.5525	803.7788	534.8088
35	1084.4618	542.0895	-
36	1649.6281	825.3165	549.8781
37	1659.5354	829.7757	553.5165
38	1111.236	555.6192	-
39	-	836.7767	557.8528
40	1722.4152	1148.579	574.1384
41	-	1157.1775	578.7394
42	1736.0581	868.5336	579.0248
43	-	874.5409	582.9797
44	1752.4896	876.2405	583.9528
45	-	1173.0034	586.7332
46	1784.7565	891.8749	595.2542
47	1802.6698	1199.6954	600.8959
48	1808.5223	904.2662	603.1795
49	-	1251.5012	626.0024
50	-	947.3724	632.253
51	1925.7411	962.8708	641.9616
52	1679.9977	839.9993	-

Table 5.4. Summary of peptides from trypsin digest that contain the tagged z1 (m/z 544.9667) as detected by the in-house extraction script.

5.5 Supplemental Information

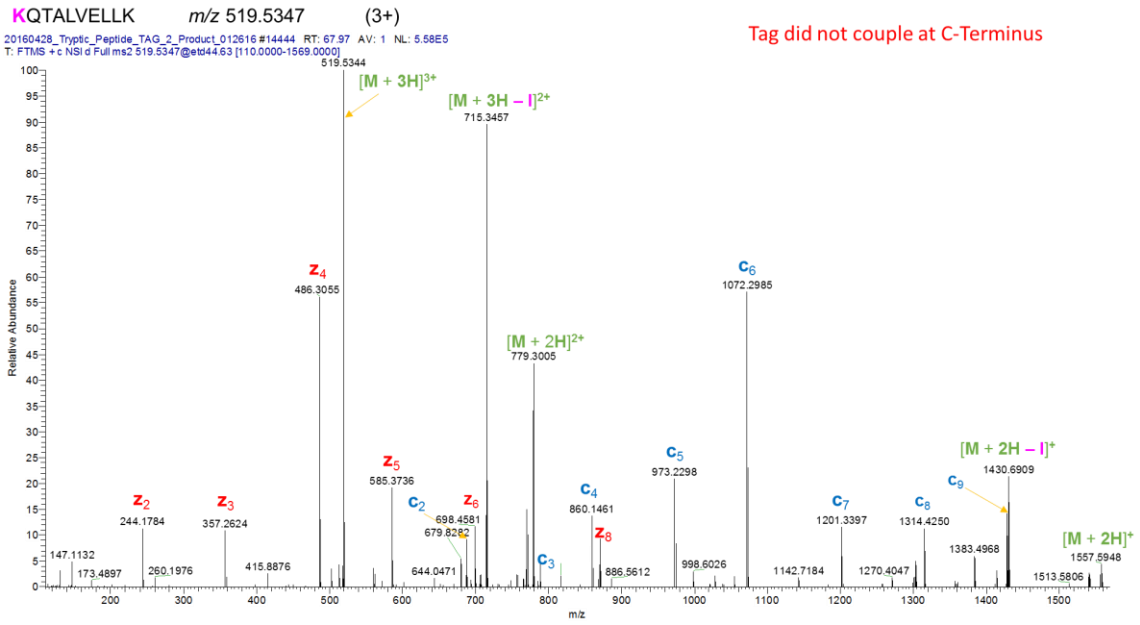


Figure 5.10. ETD-MS2 of Derivatized Trypsin-Digested KQTALVELLK Detected in Tagged Sample by Database Search. (Note that this peptide tagged at the N-terminal lysine rather than the targeted C-terminal lysine)

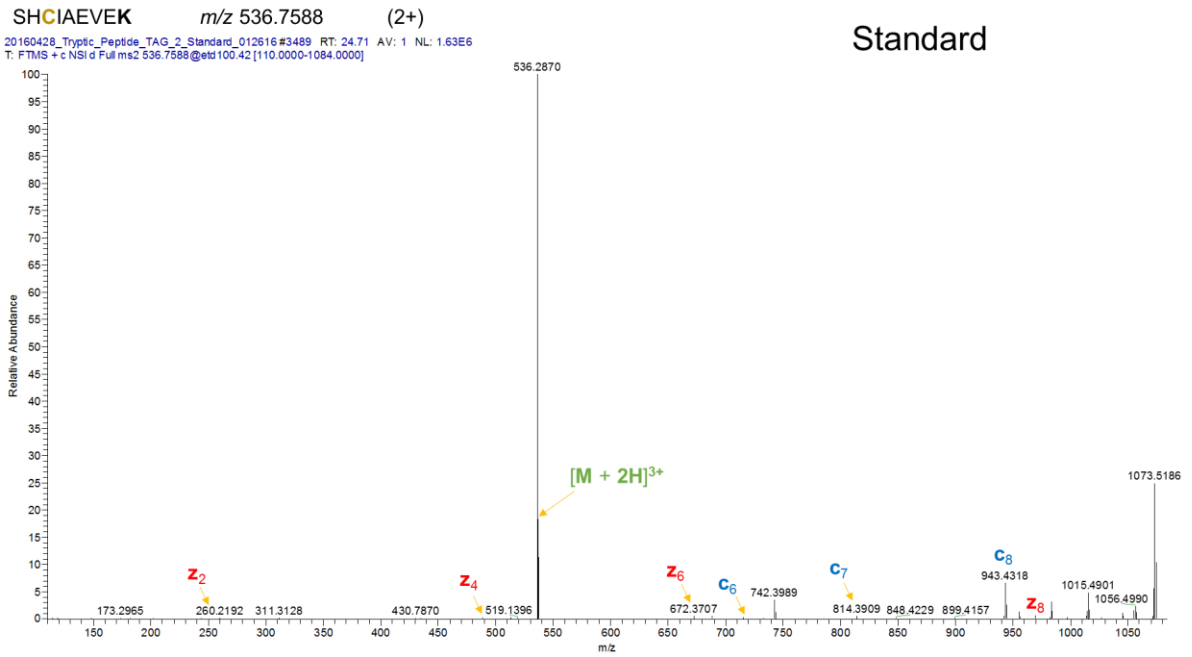


Figure 5.11. ETD-MS2 of Underivatized Trypsin-Digested SHCIAEVEK Detected in Standard by Database Search. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow, to prevent the reformation of disulfide bonds)

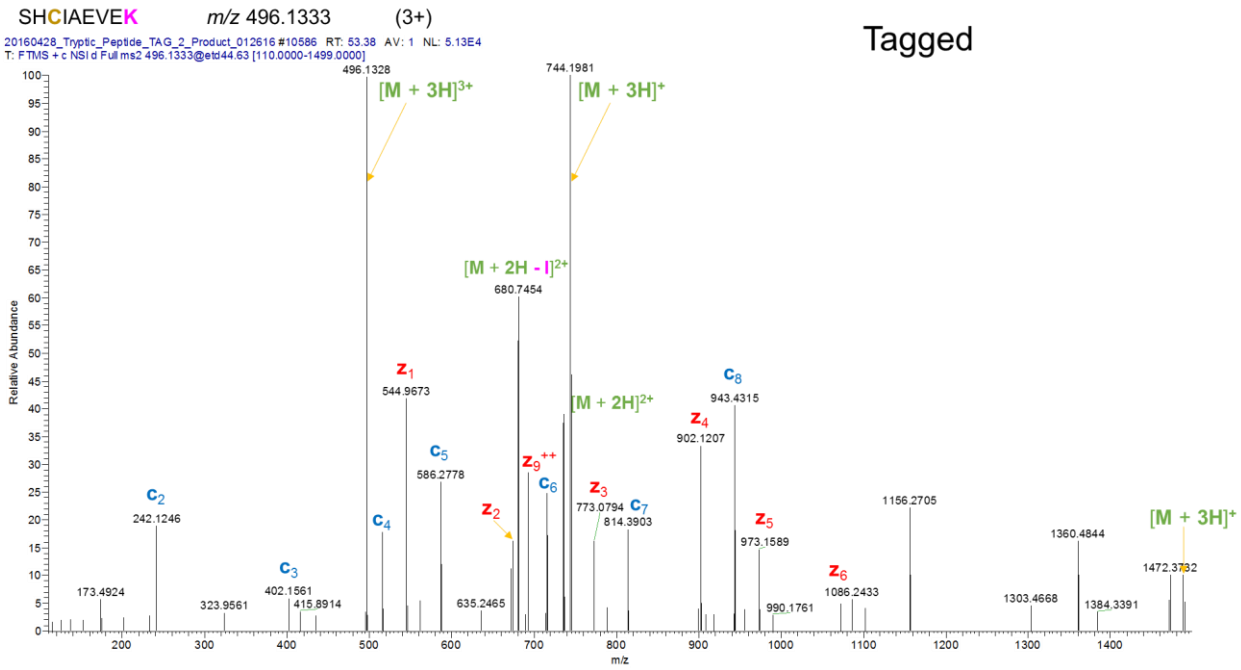


Figure 5.12. ETD-MS2 of Derivatized Trypsin-Digested SHCIAEVEK Detected in Tagged Sample by Database Search. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow, to prevent the reformation of disulfide bonds)

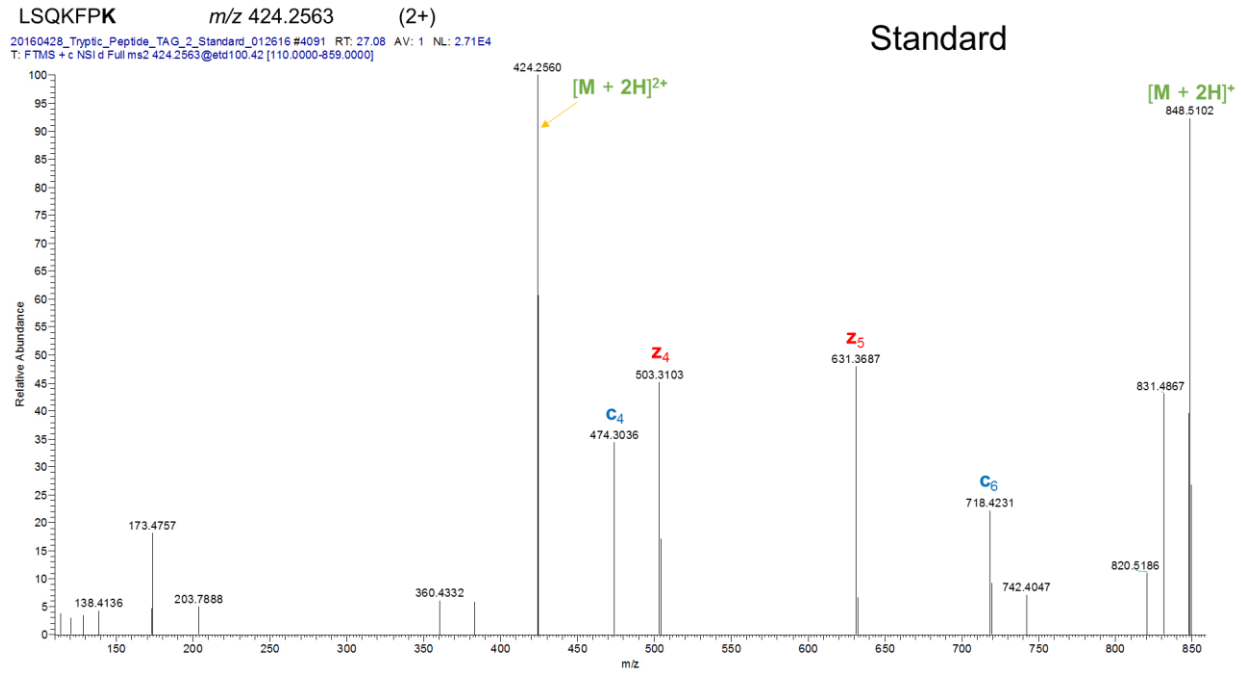


Figure 5.13. ETD-MS2 of Underivatized Trypsin-Digested LSQKFPK Detected in Standard by Database Search.

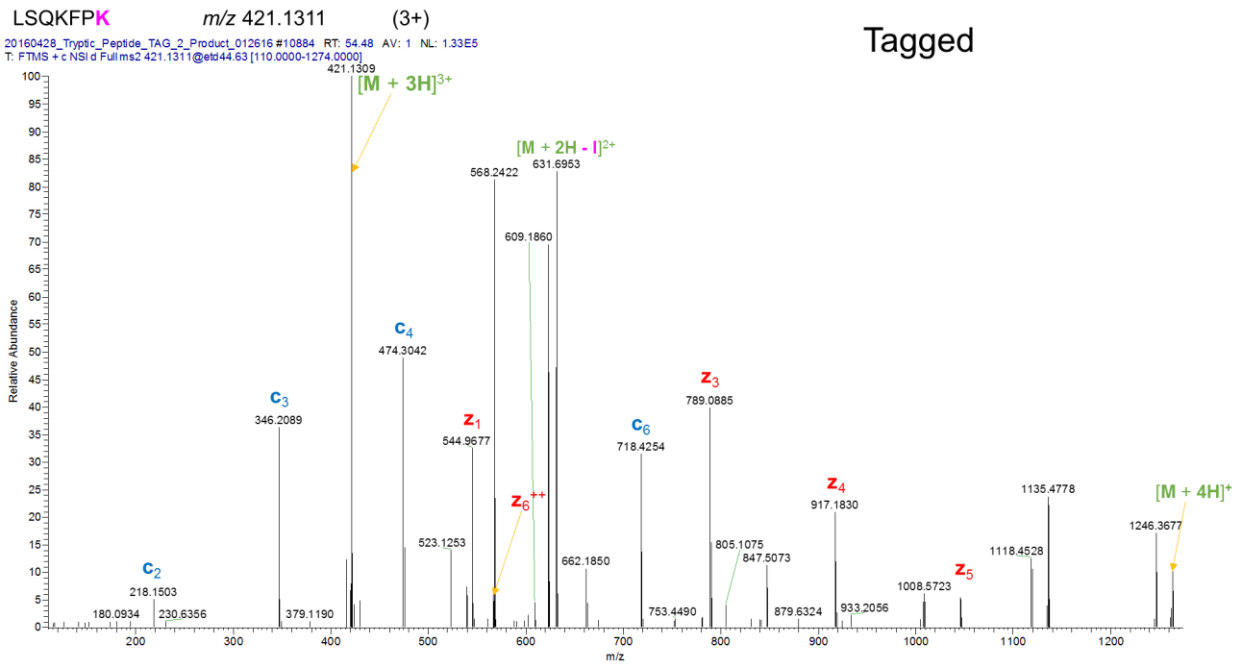


Figure 5.14. ETD-MS2 of Derivatized Trypsin-Digested LSQKFPK Detected in Tagged Sample by Database Search.

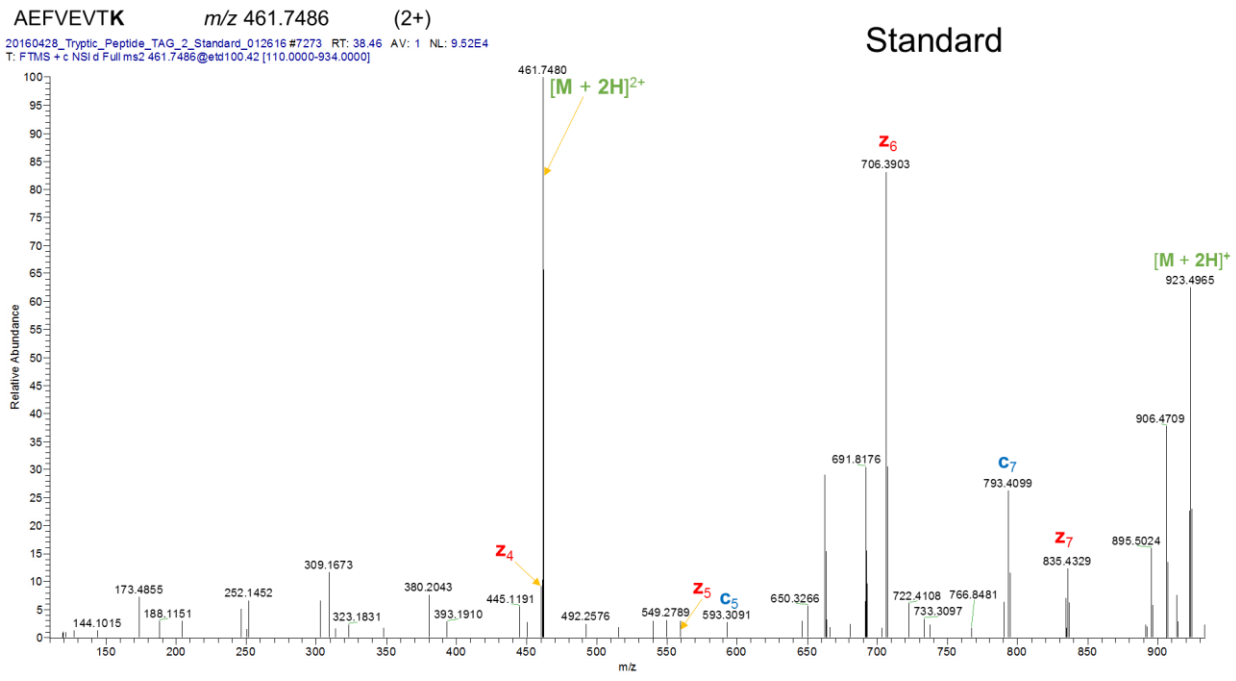


Figure 5.15. ETD-MS2 of Underivatized Trypsin-Digested AEFVEVTK Detected in Standard by Database Search.

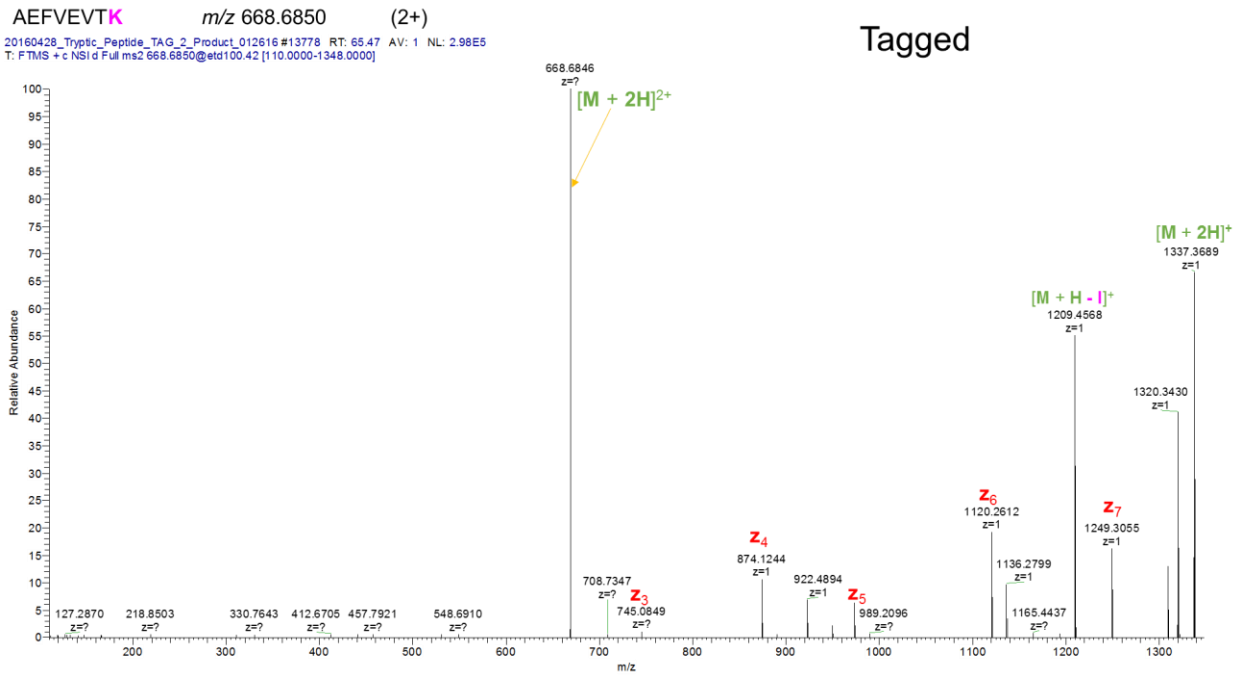


Figure 5.16. ETD-MS2 of Derivatized Trypsin-Digested AEFVEVTK Detected in Tagged Sample by Database Search.

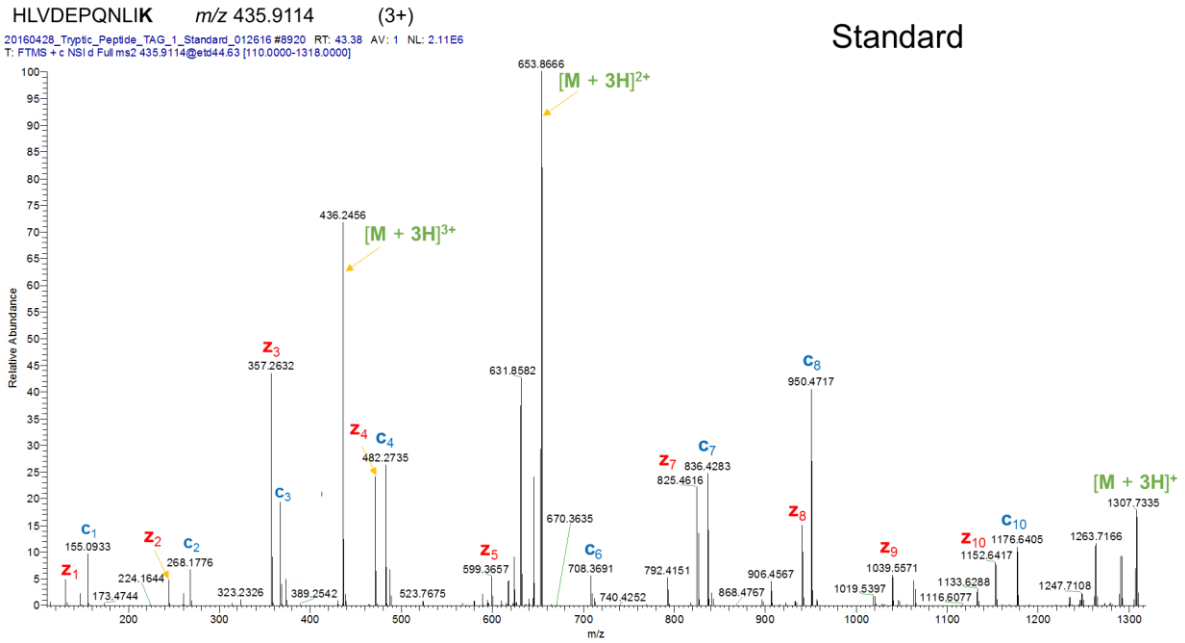


Figure 5.17. ETD-MS2 of Underivatized Trypsin-Digested HLVDEPQNLIK Detected in Standard by Database Search.

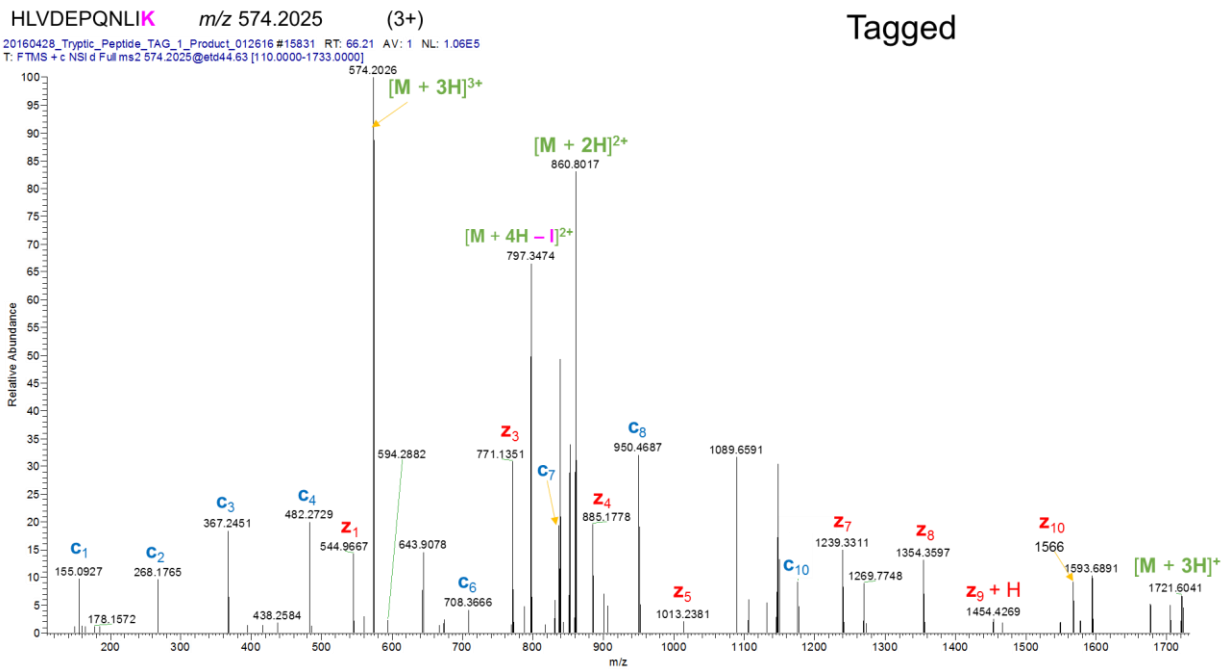


Figure 5.18. ETD-MS2 of Derivatized Trypsin-Digested HLVDEPQNLIK Detected in Tagged Sample by Database Search.

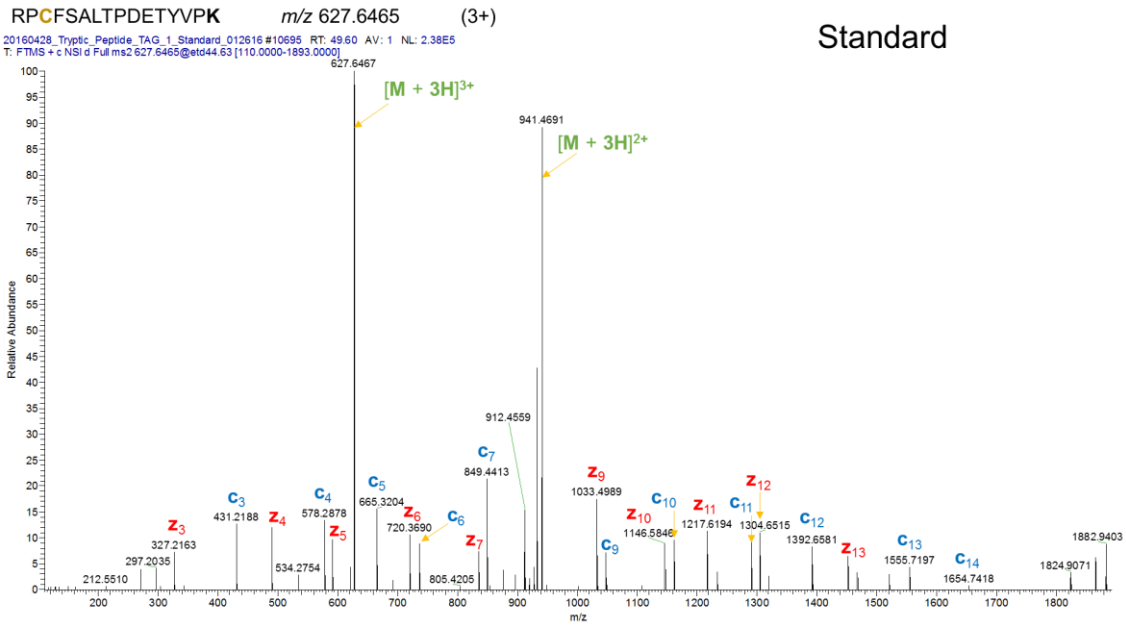


Figure 5.19. ETD-MS2 of Underivatized Trypsin-Digested RPCFSALTPDETYVPK Detected in Standard by Database Search. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow, to prevent the reformation of disulfide bonds)

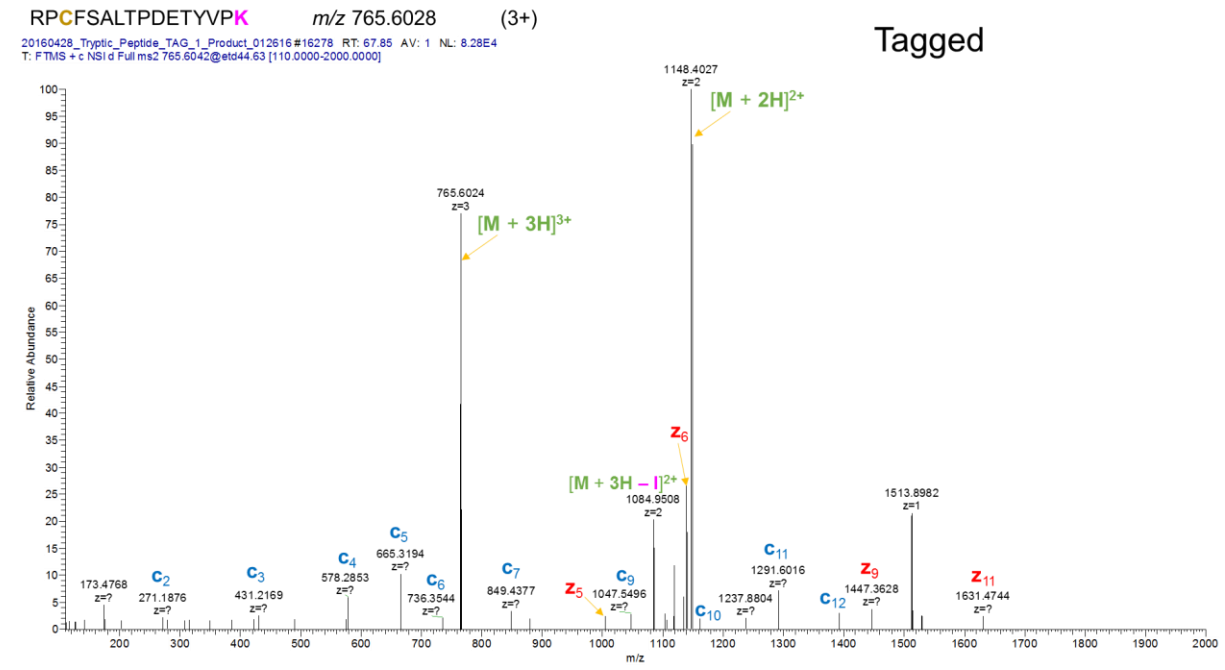


Figure 5.20. ETD-MS2 of Derivatized Trypsin-Digested RPCFSALTPDETYVPK Detected in Tagged Sample by Database Search. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow)

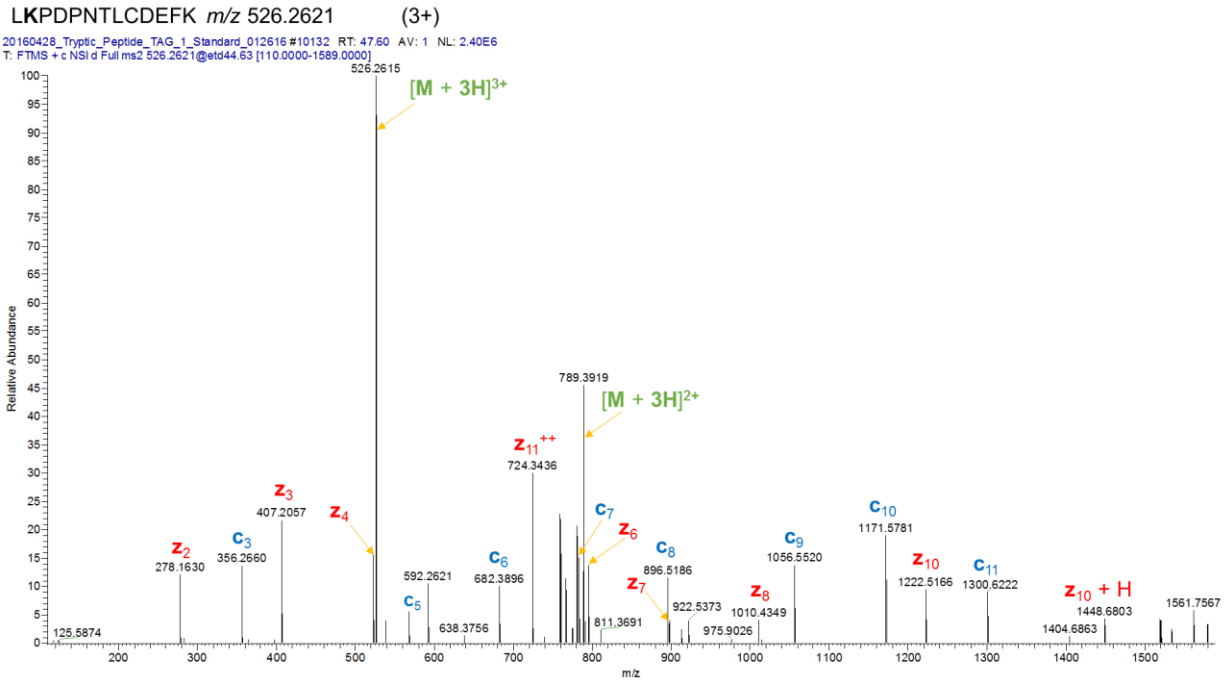


Figure 5.21. ETD-MS2 of Underivatized Trypsin-Digested, miscleaved LKPDPNTLCDEFK Detected in Standard by Database Search.

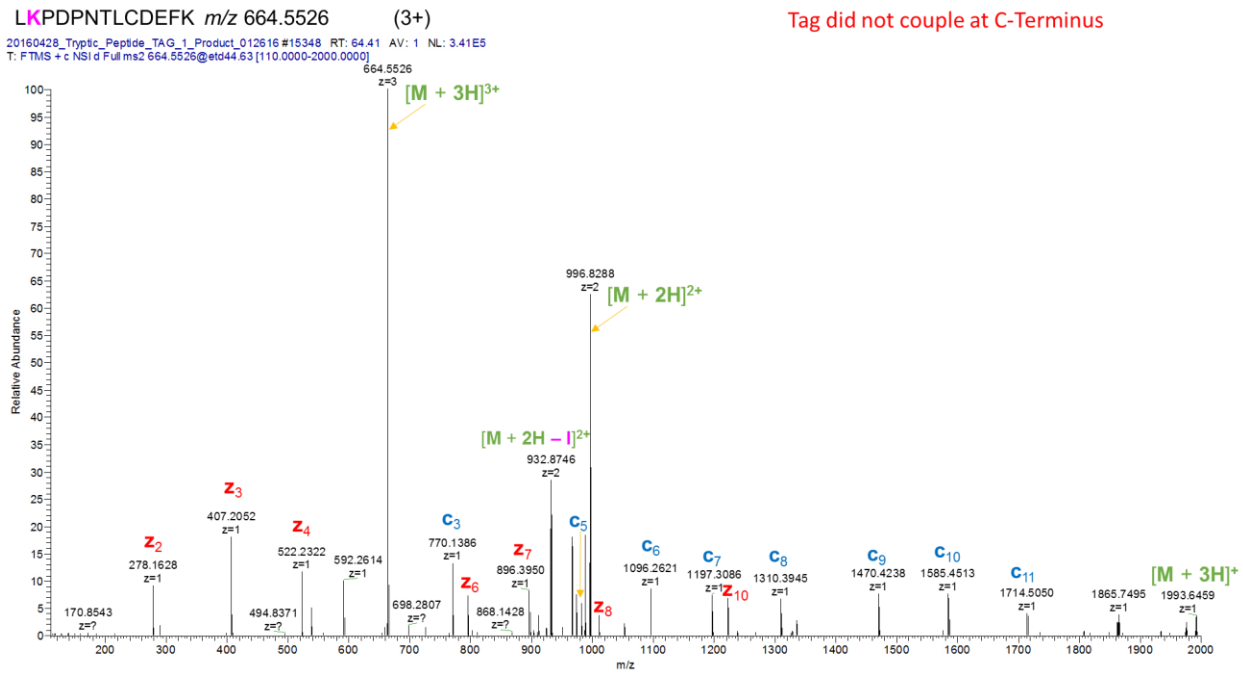


Figure 5.22. ETD-MS2 of Derivatized Trypsin-Digested, miscleaved LKPDPNTLCDEFK Detected in Tagged Sample by Database Search. (Note that in the case illustrated above the tag was found on the Lys residue at the site of miscleavage rather than at the C-terminal Lys)

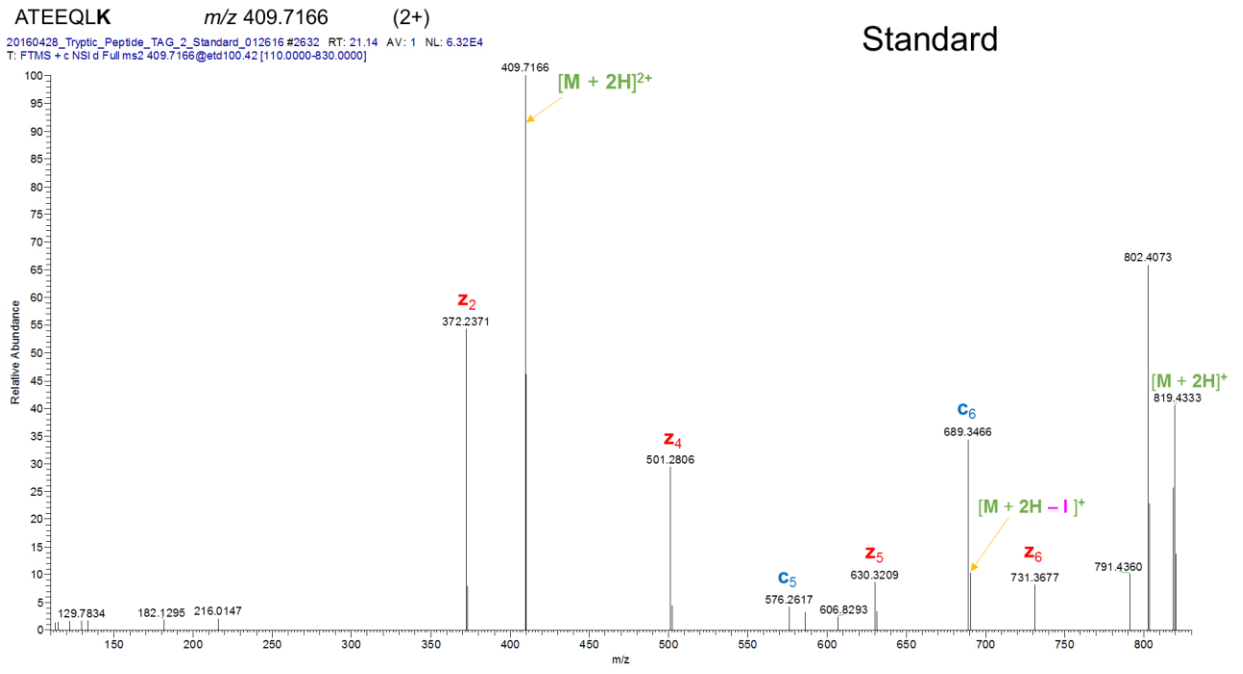


Figure 5.23. ETD-MS2 of Underivatized Trypsin-Digested ATEEQLK Detected in Standard by Database Search.

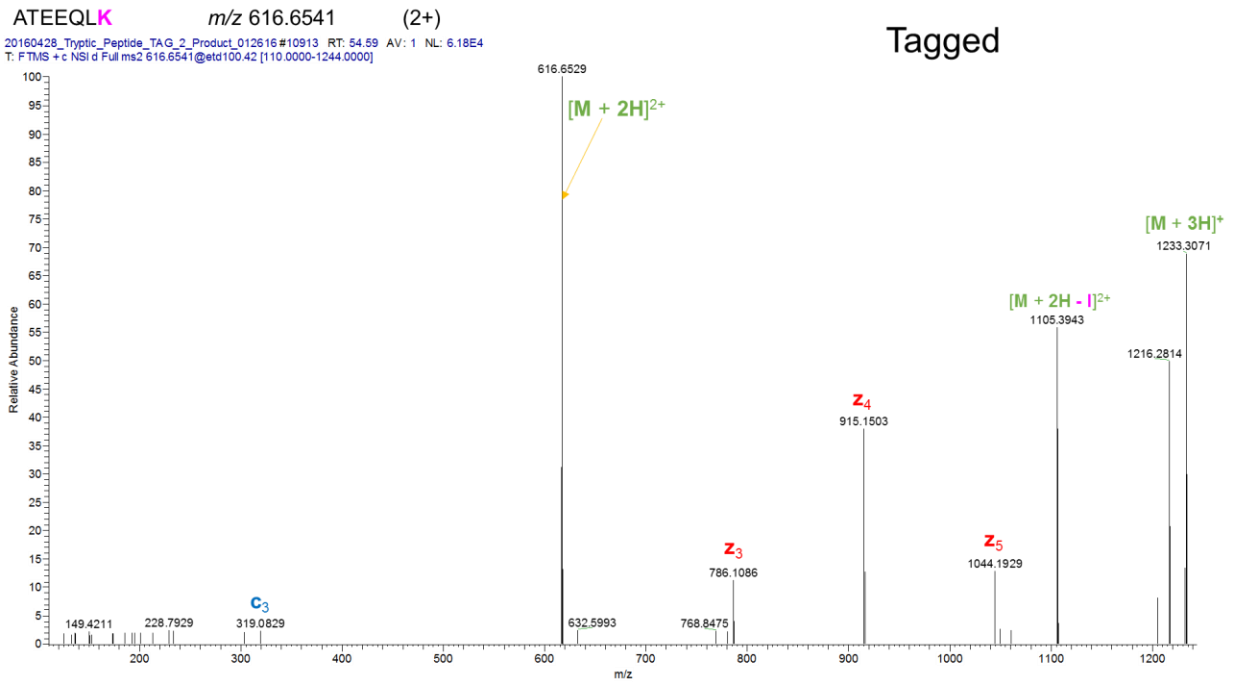


Figure 5.24. ETD-MS2 of Derivatized Trypsin-Digested ATEEQLK Detected in Tagged Sample by Database Search.

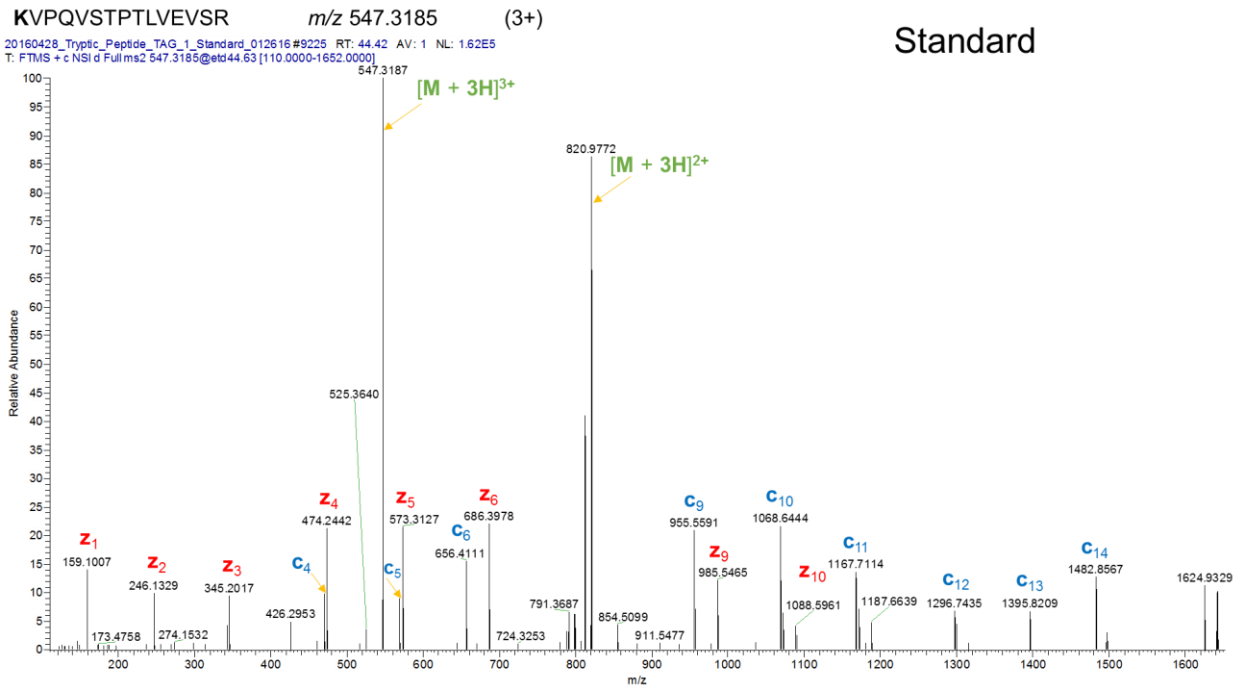


Figure 5.25. ETD-MS2 of Underivatized Trypsin-Digested KVPQVSTPTLVEVSR Detected in Standard by Database Search.

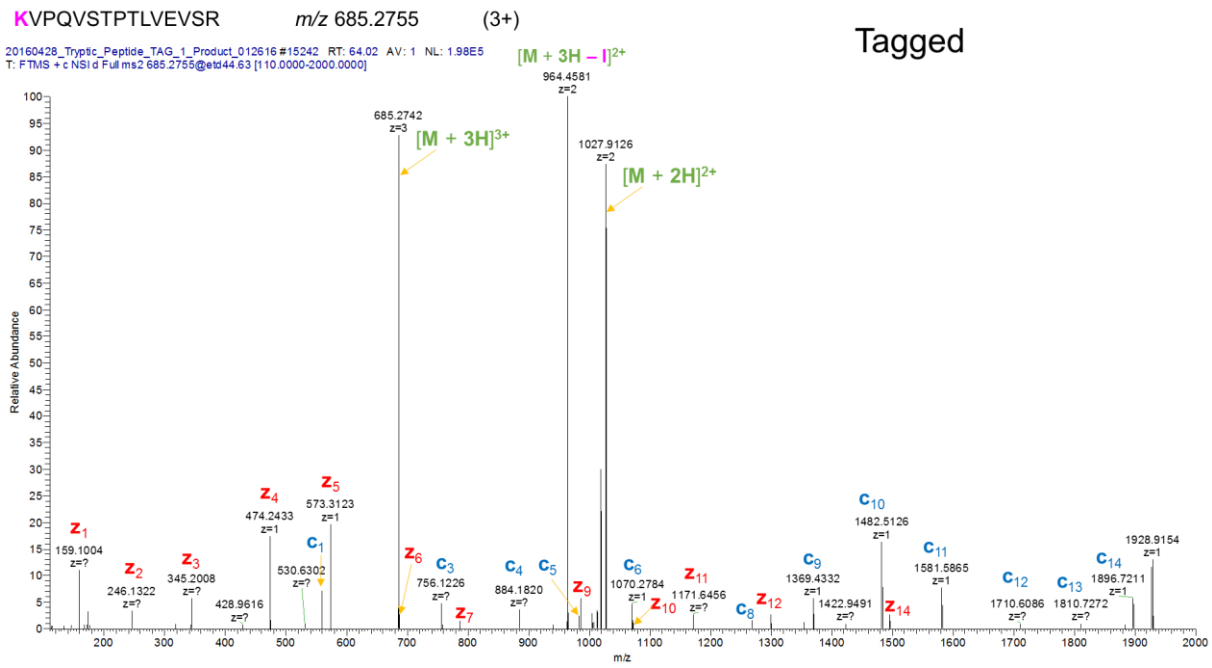


Figure 5.26. ETD-MS2 of Underivatized Trypsin-Digested KVPQVSTPTLVEVSR Detected in Standard by Database Search. (Note that in the case illustrated above the N-terminal lysine is tagged and the peptide is Arg-terminated, meaning it remained in solution despite the washing steps)

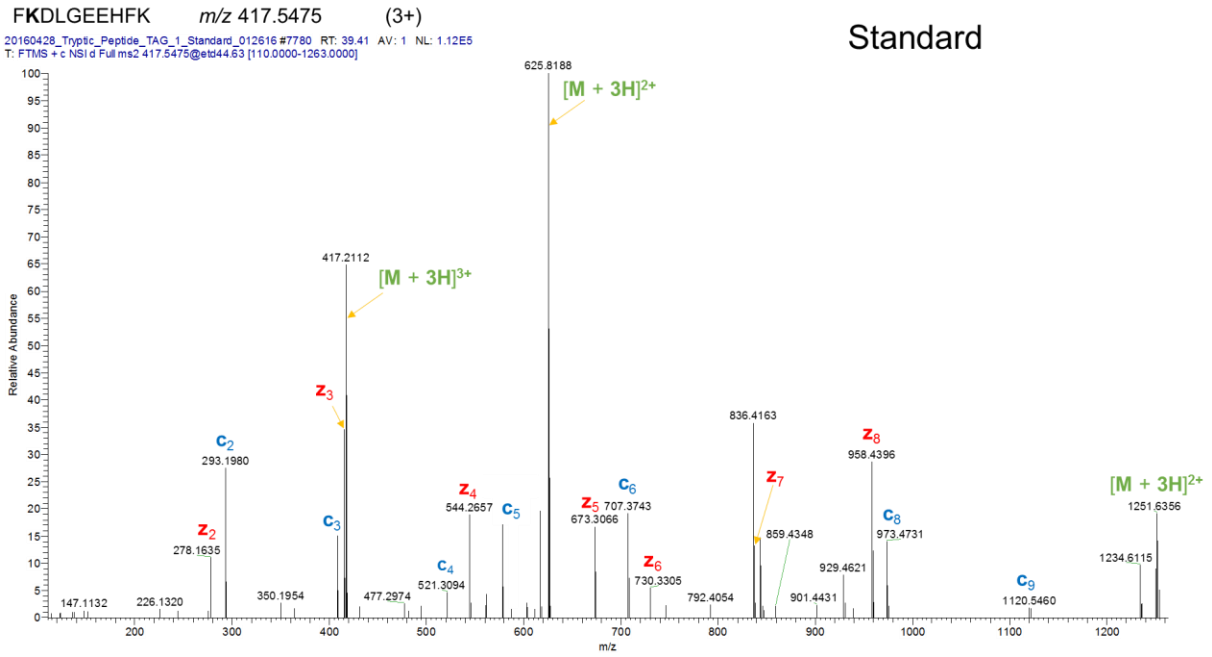


Figure 5.27. ETD-MS2 of Underivatized Trypsin-Digested FKDLGEEHFK Detected in Standard by Database Search.

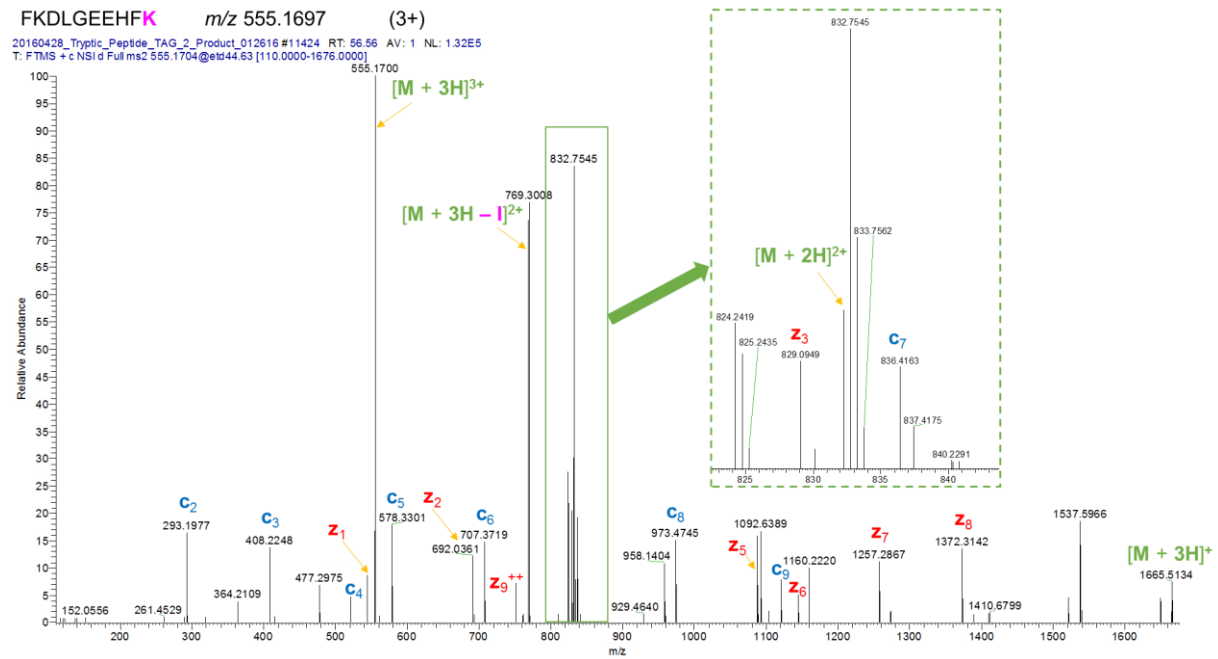


Figure 5.28. ETD-MS2 of Derivatized Trypsin-Digested FKDLGEEHFK Detected in Tagged Sample by Database Search.

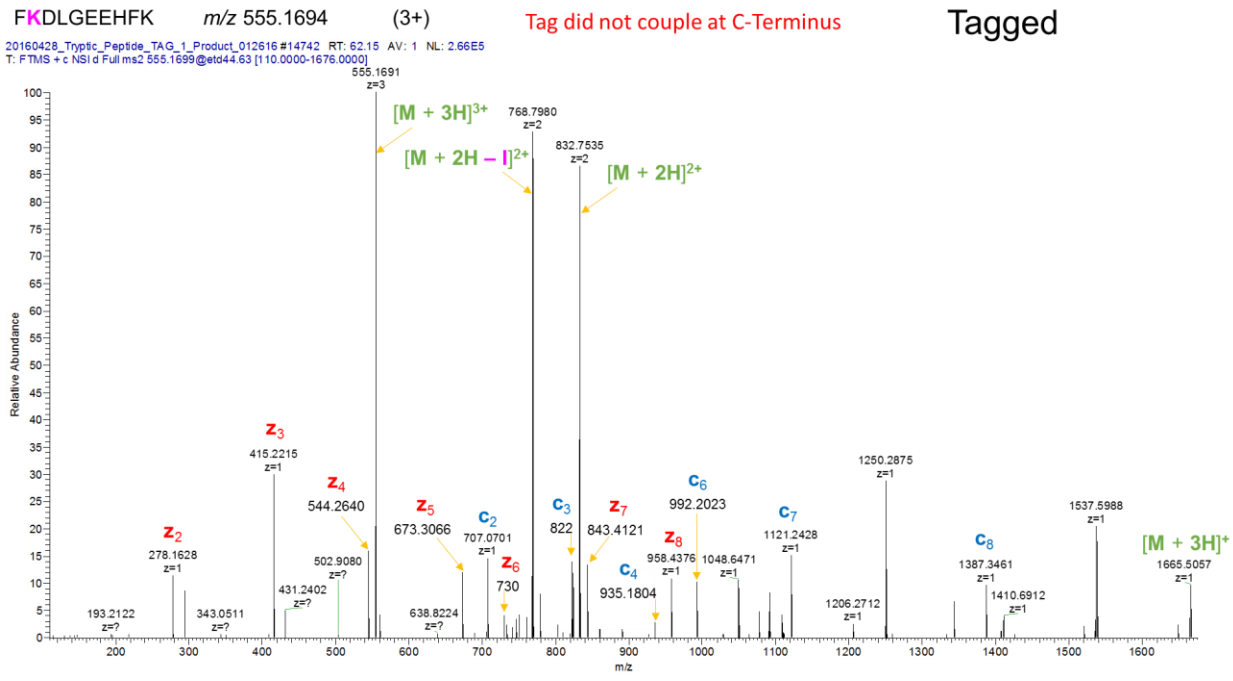


Figure 5.29. ETD-MS2 of Derivatized Trypsin-Digested FKDLGEEHFK Detected in Tagged Sample by Database Search. Peptide Tagged at Miscleavage Site, not at C-terminal LysC.

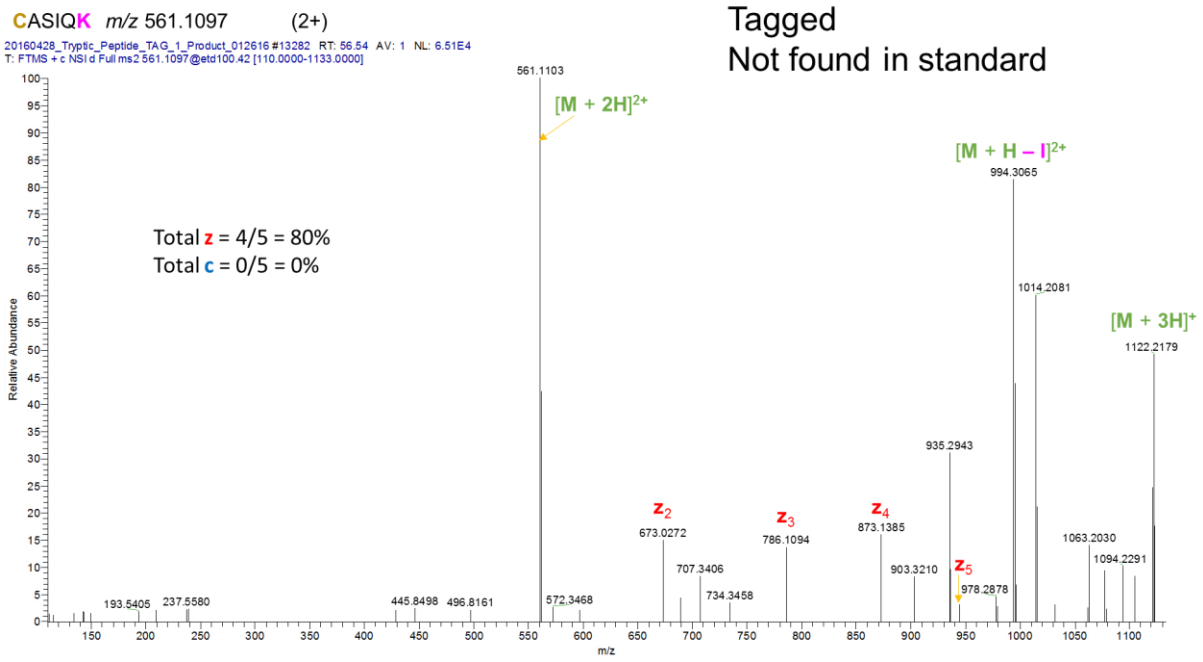


Figure 5.30. ETD-MS2 of Derivatized Trypsin-Digested CASIQK Detected in Tagged Sample by Database Search. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow, to prevent the reformation of disulfide bonds)

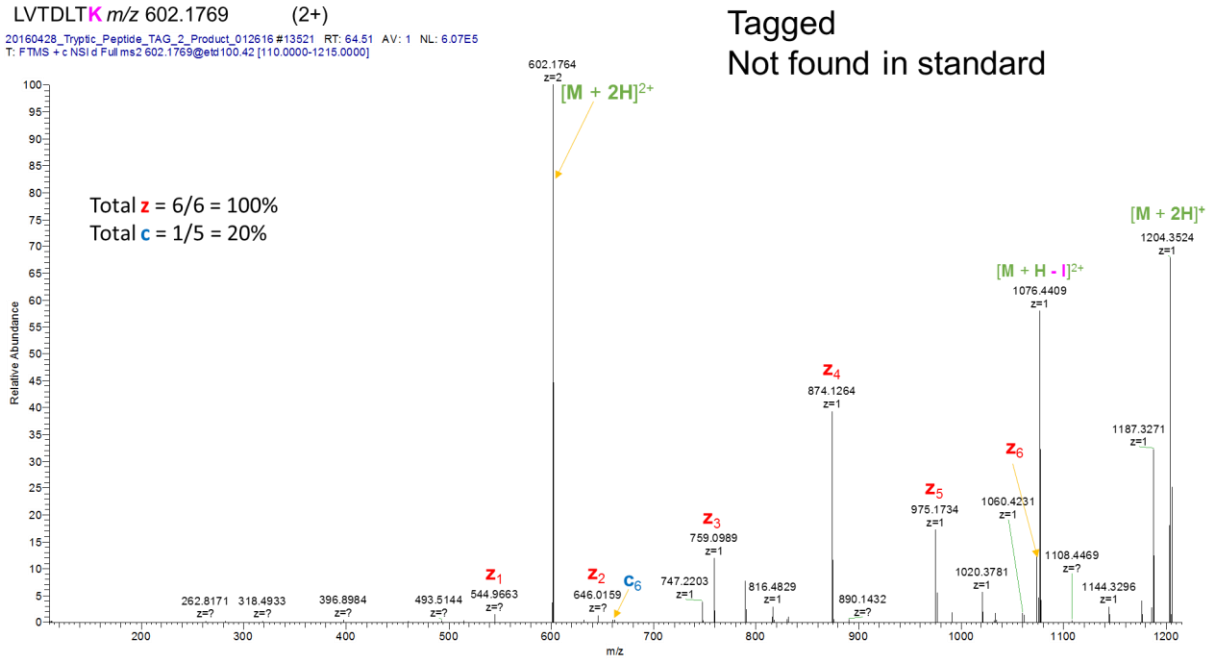


Figure 5.31. ETD-MS2 of Derivatized Trypsin-Digested LVTDLTK Detected in Tagged Sample by Database Search. (Note that this peptide was detected in the tagged sample but was not detected in the digest standard)

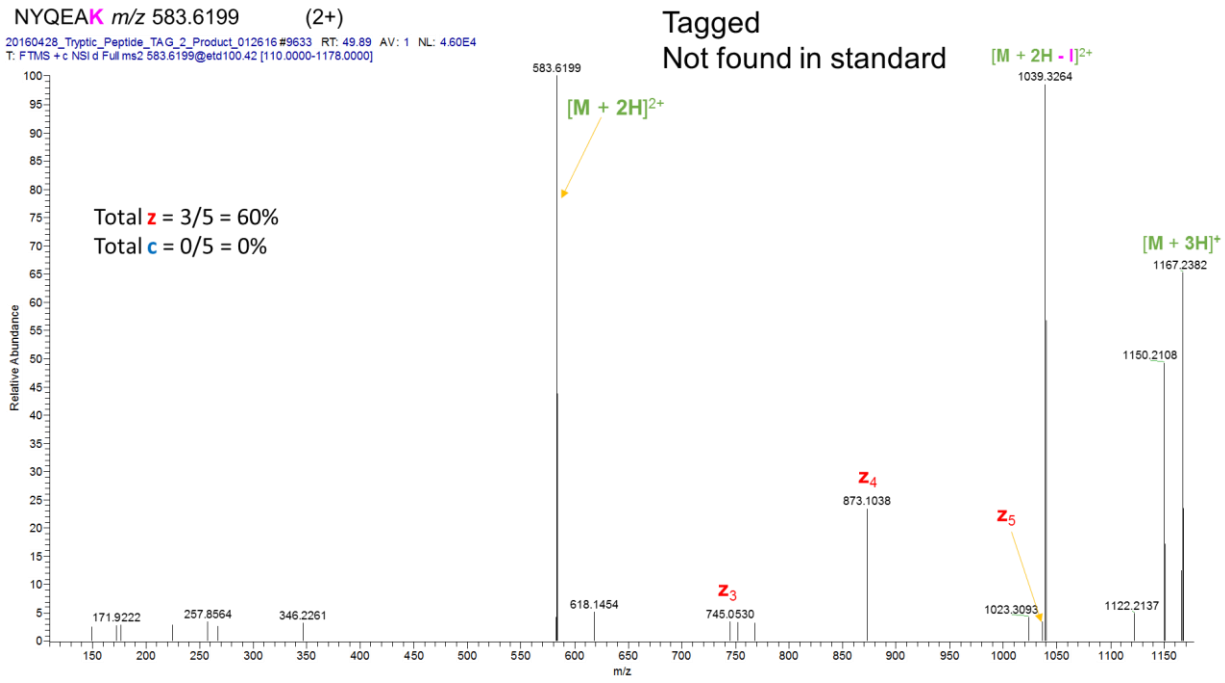


Figure 5.32. ETD-MS2 of Derivatized Trypsin-Digested NYQEAK Detected in Tagged Sample by Database Search.

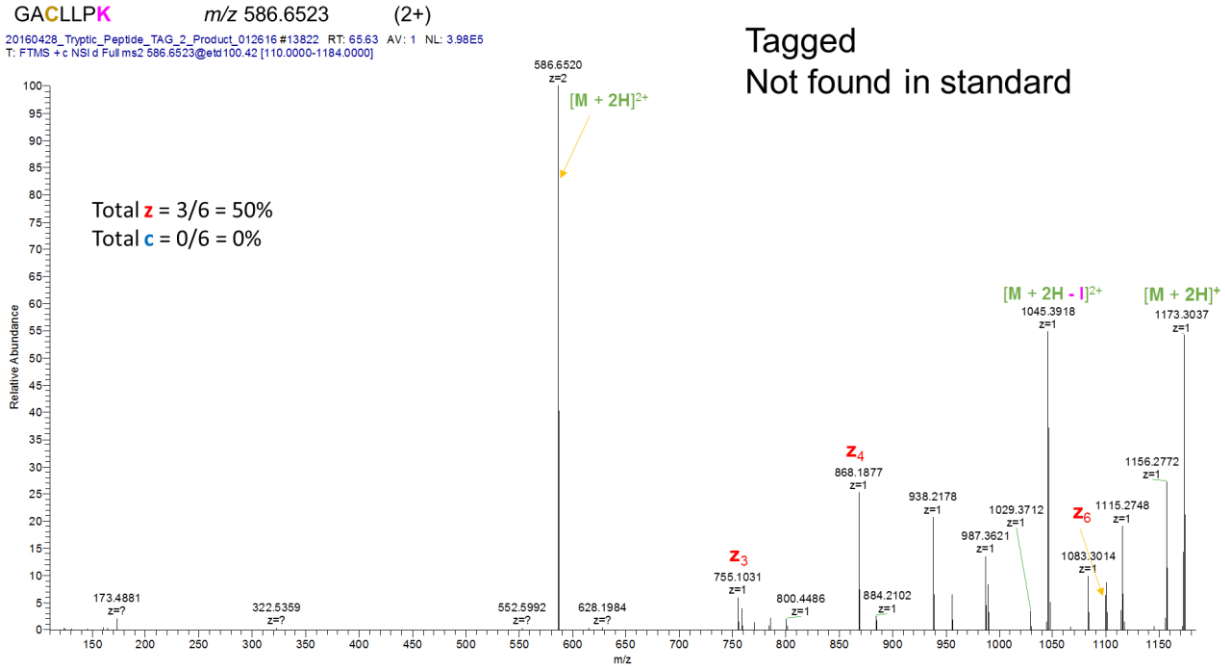


Figure 5.33. ETD-MS2 of Derivatized Trypsin-Digested GACLLPK Detected in Tagged Sample. (Note that this peptide is alkylated at the cysteine residue, highlighted in yellow, to prevent the reformation of disulfide bonds)

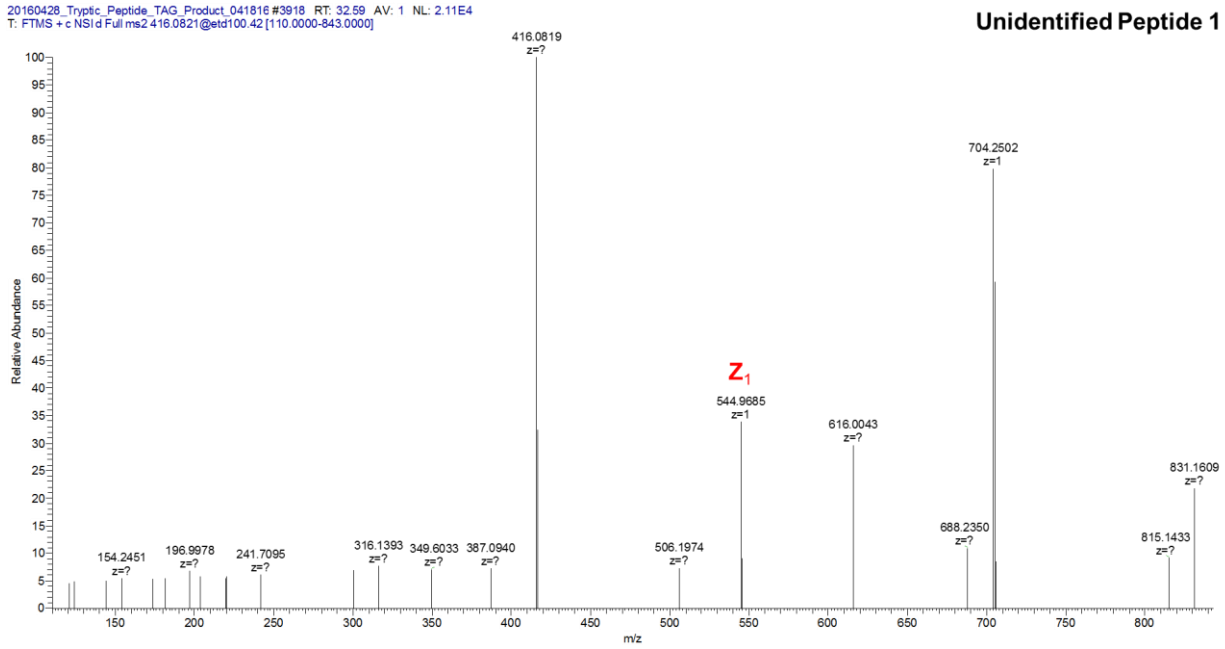


Figure 5.34. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 1, Undetected by Database but Detected Using In-House Developed Script.

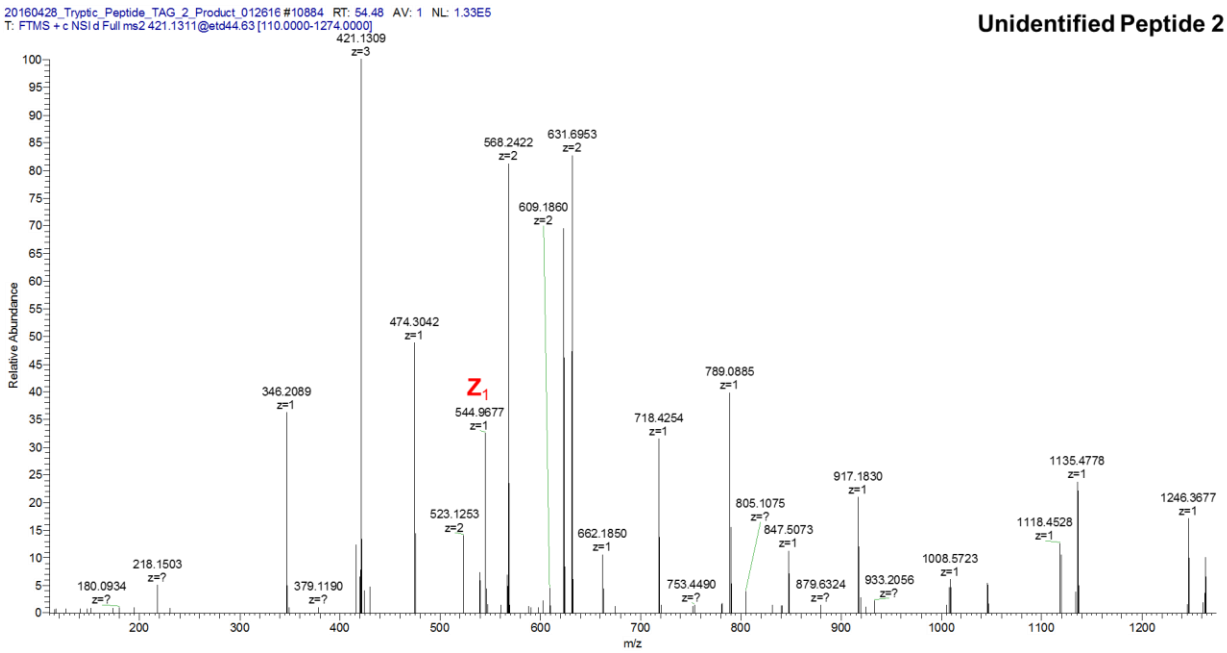


Figure 5.35. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 2, Undetected by Database but Detected Using In-House Developed Script.

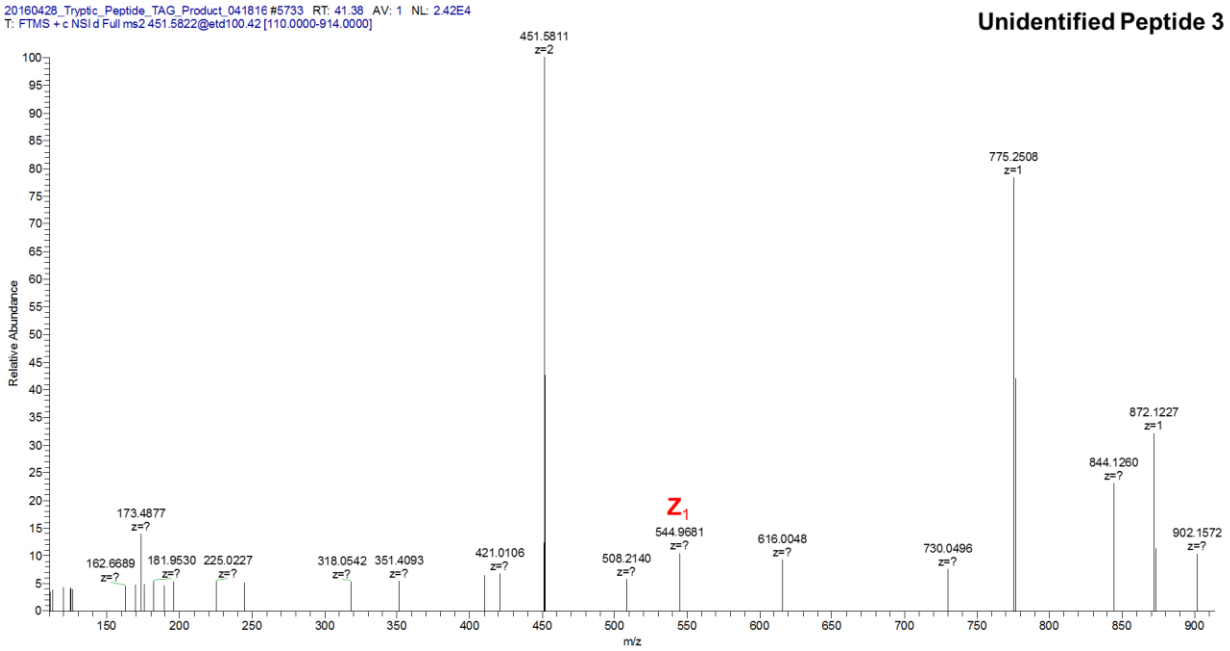


Figure 5.36. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 3, Undetected by Database but Detected Using In-House Developed Script.

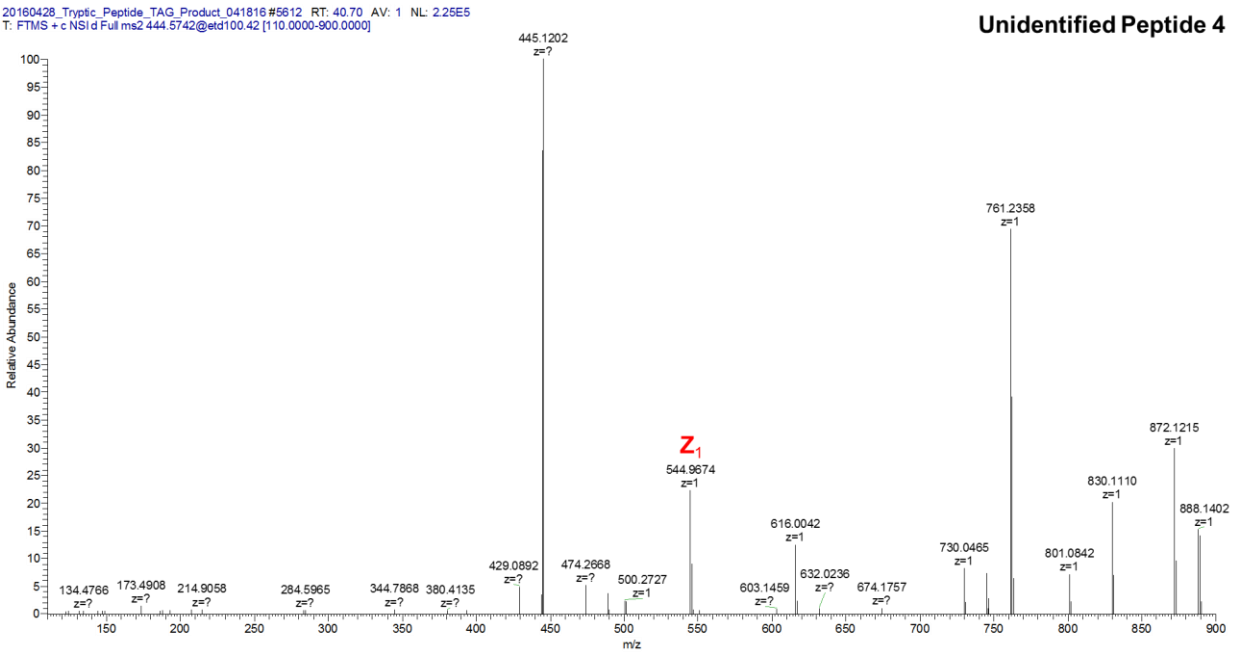


Figure 5.37. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 4, Undetected by Database but Detected Using In-House Developed Script.

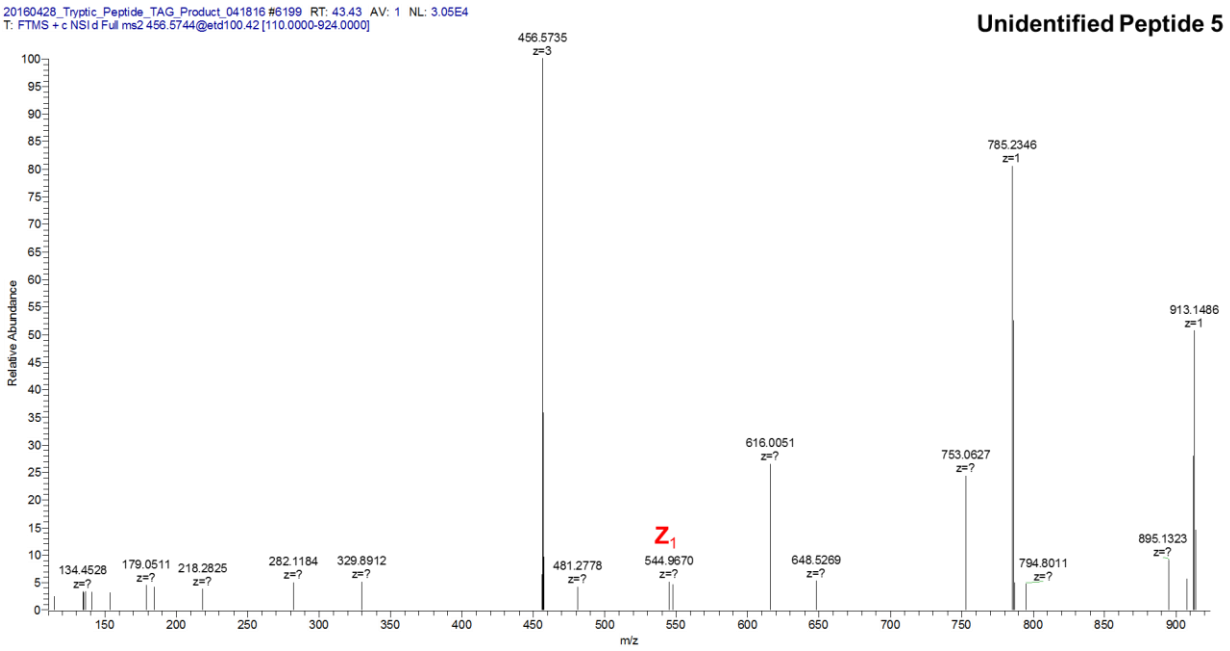


Figure 5.38. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 5, Undetected by Database but Detected Using In-House Developed Script.

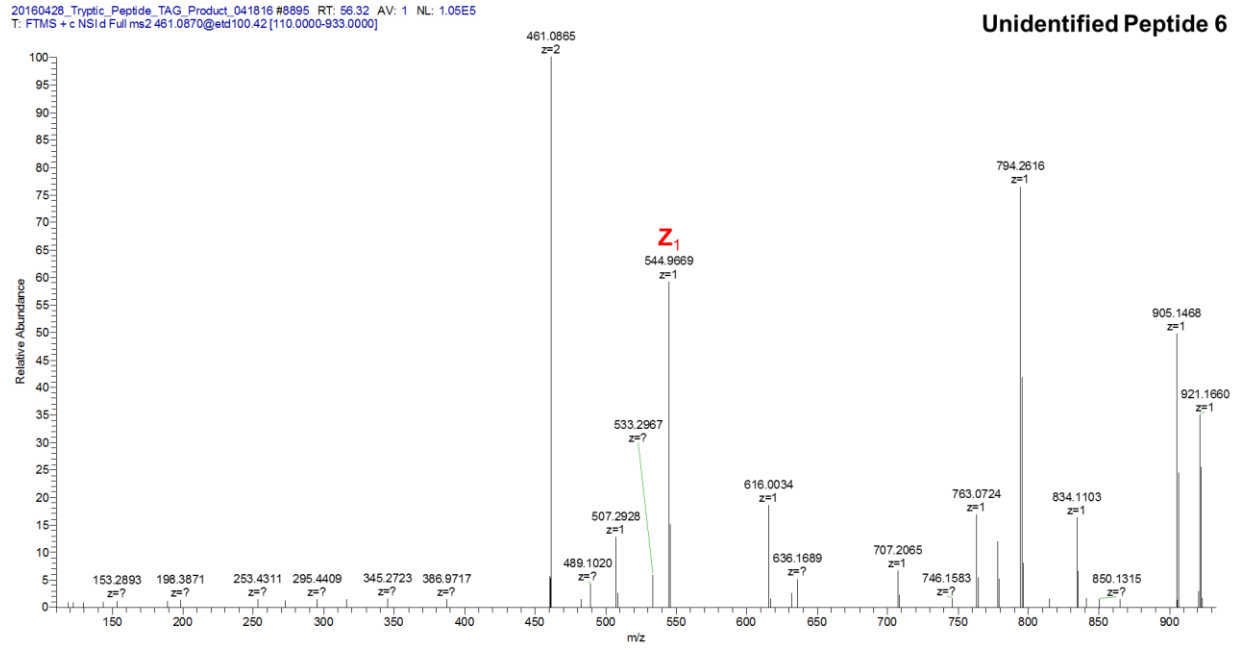


Figure 5.39. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 6, Undetected by Database but Detected Using In-House Developed Script.

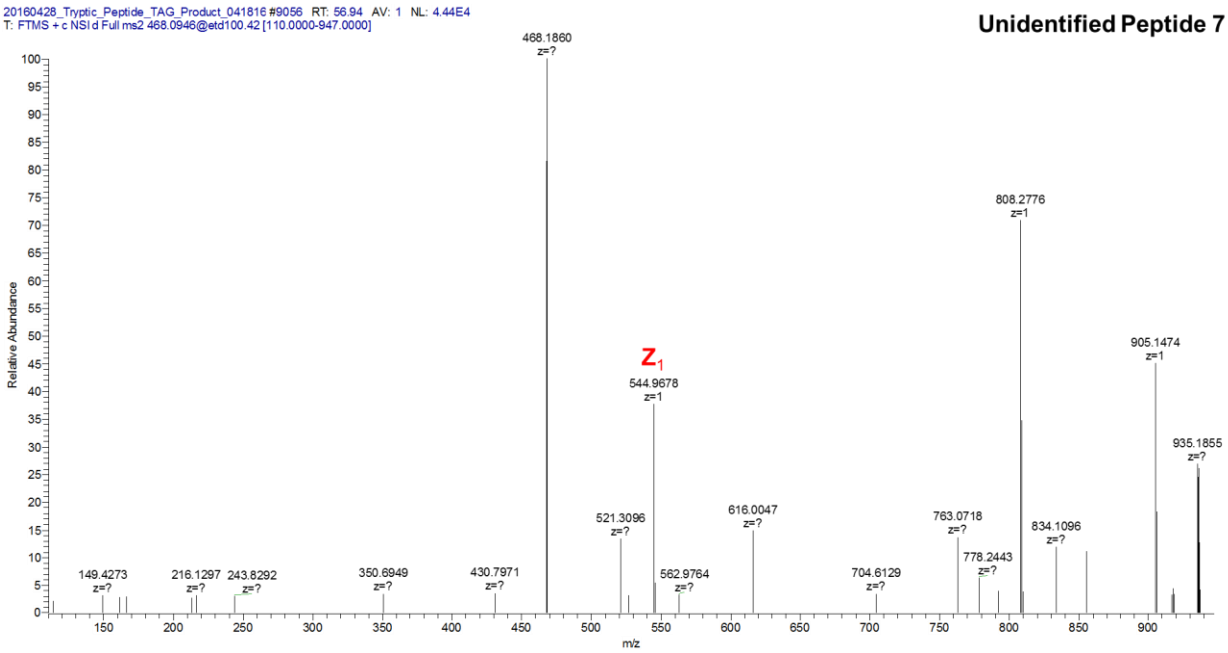


Figure 5.40. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 7, Undetected by Database but Detected Using In-House Developed Script.

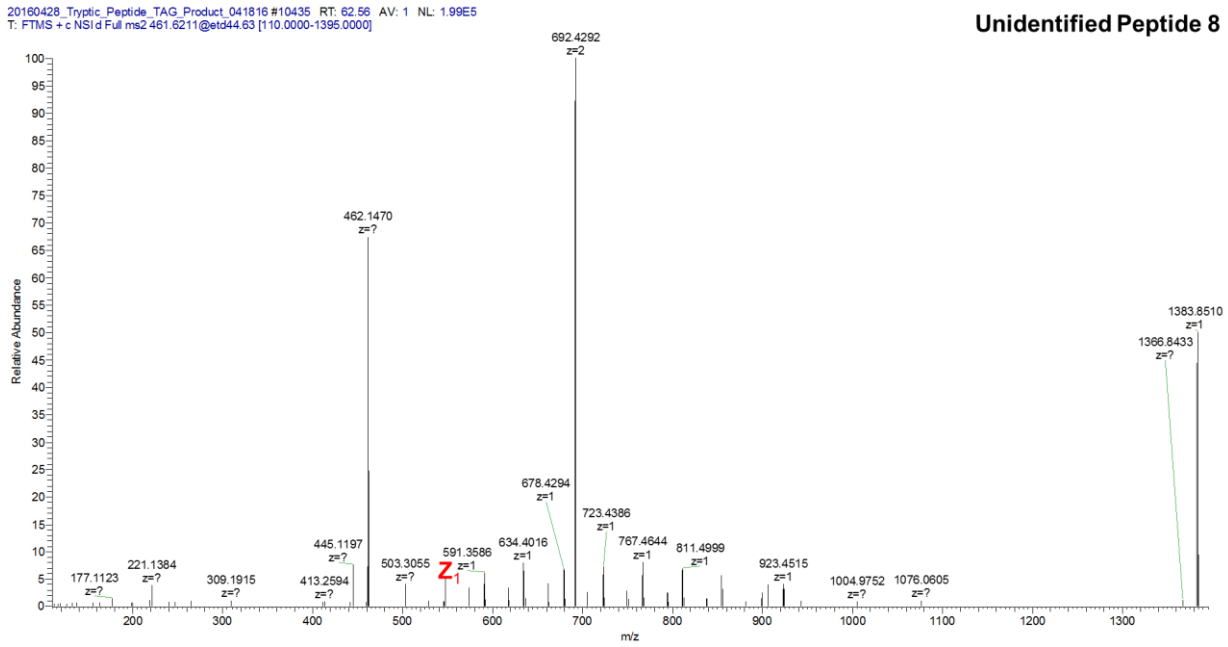


Figure 5.41. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 8, Undetected by Database but Detected Using In-House Developed Script.

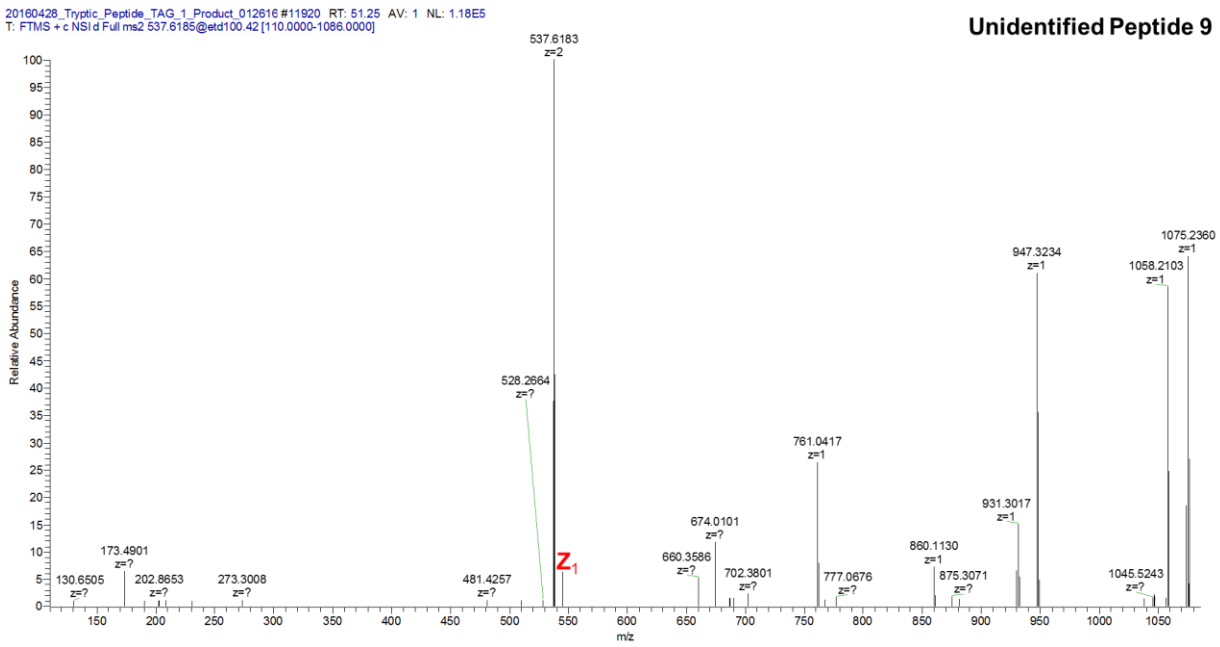


Figure 5.42. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 9, Undetected by Database but Detected Using In-House Developed Script.

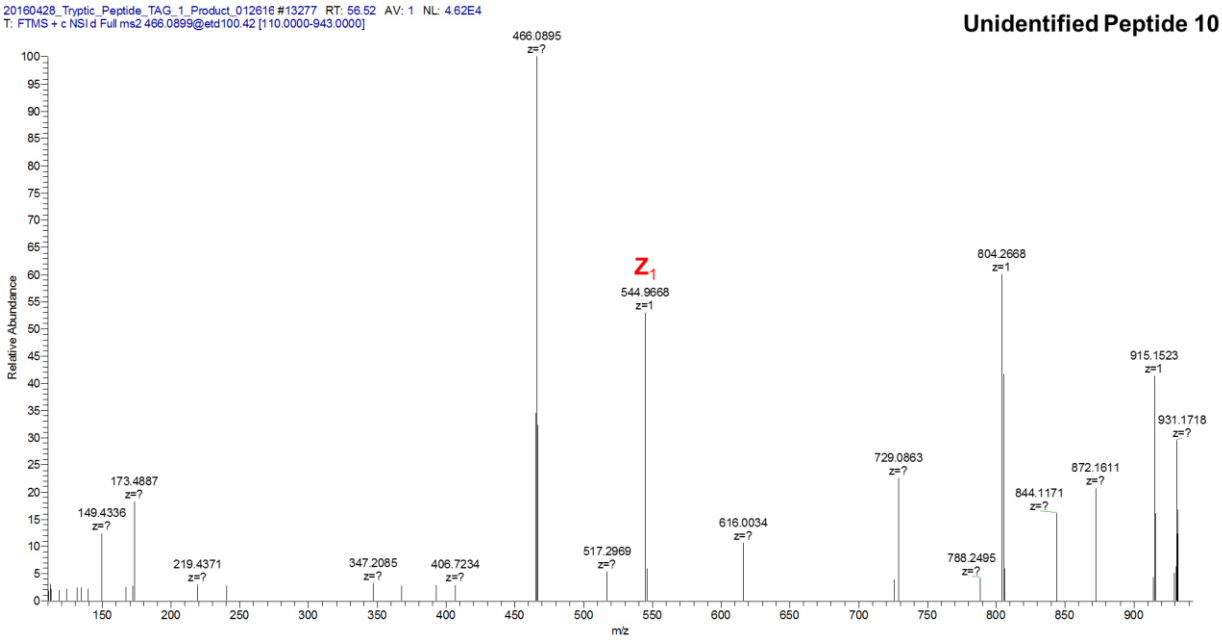


Figure 5.43. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 10, Undetected by Database but Detected Using In-House Developed Script.

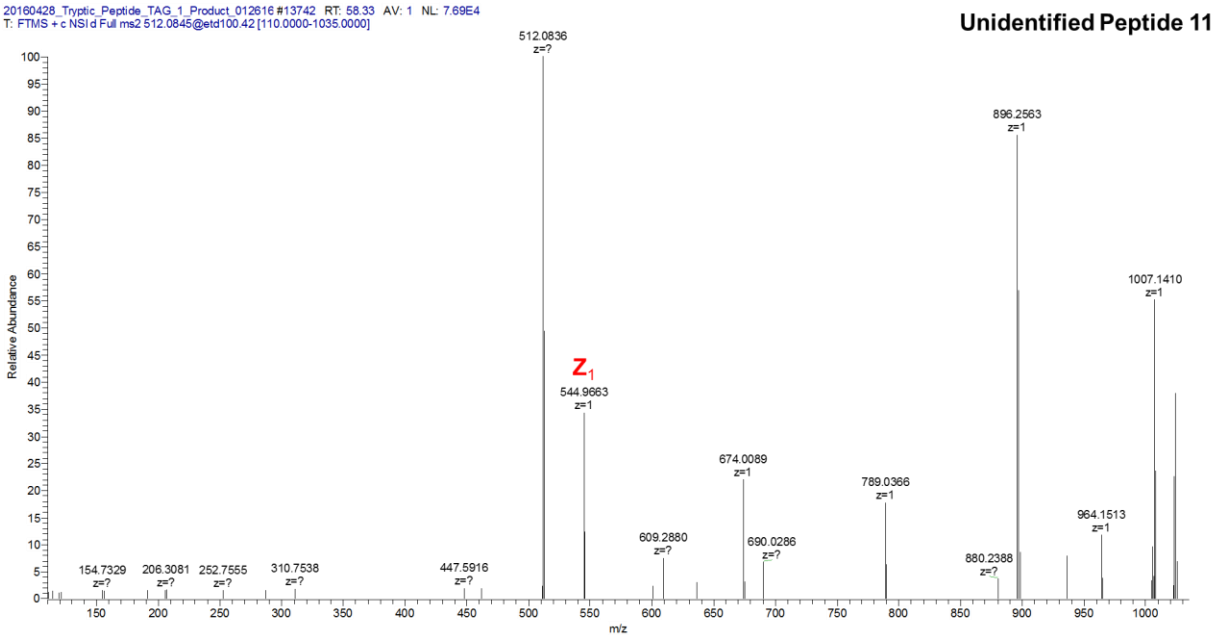


Figure 5.44. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 11, Undetected by Database but Detected Using In-House Developed Script.

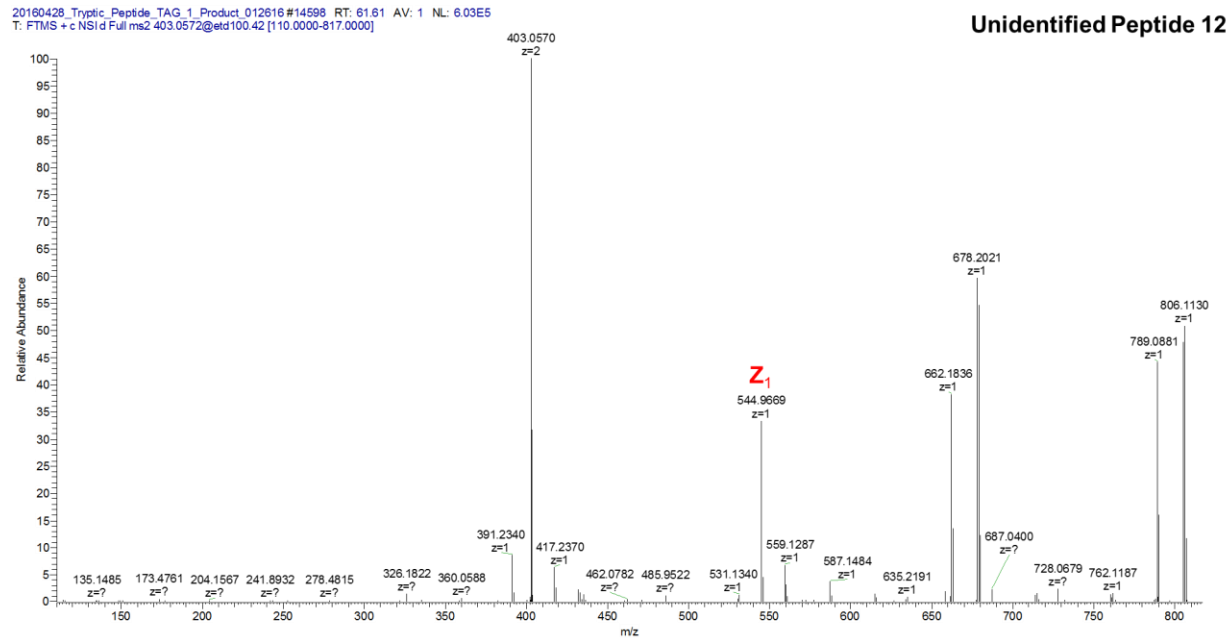


Figure 5.45. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 12, Undetected by Database but Detected Using In-House Developed Script.

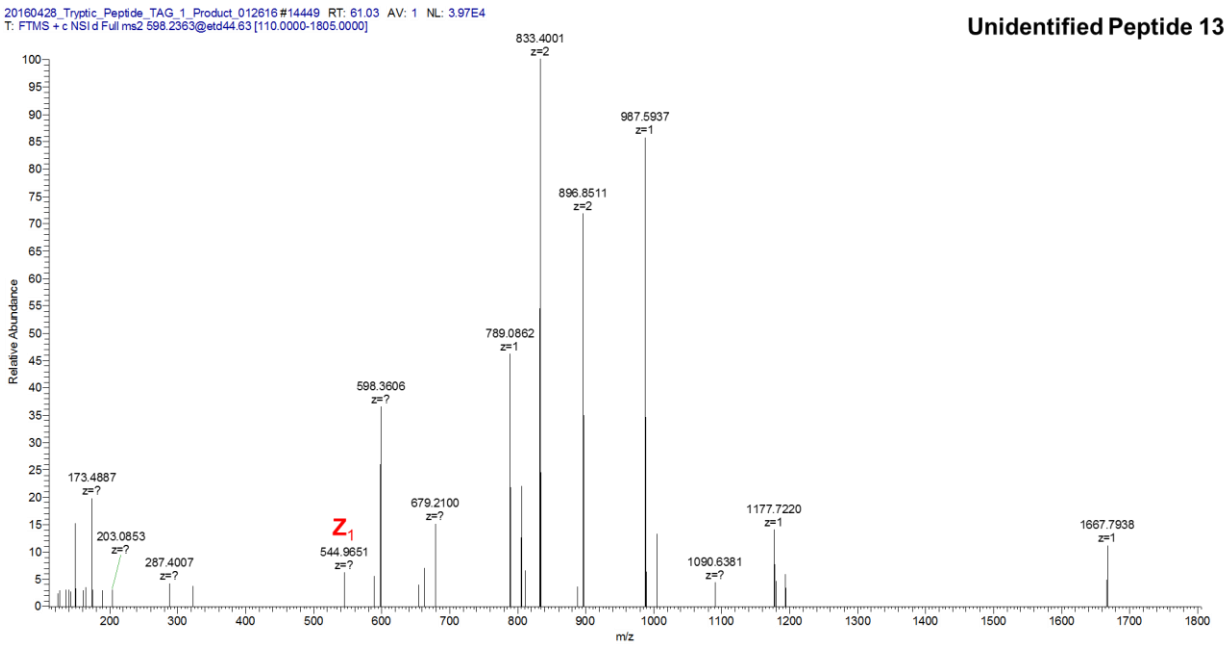


Figure 5.46. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 13, Undetected by Database but Detected Using In-House Developed Script.

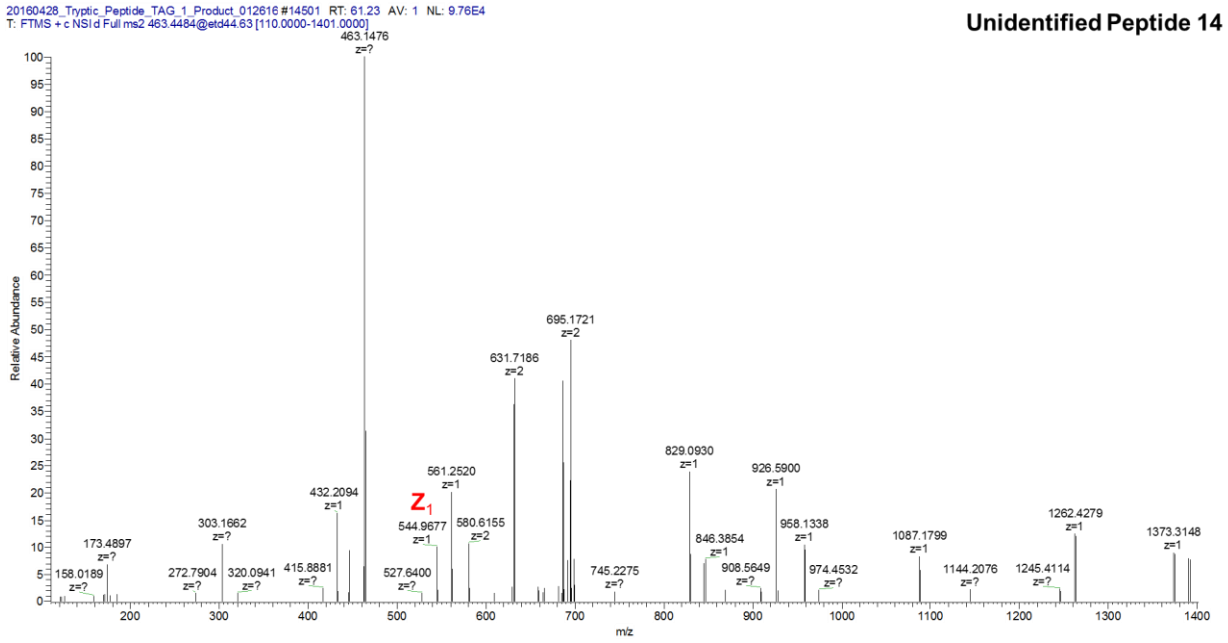


Figure 5.47. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 14, Undetected by Database but Detected Using In-House Developed Script.

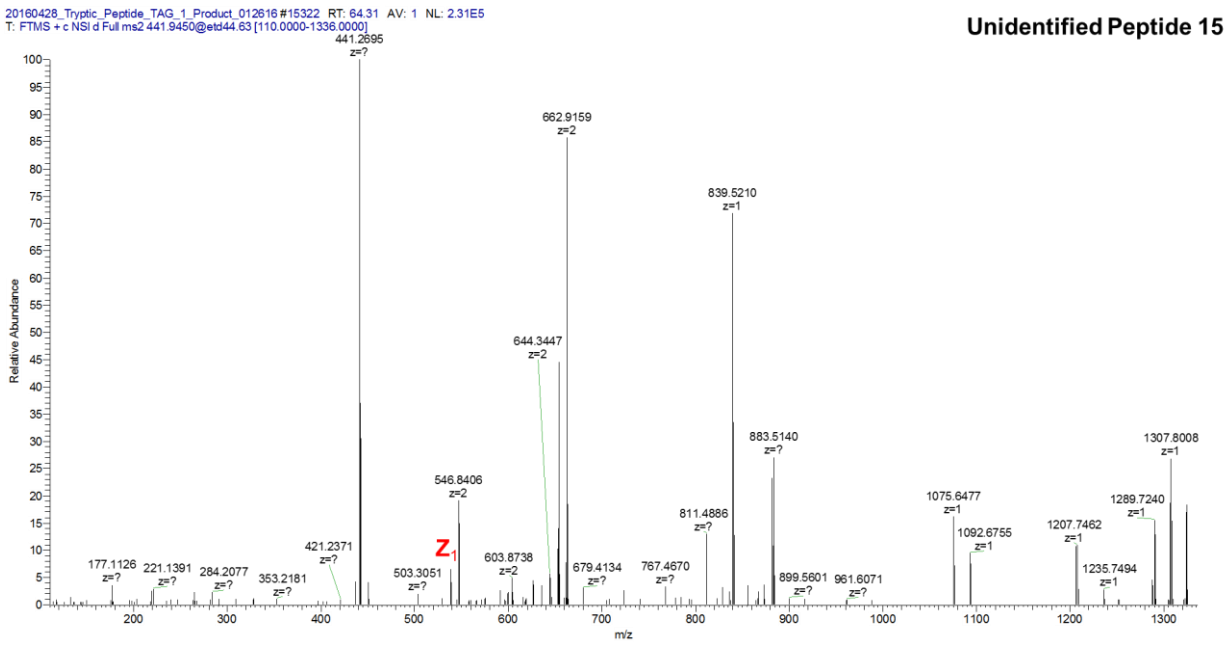


Figure 5.48. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 15, Undetected by Database but Detected Using In-House Developed Script.

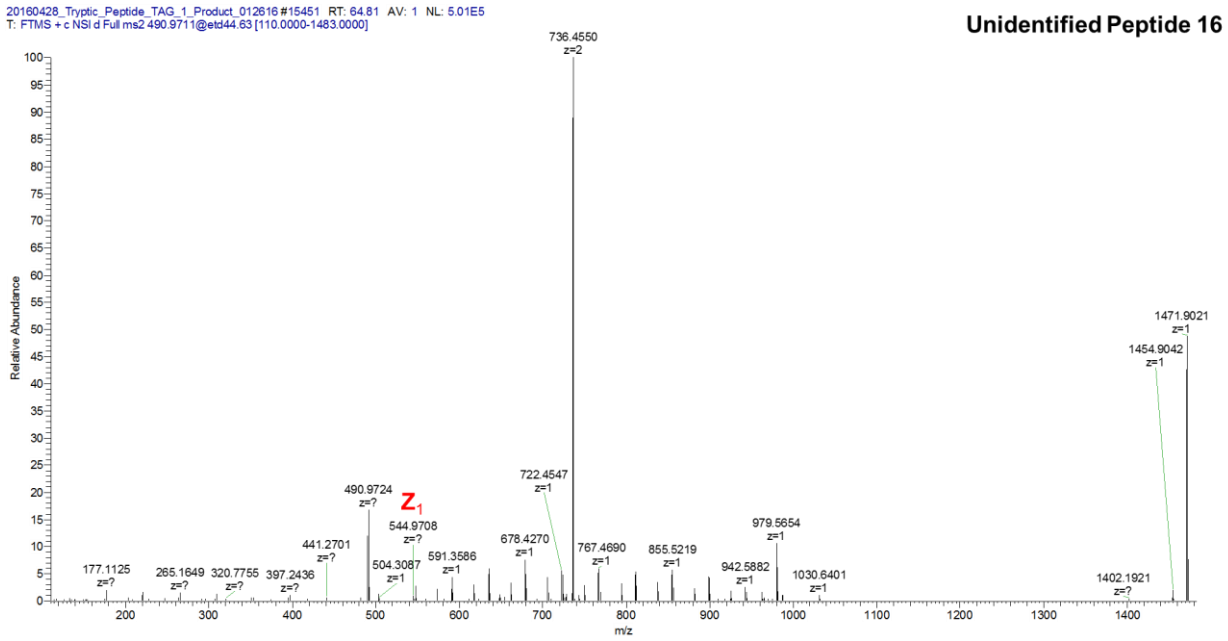


Figure 5.49. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 16, Undetected by Database but Detected Using In-House Developed Script.

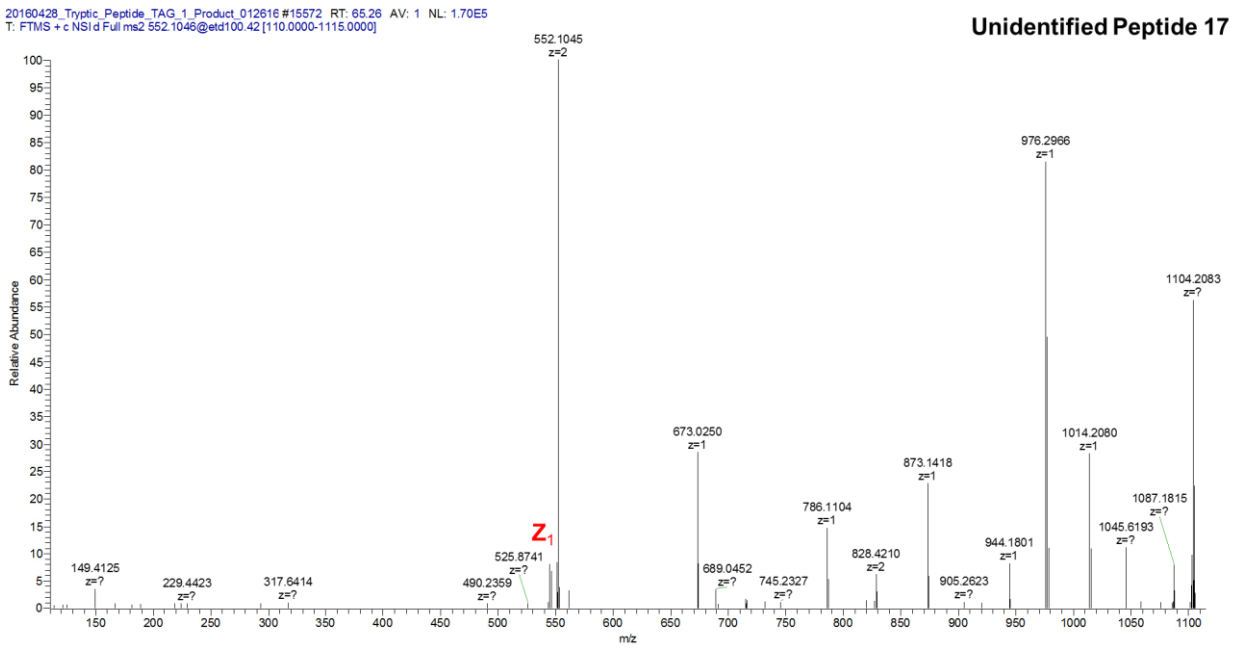


Figure 5.50. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 17, Undetected by Database but Detected Using In-House Developed Script.

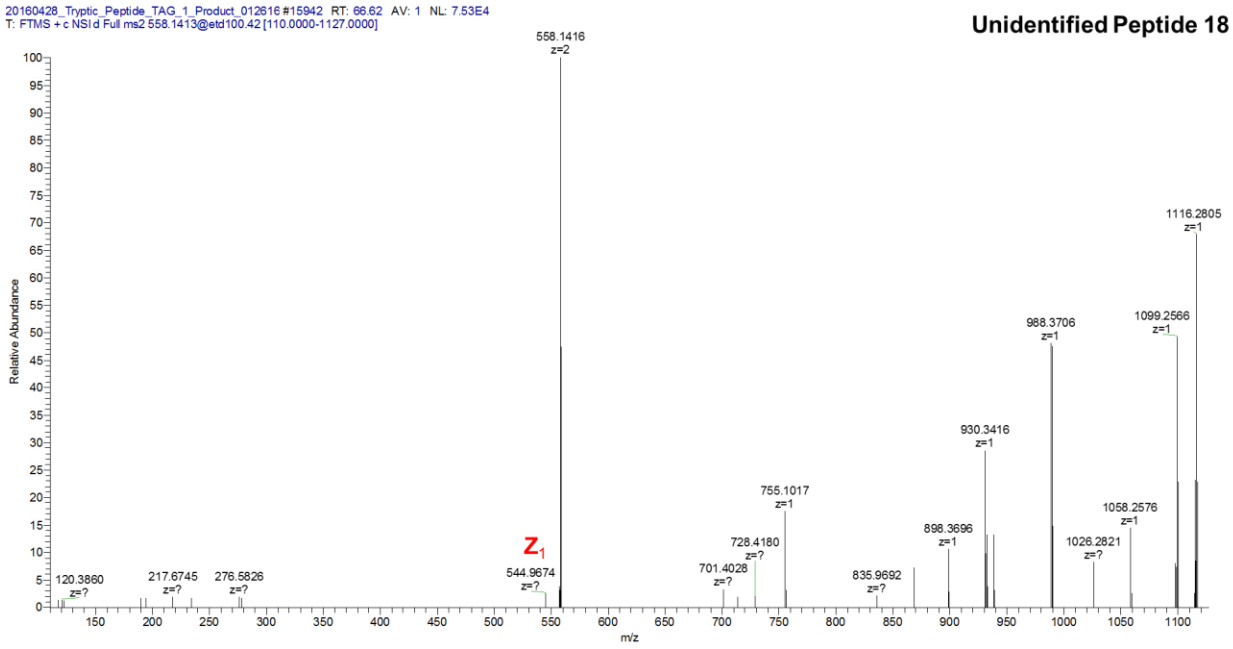


Figure 5.51. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 18, Undetected by Database but Detected Using In-House Developed Script.

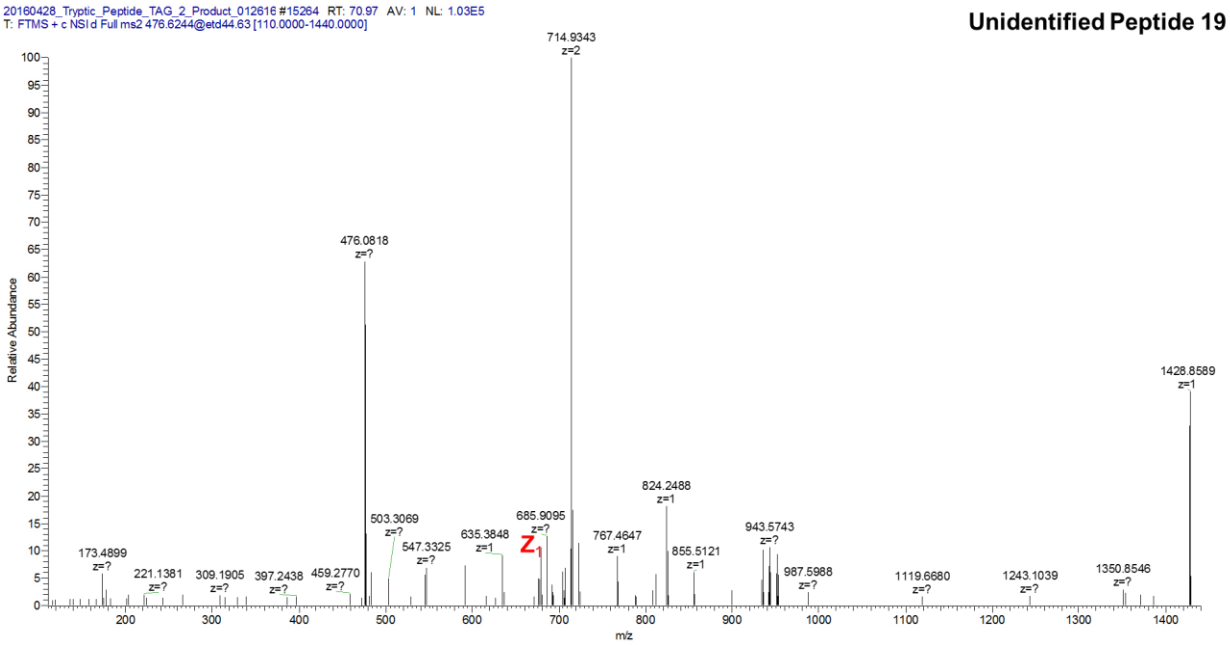


Figure 5.52. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 19, Undetected by Database but Detected Using In-House Developed Script.

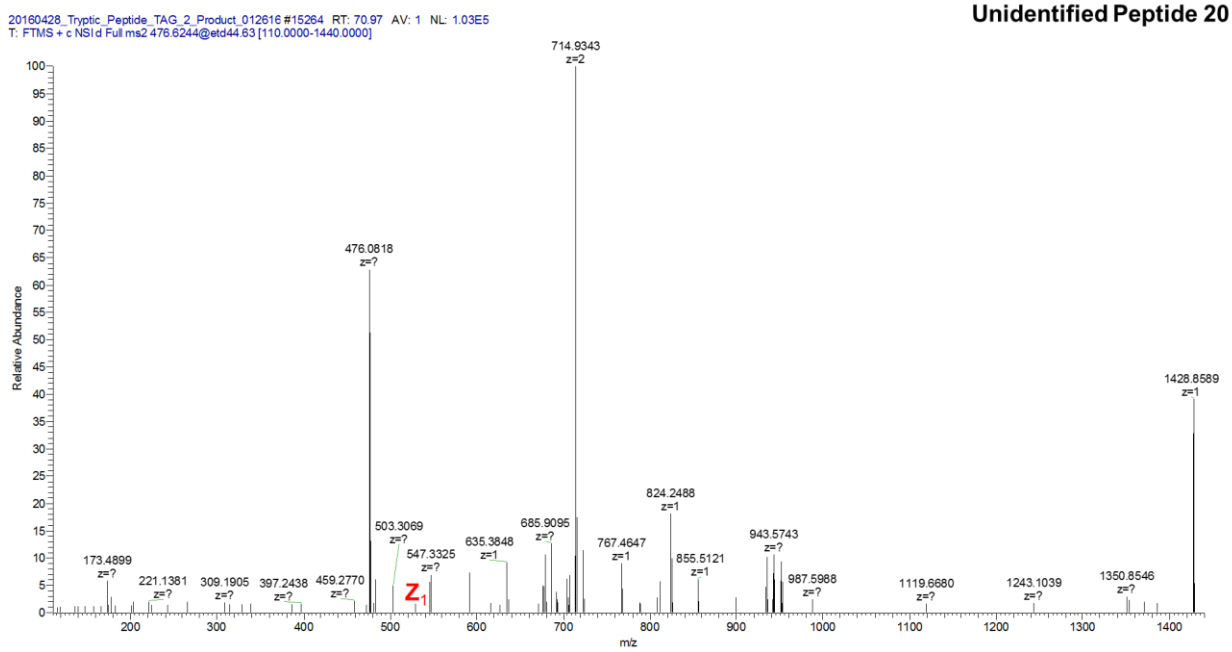


Figure 5.53. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 20, Undetected by Database but Detected Using In-House Developed Script.

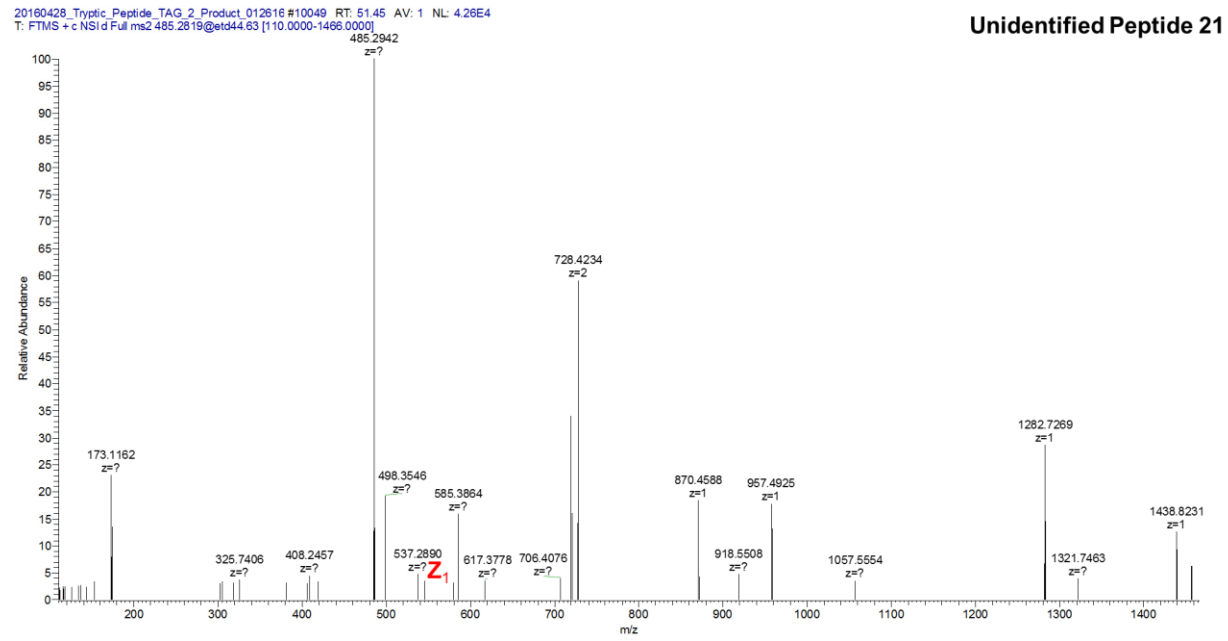


Figure 5.54. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 21, Undetected by Database but Detected Using In-House Developed Script.

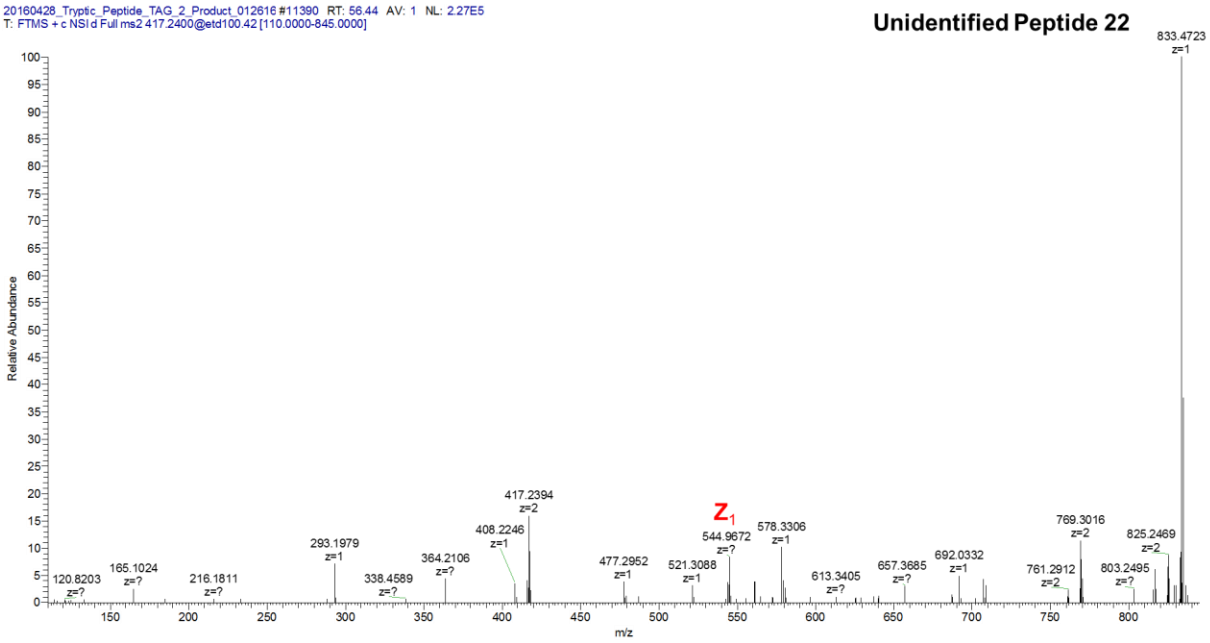


Figure 5.55. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 22, Undetected by Database but Detected Using In-House Developed Script.

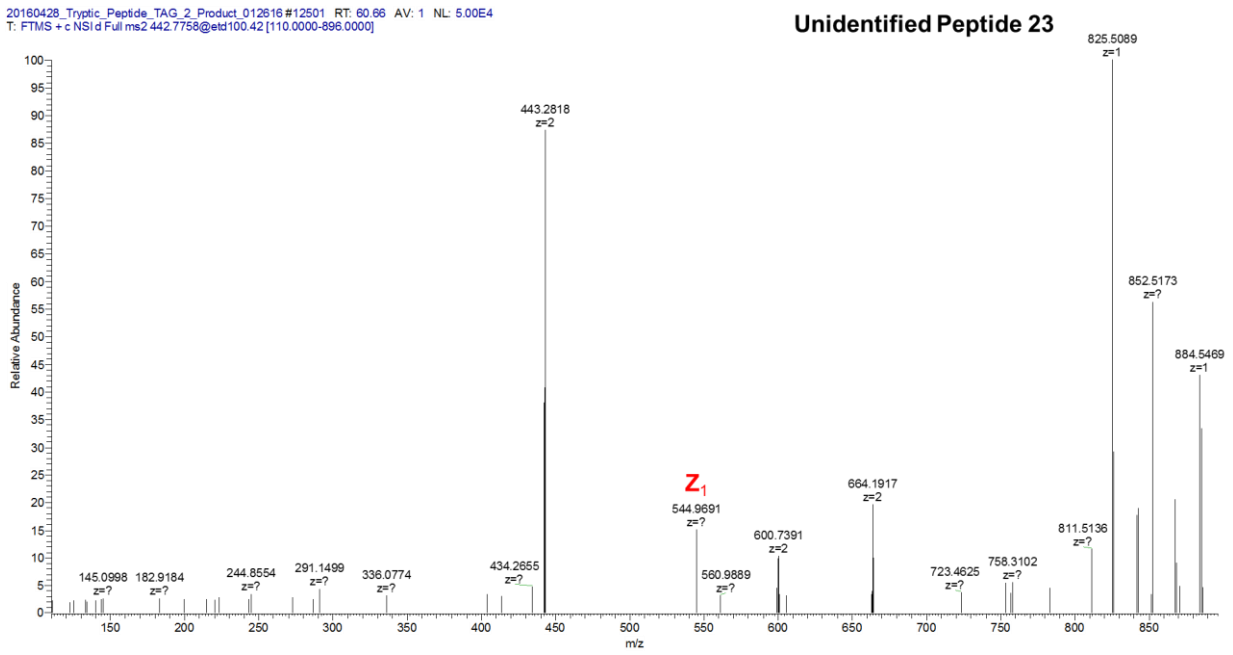


Figure 5.56. ETD-MS2 of Trypsin-Digested Unidentified Tagged Peptide 23, Undetected by Database but Detected Using In-House Developed Script.

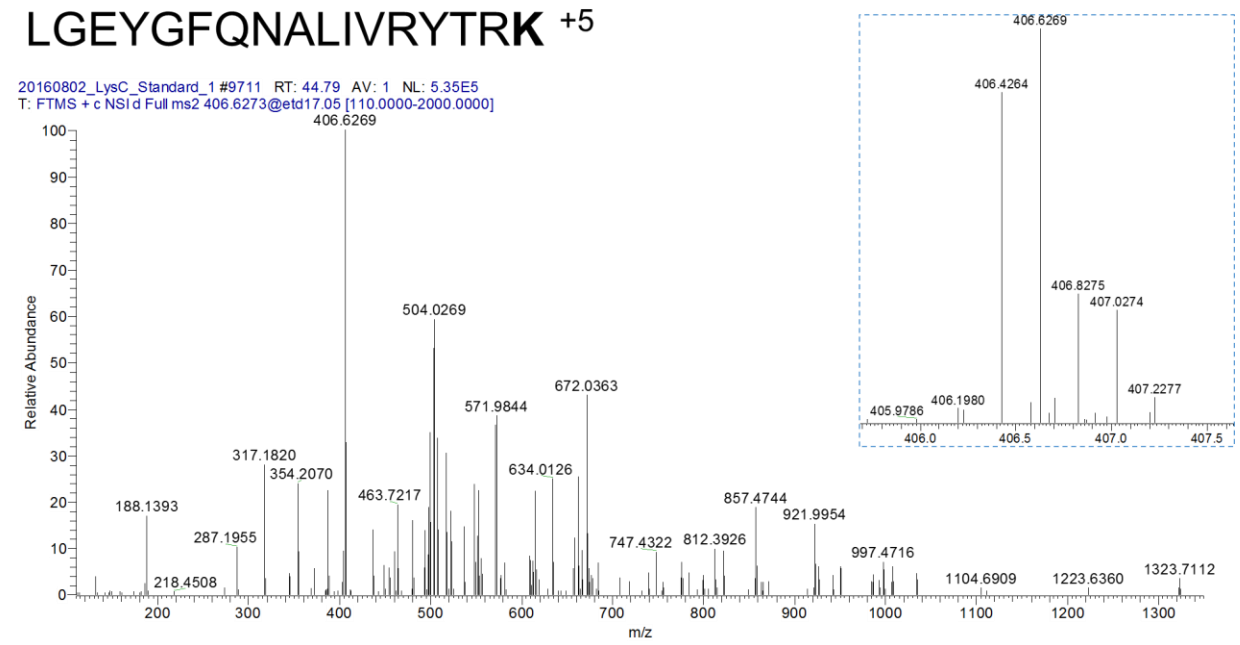


Figure 5.57. ETD-MS2 of Underivatized LysC-Digested LGEYGFQNALIVRYTRK Detected in Standard by Database Search.

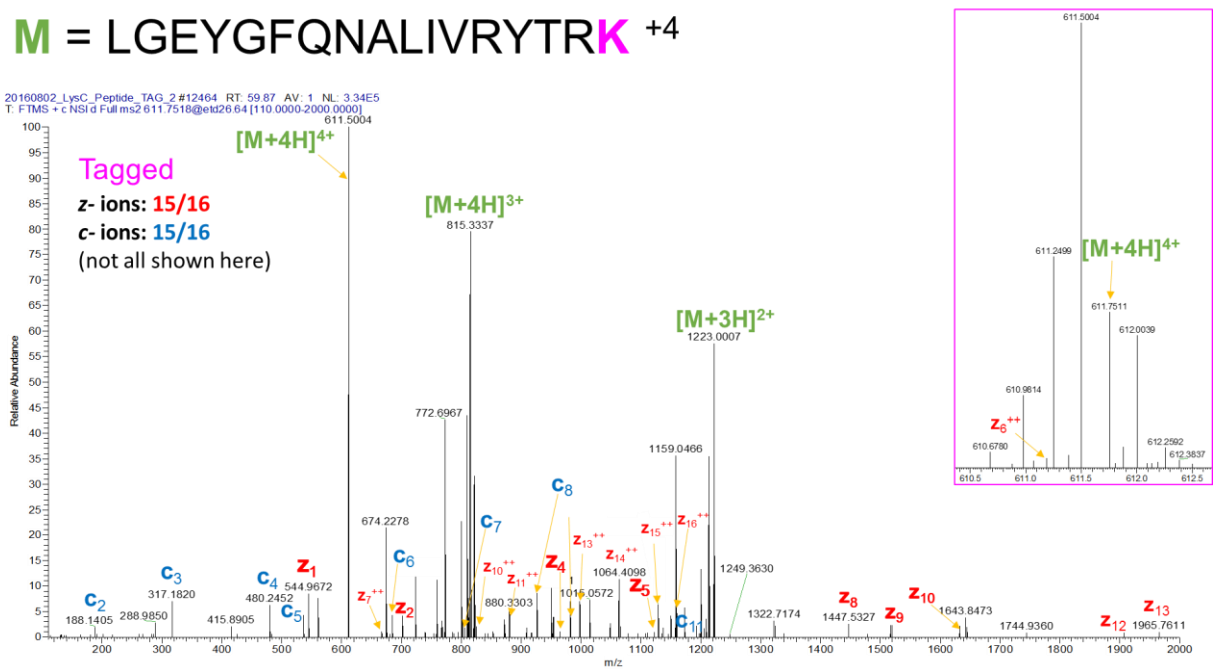


Figure 5.58. ETD-MS2 of Derivatized LysC-Digested LGEYGFQNALIVRYTRK Detected in Tagged Sample by Database Search.

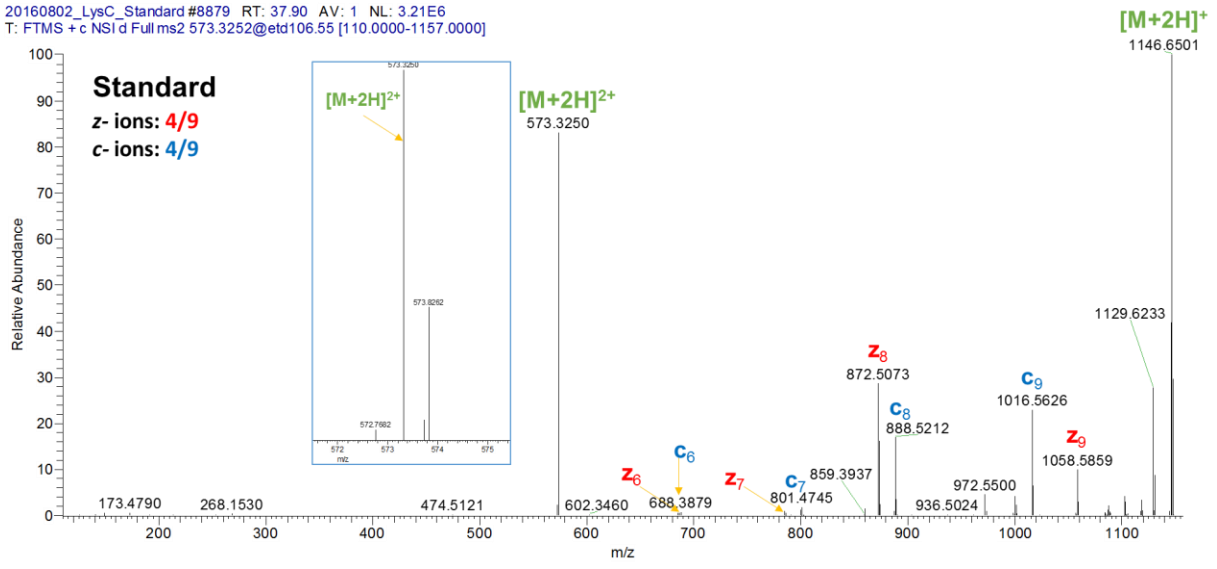


Figure 5.59. ETD-MS2 of Underivatized LysC-Digested AWSVARLSQK Detected in Standard by Database Search.

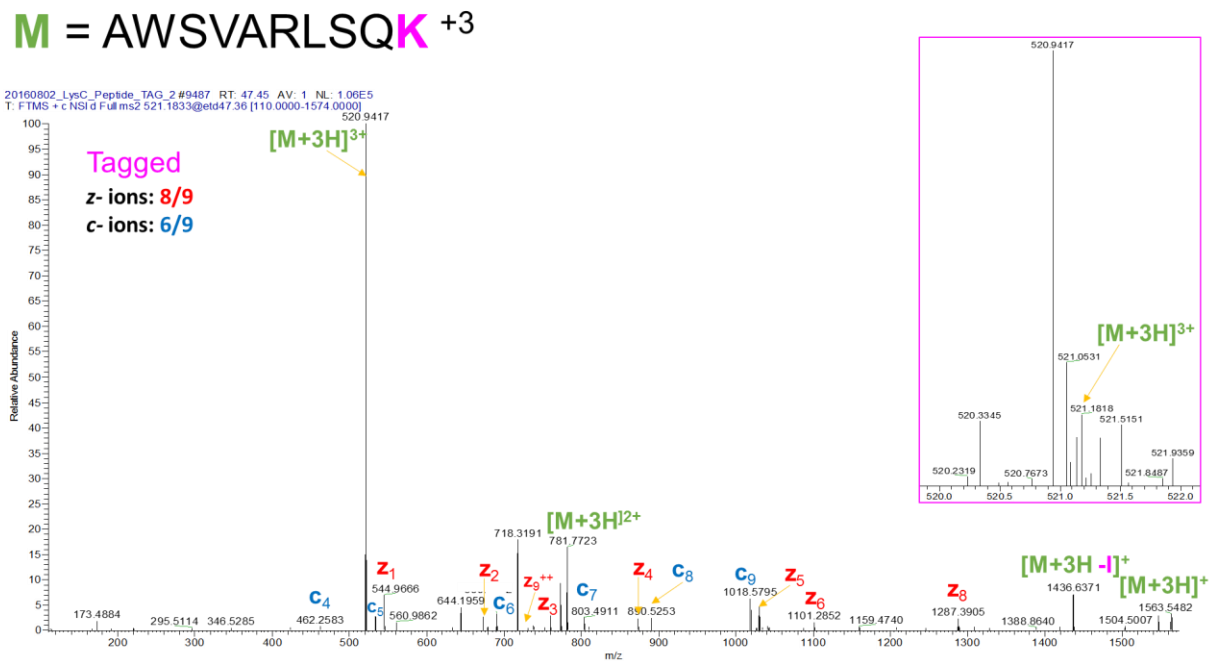


Figure 5.60. ETD-MS2 of Derivatized LysC-Digested AWSVARLSQK Detected in Tagged Sample by Database Search.

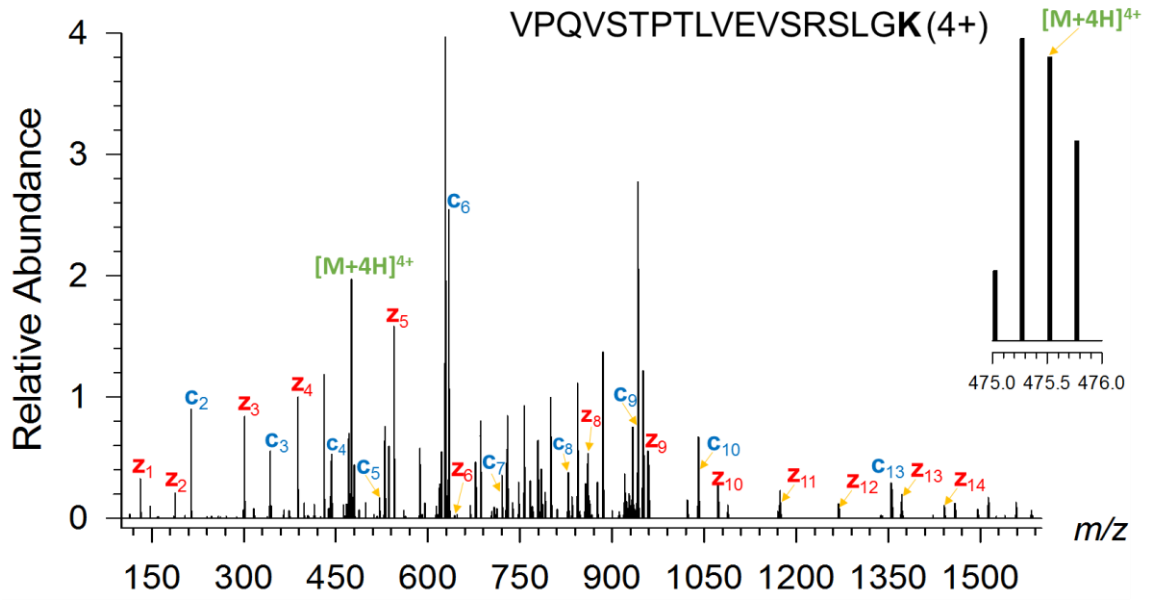


Figure 5.61. ETD-MS2 of Underivatized LysC-Digested VPQVSTPTLVEVSRSLGK Detected in Standard by Database Search.

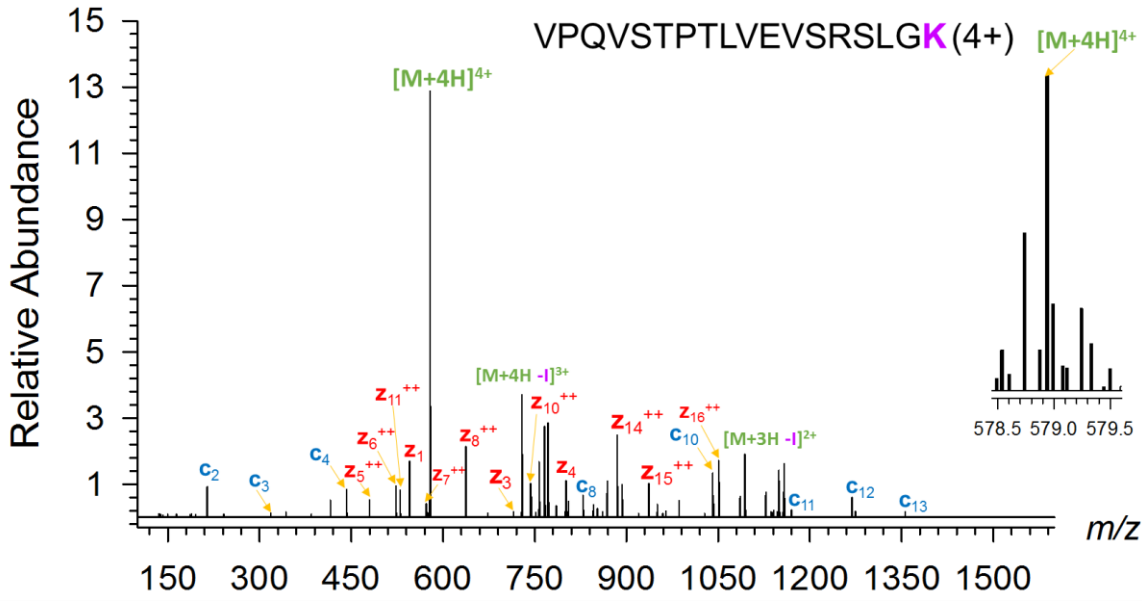


Figure 5.62. ETD-MS2 of Derivatized LysC-Digested VPQVSTPTLVEVSRSLGK Detected in Tagged Sample by Database Search.

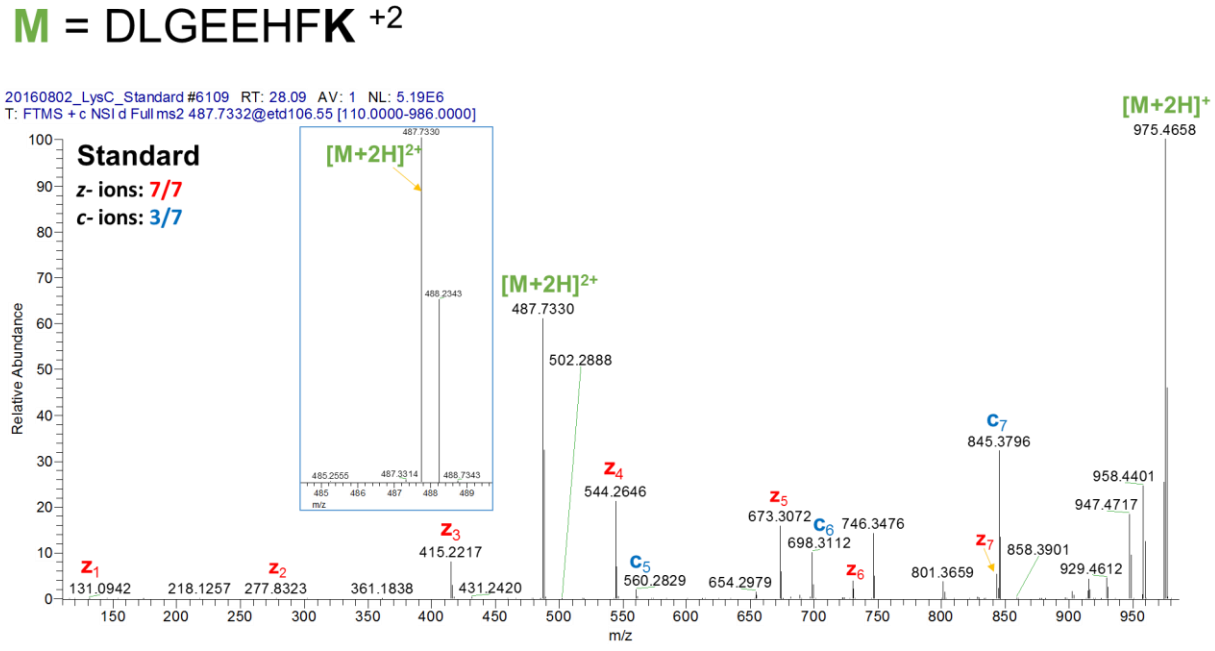


Figure 5.63. ETD-MS2 of Underivatized LysC-Digested DLGEEHFK Detected in Standard by Database Search.

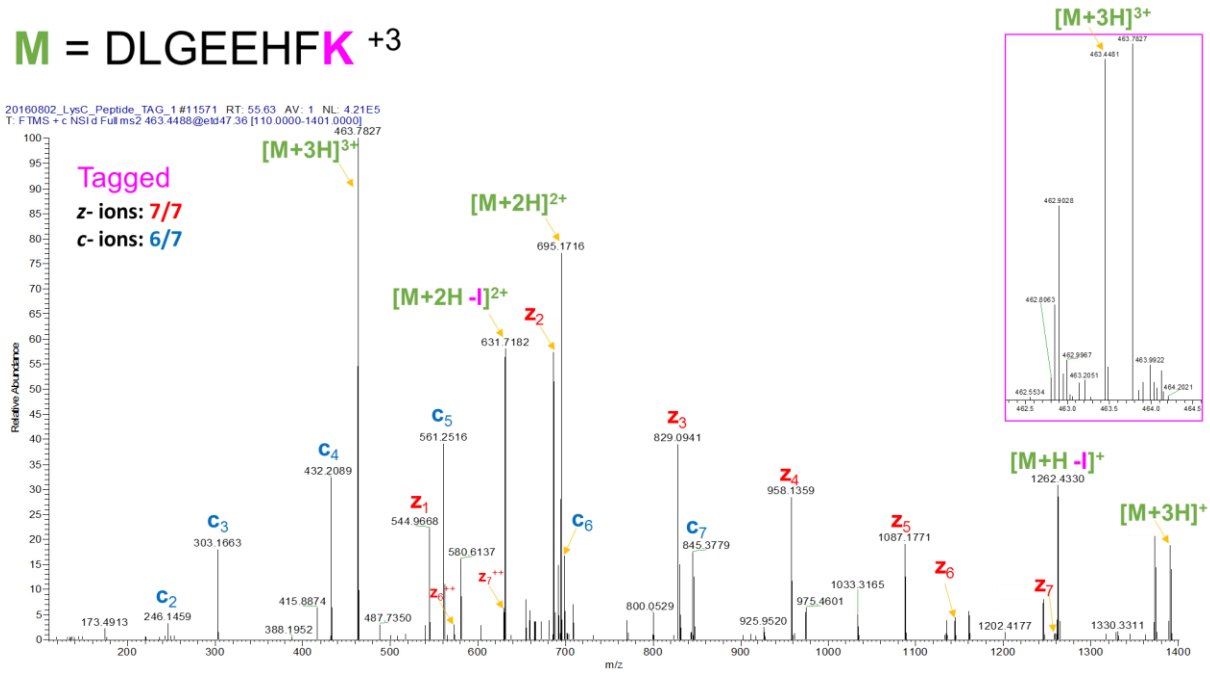


Figure 5.64. ETD-MS2 of Derivatized LysC-Digested DLGEEHFK Detected in Tagged Sample by Database Search.

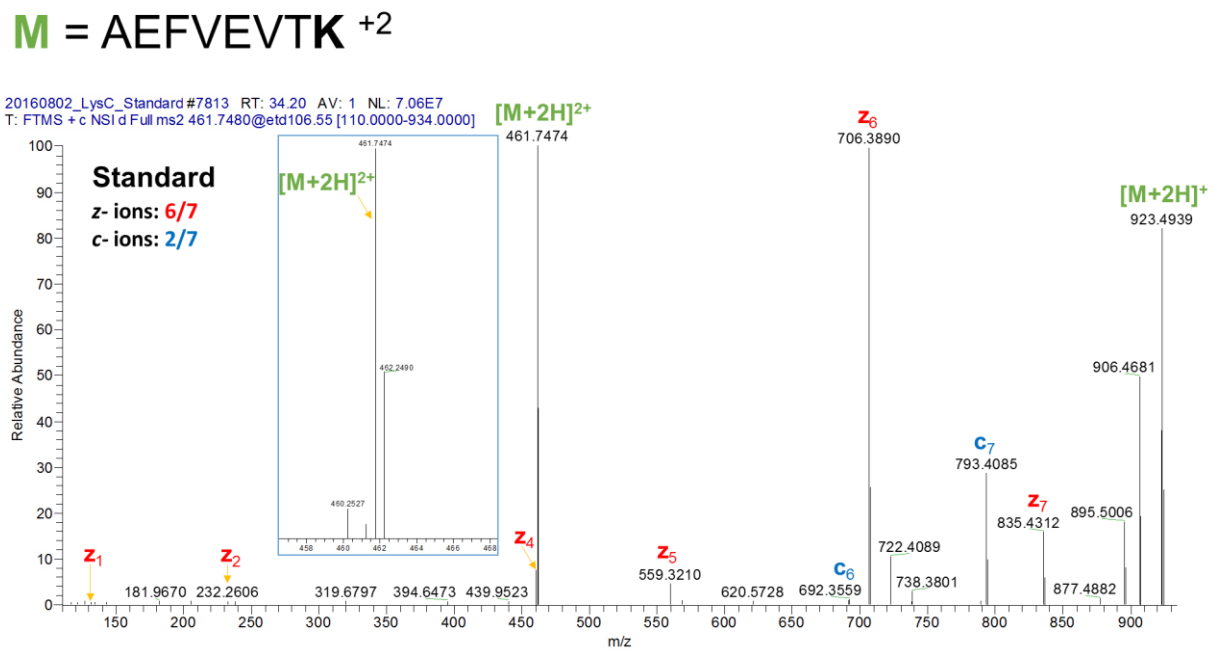


Figure 5.65. ETD-MS2 of Underivatized LysC-Digested AEFVEVTK Detected in Standard by Database Search.

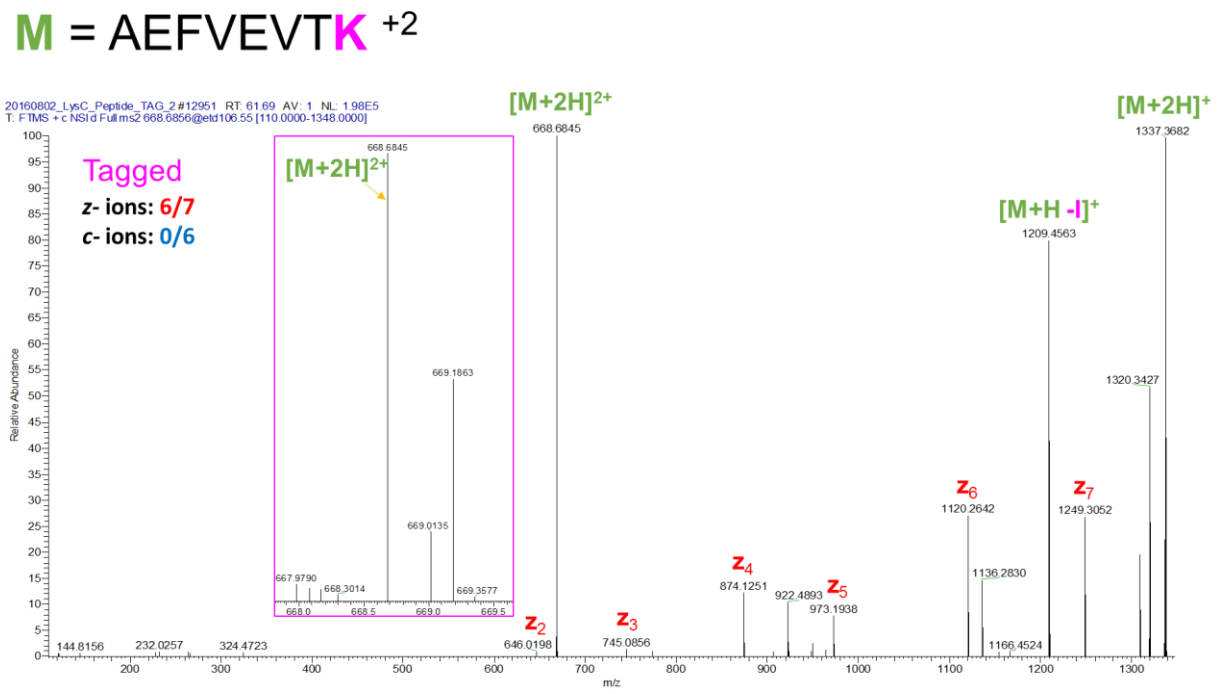


Figure 5.66. ETD-MS2 of Derivatized LysC-Digested AEFVEVTK Detected in Tagged Sample by Database Search.

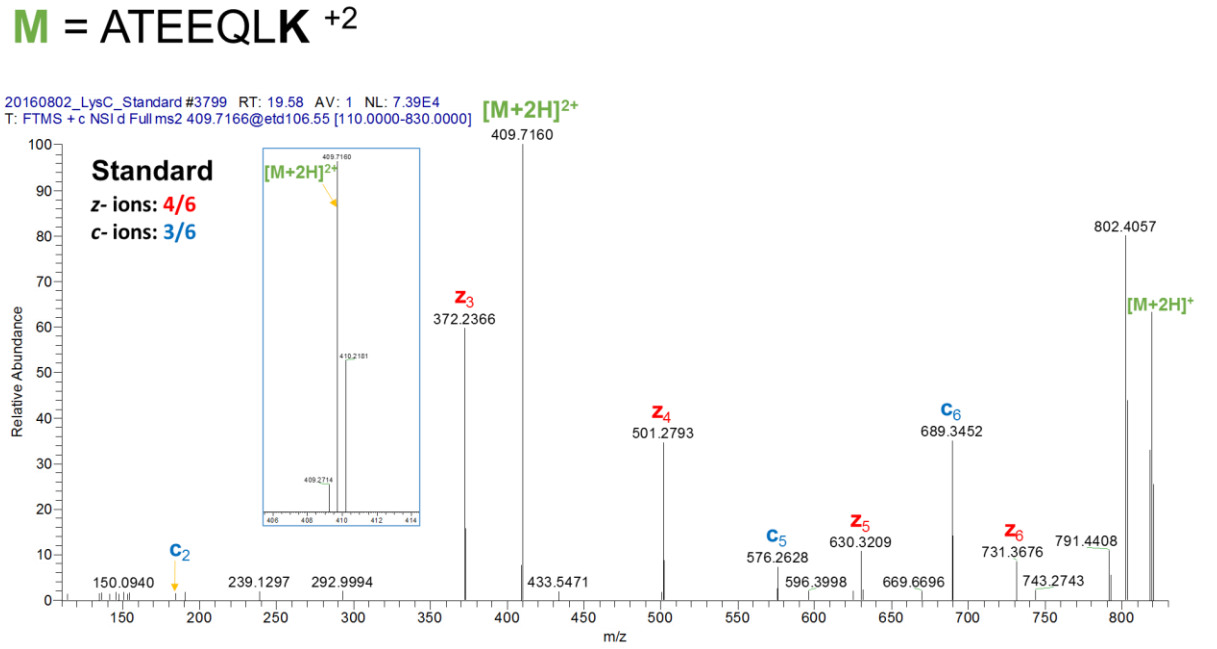


Figure 5.67. ETD-MS2 of Underivatized LysC-Digested ATEEQLK Detected in Standard by Database Search.

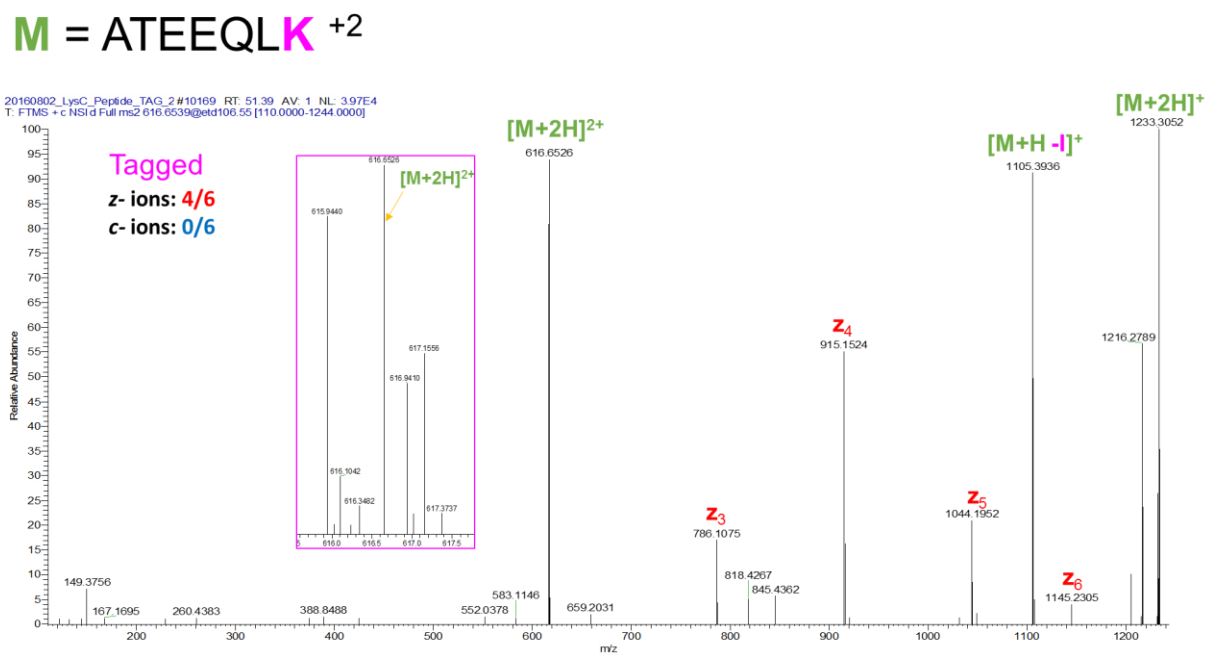


Figure 5.68. ETD-MS2 of Derivatized LysC-Digested ATEEQLK Detected in Tagged Sample by Database Search.

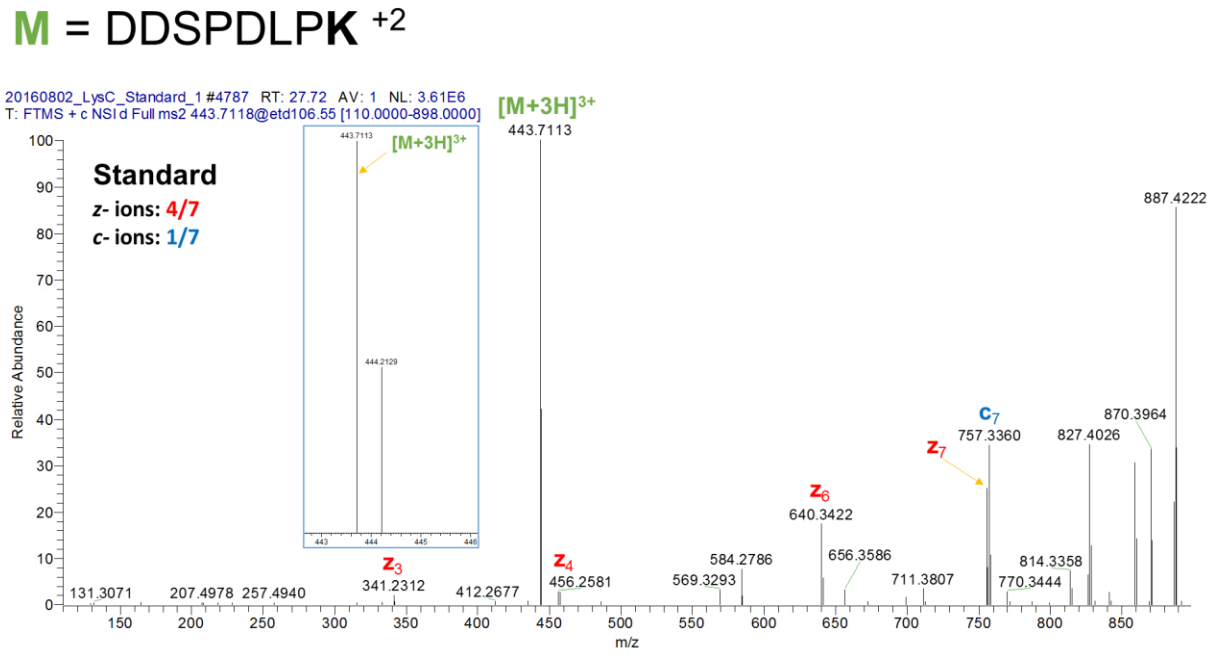


Figure 5.69. ETD-MS2 of Underivatized LysC-Digested DDSPDLPK Detected in Standard by Database Search.

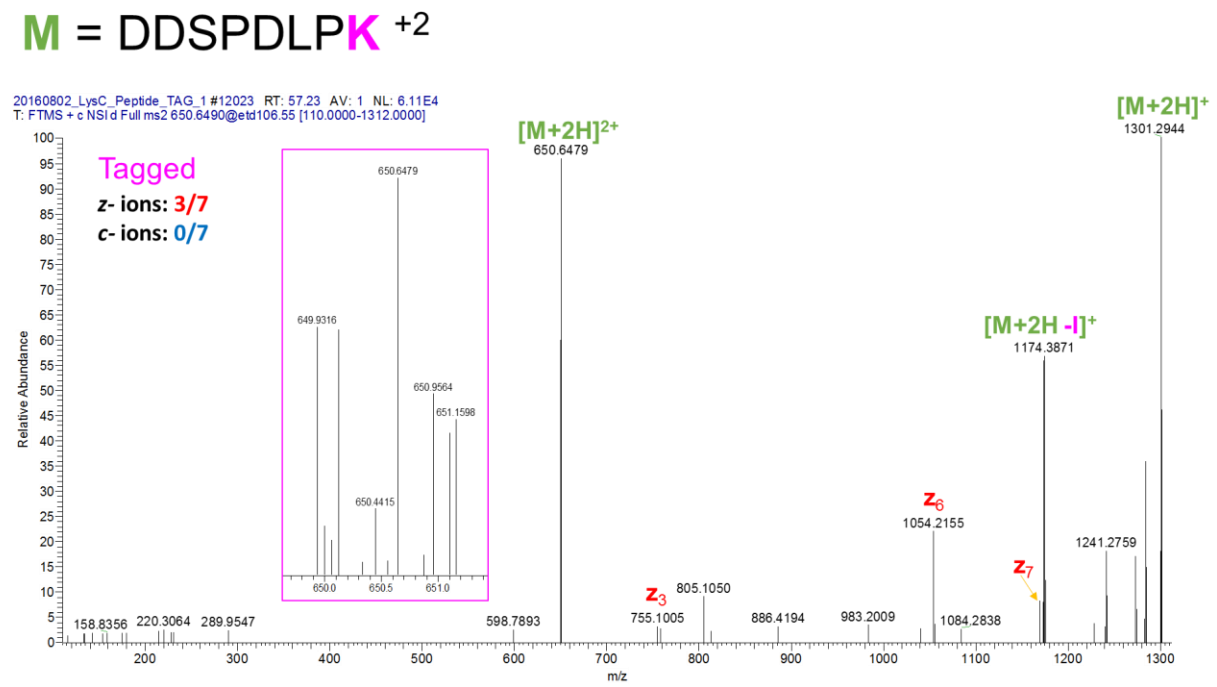


Figure 5.70. ETD-MS2 of Derivatized LysC-Digested DDSPDLPK Detected in Tagged Sample by Database Search.

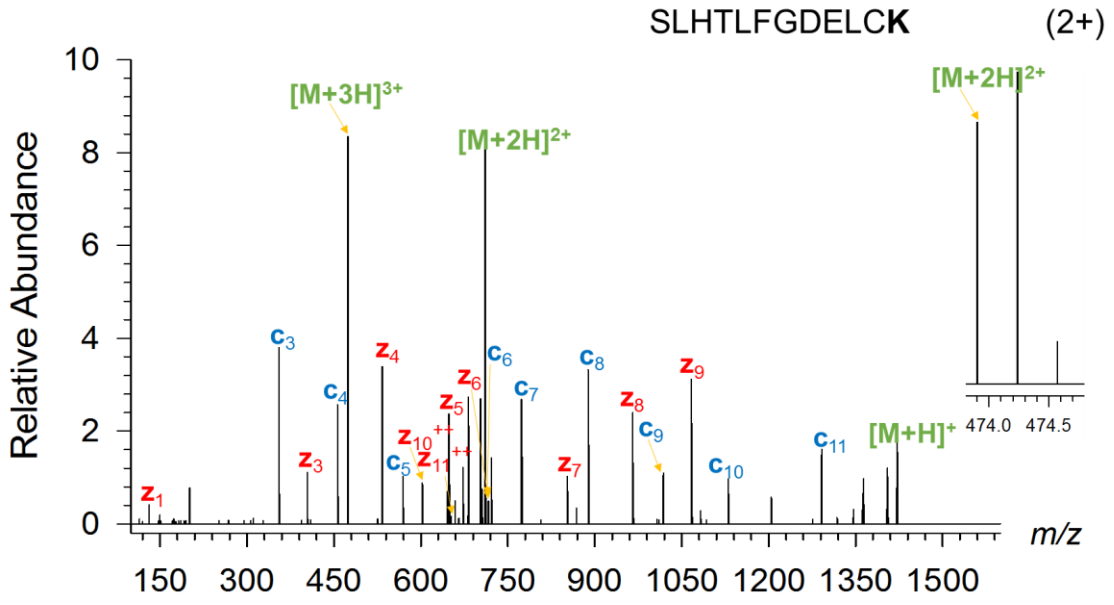


Figure 5.71. ETD-MS2 of Underivatized LysC-Digested SLHTLFGDELCK Detected in Standard.

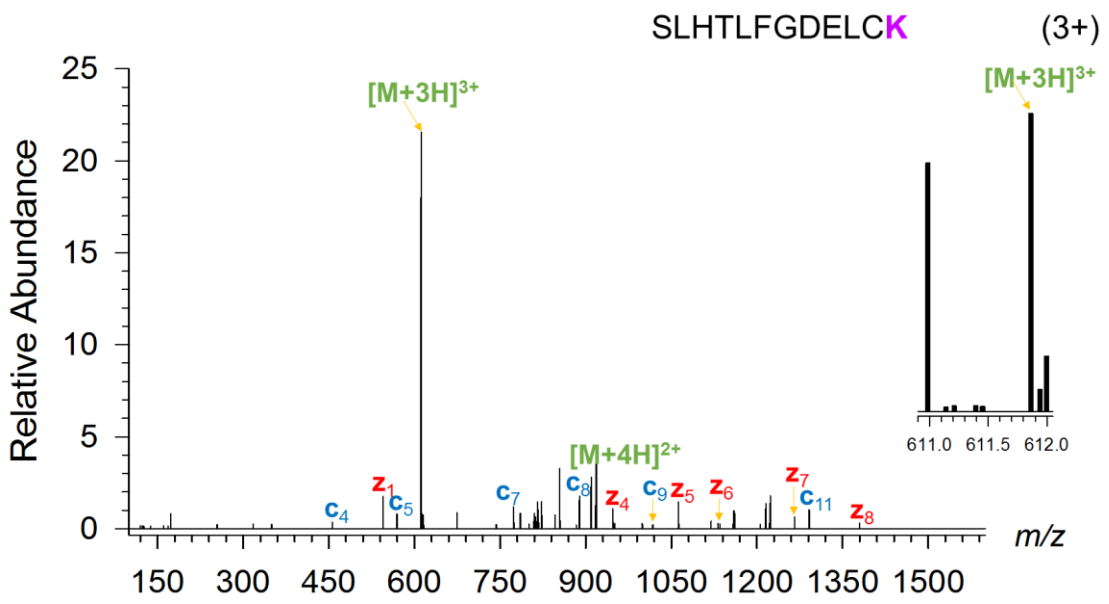


Figure 5.72. ETD-MS2 of Derivatized LysC-Digested SLHTLFGDELCK Detected in Tagged Sample.

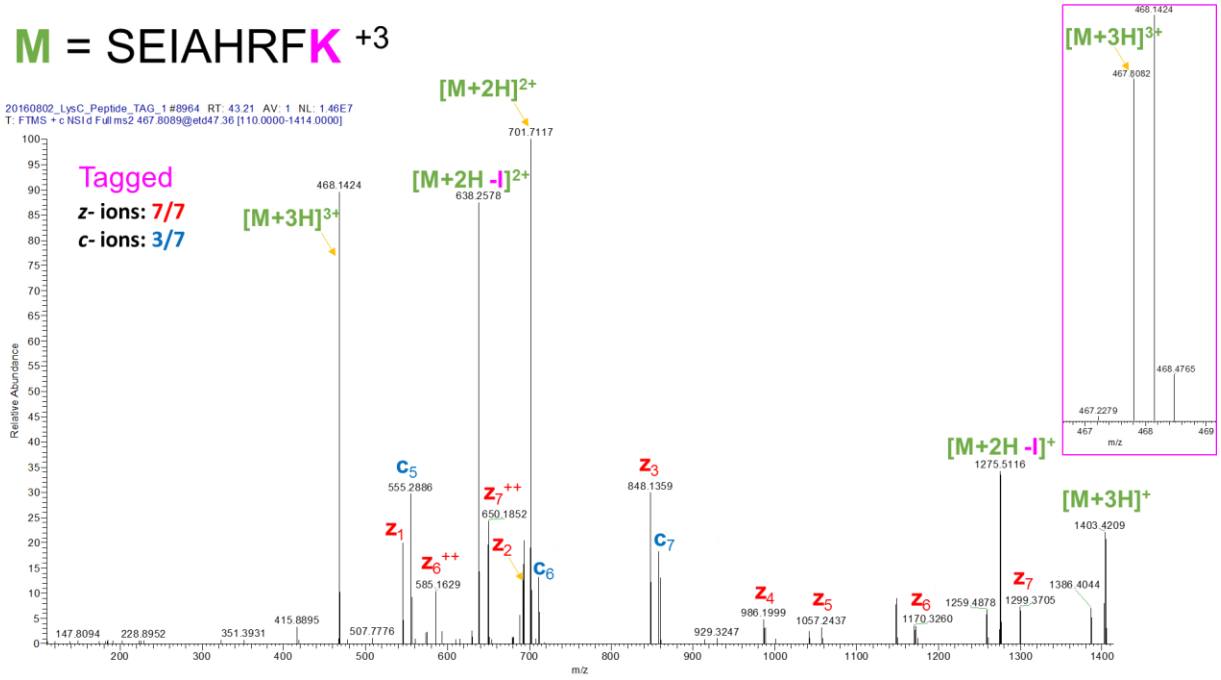


Figure 5.73. ETD-MS2 of Derivatized LysC-Digested SEIAHRFK Detected in Tagged Sample by Database Search. (Note that the untagged counterpart was not detected in the standard).

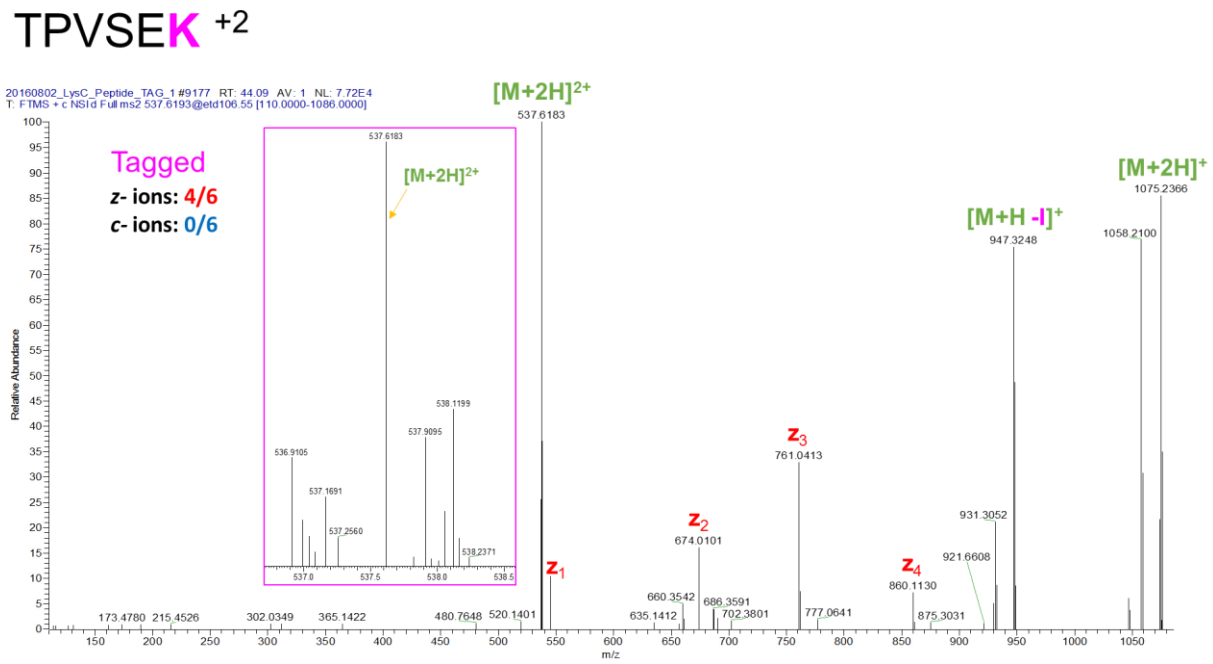


Figure 5.74. ETD-MS2 of Derivatized LysC-Digested TPVSEK Detected in Tagged Sample by Database Search. (Note that the untagged counterpart was not detected in the standard)

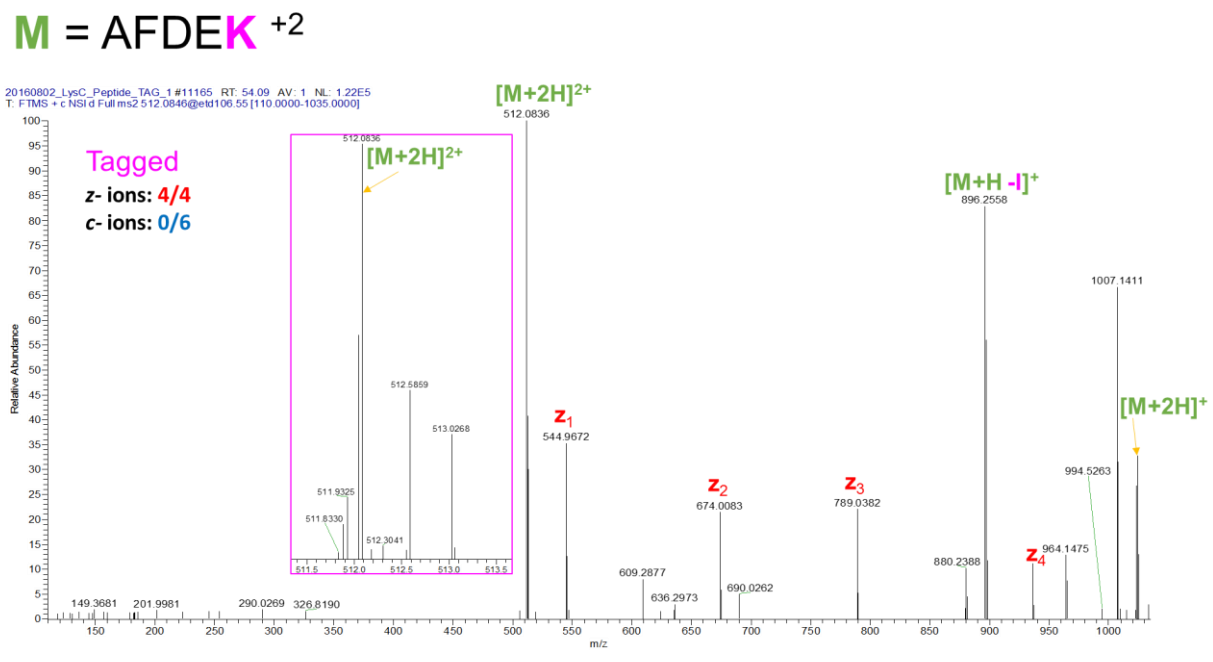


Figure 5.75. ETD-MS2 of Derivatized LysC-Digested AFDEK Detected in Tagged Sample by Database Search. (Note that the untagged counterpart was not detected in the standard)

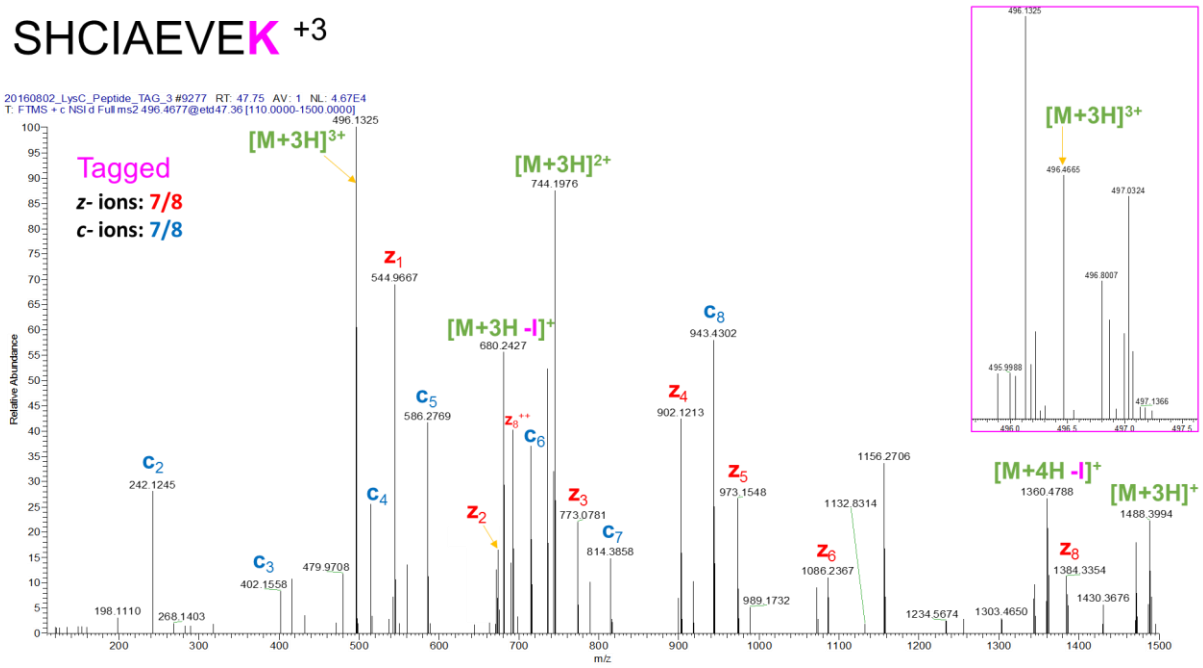


Figure 5.76. ETD-MS2 of Derivatized LysC-Digested SHCIAEVEK Detected in Tagged Sample by Database Search. (Note that the untagged counterpart was not detected in the standard).

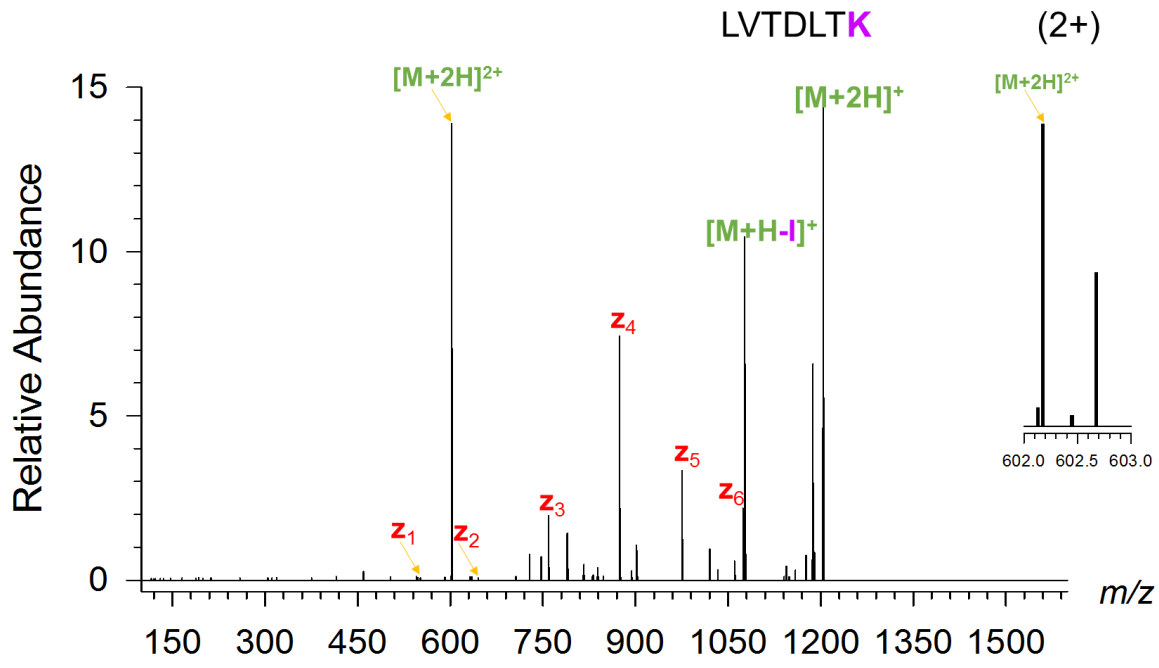


Figure 5.77. ETD-MS2 of Derivatized LysC-Digested LVTDLTK Detected in Tagged Sample. (Note that the untagged counterpart was not detected in the standard)

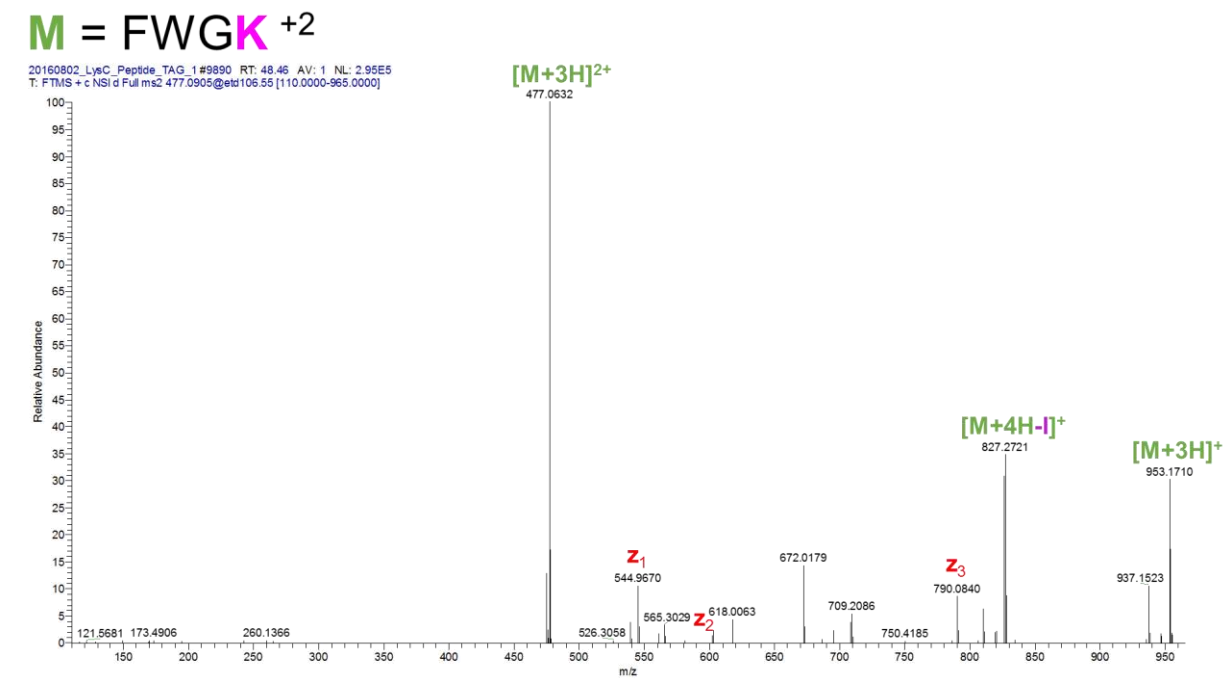


Figure 5.78. ETD-MS2 of Tagged LysC-Digested FWGK Detected in Tagged Sample.

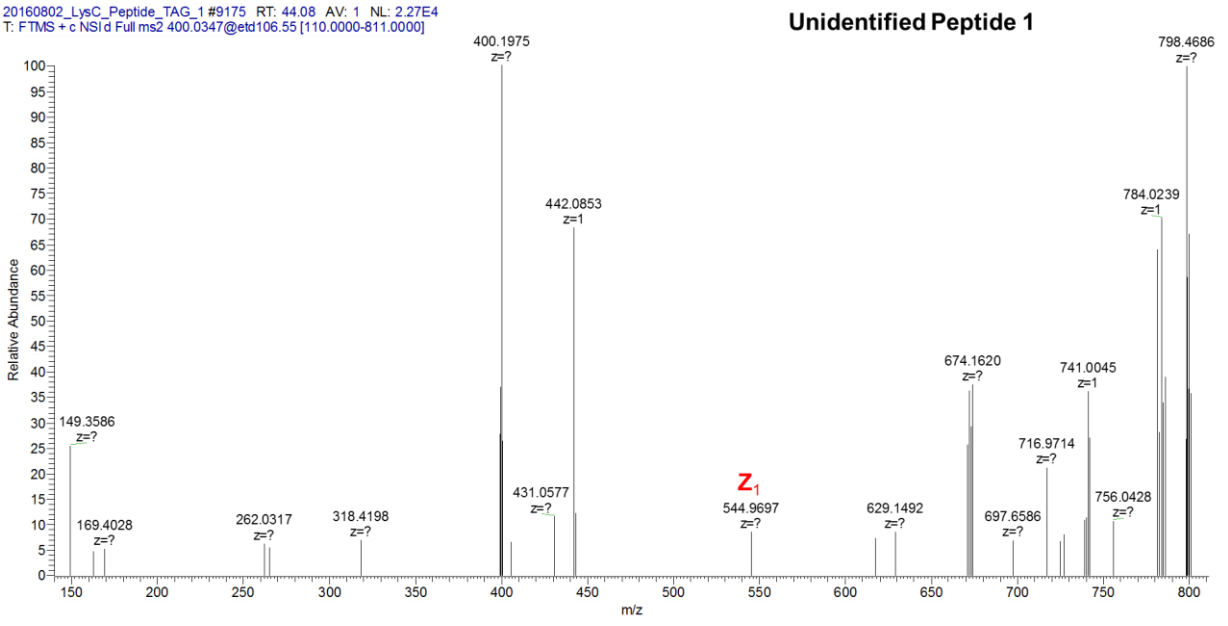


Figure 5.79. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 1, Undetected by Database but Detected Using In-House Developed Script.

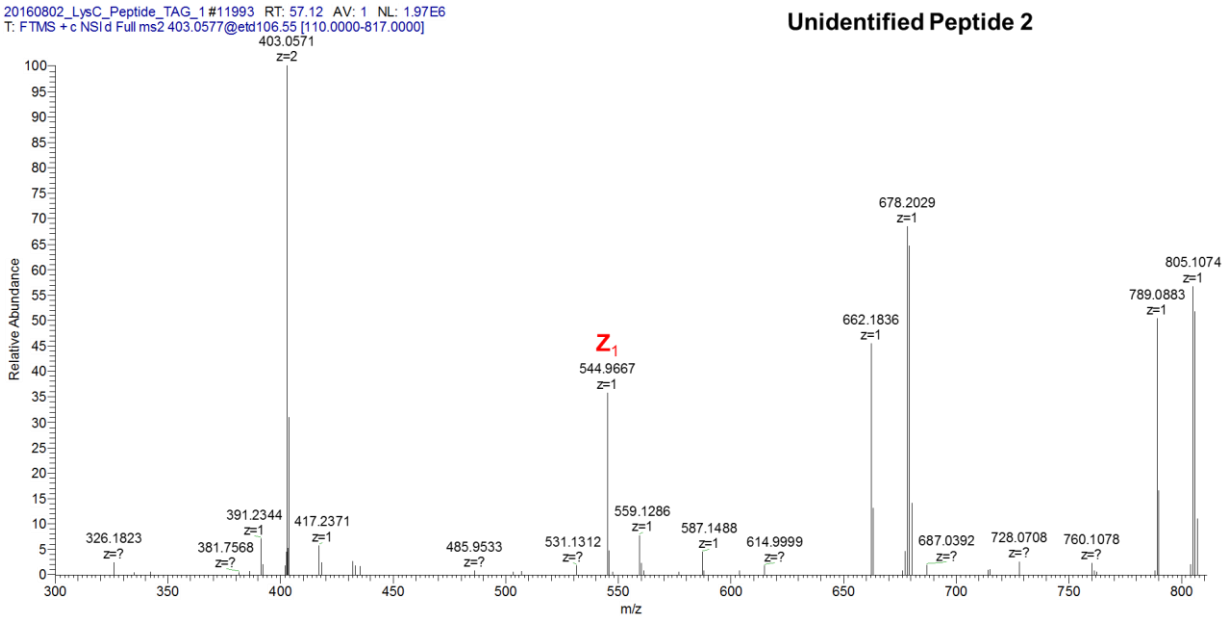


Figure 5.80. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 2, Undetected by Database but Detected Using In-House Developed Script.

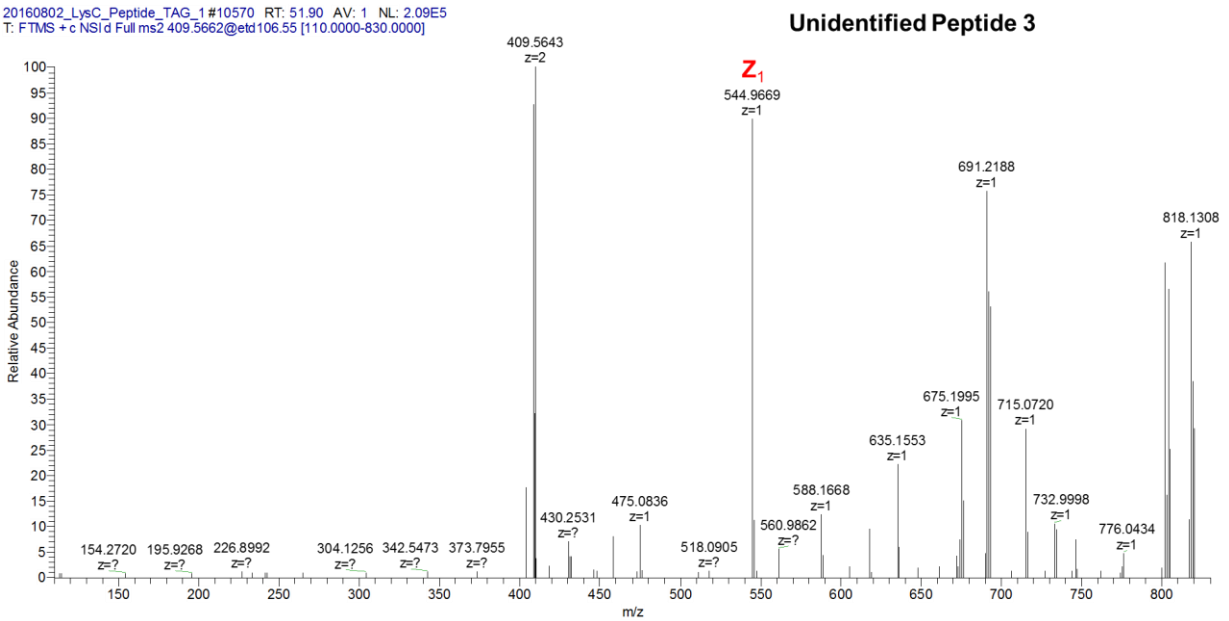


Figure 5.81. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 3, Undetected by Database but Detected Using In-House Developed Script.

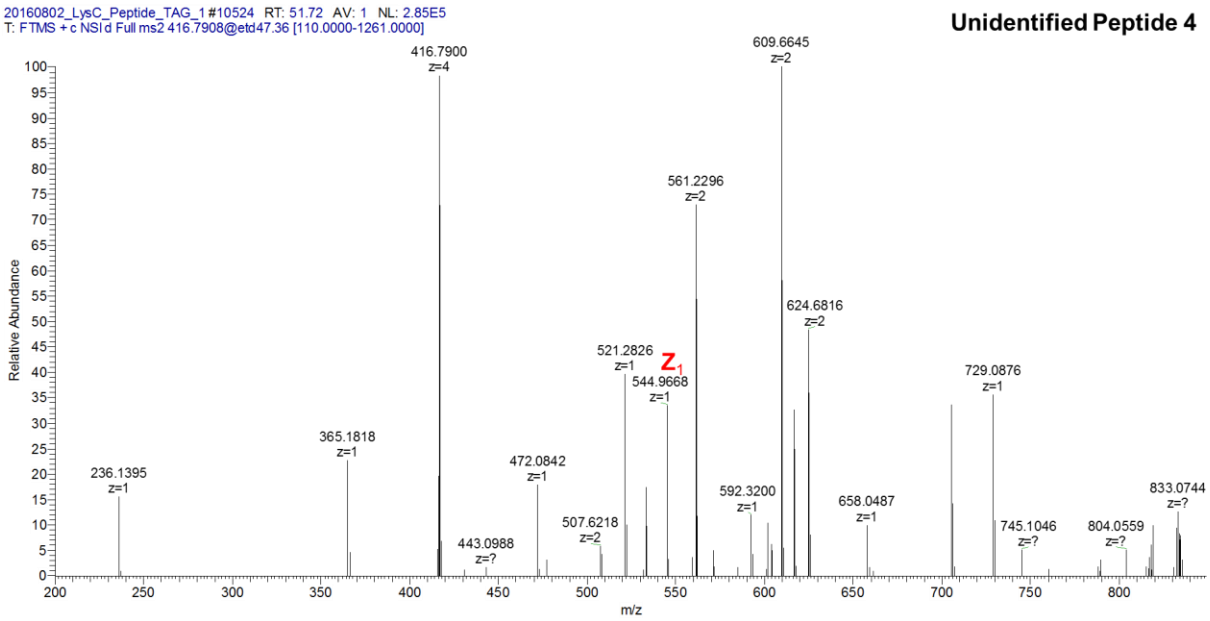


Figure 5.82. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 4, Undetected by Database but Detected Using In-House Developed Script.

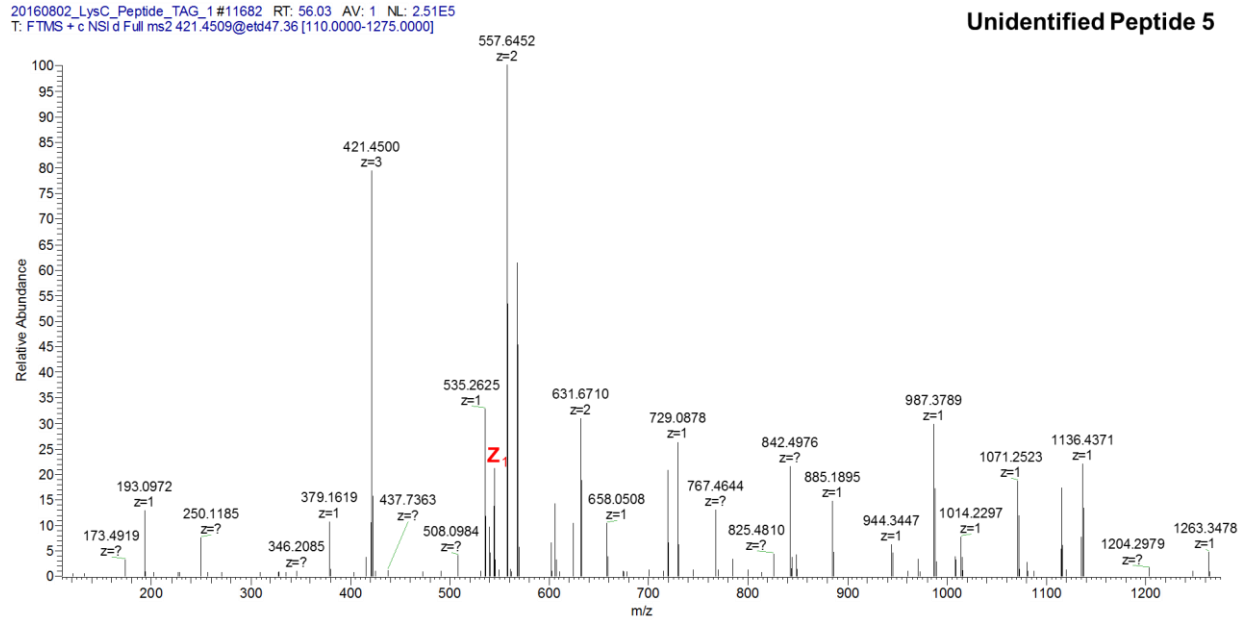


Figure 5.83. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 5, Undetected by Database but Detected Using In-House Developed Script.

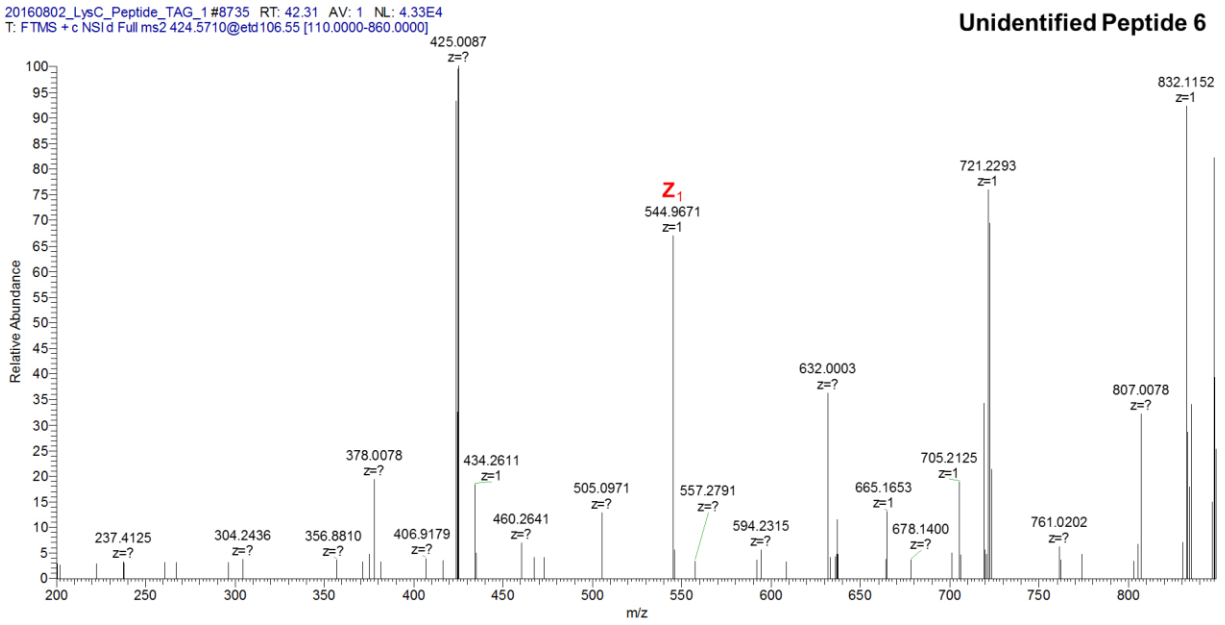


Figure 5.84. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 6, Undetected by Database but Detected Using In-House Developed Script.

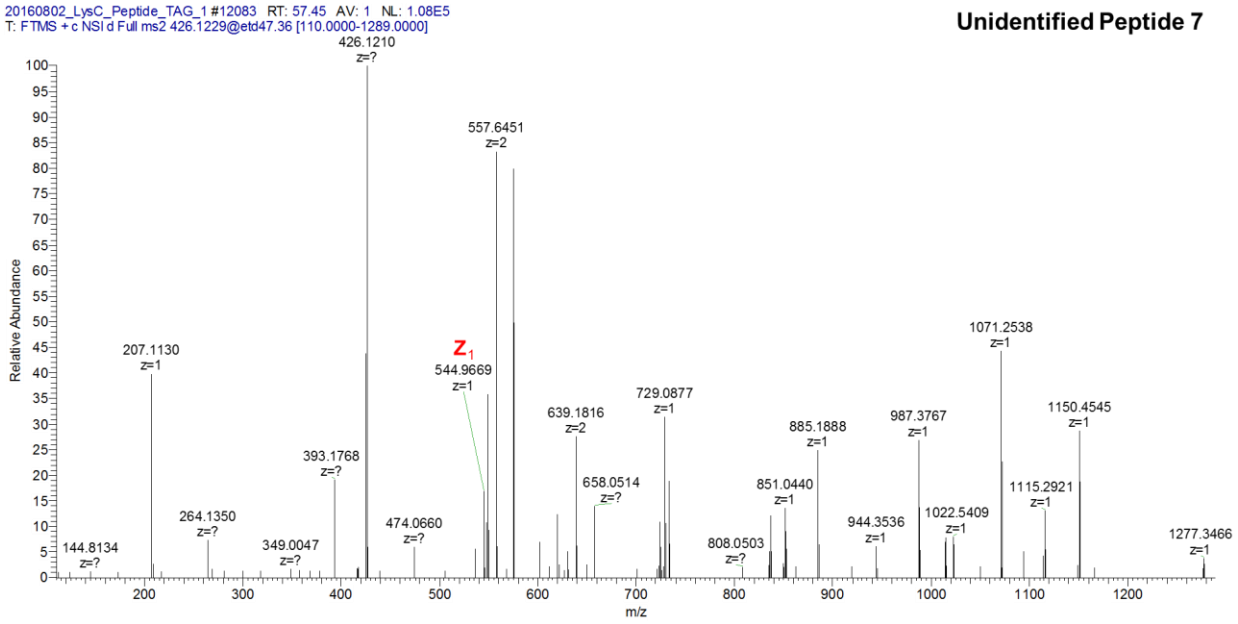


Figure 5.85. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 7, Undetected by Database but Detected Using In-House Developed Script.

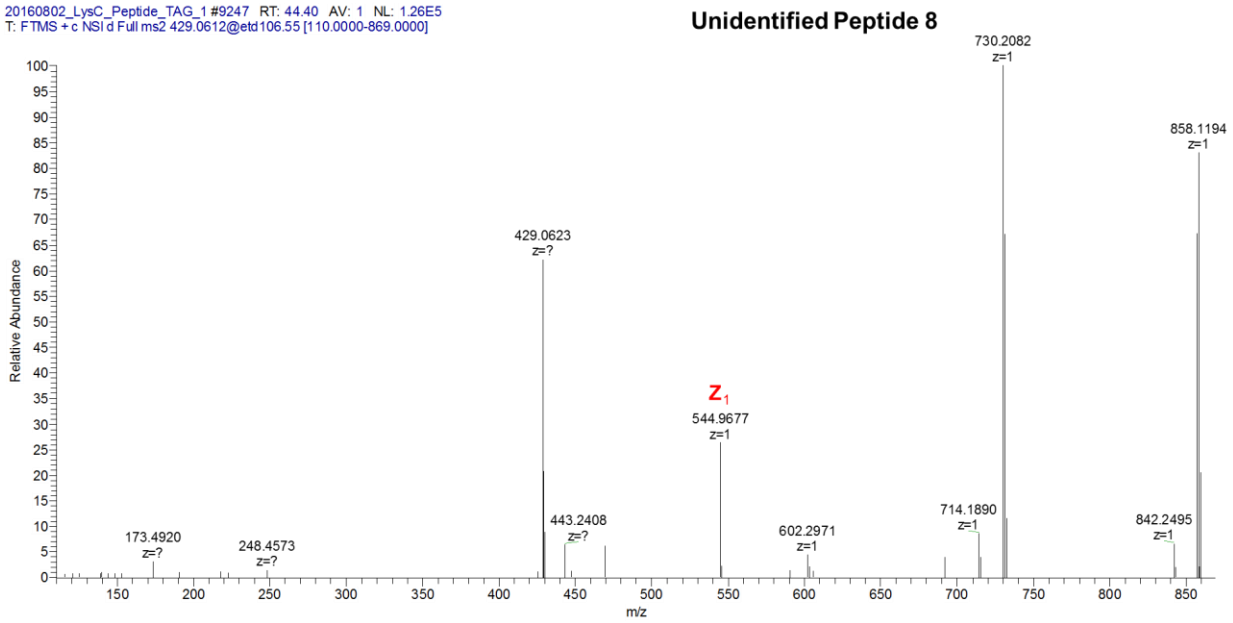


Figure 5.86. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 8, Undetected by Database but Detected Using In-House Developed Script.

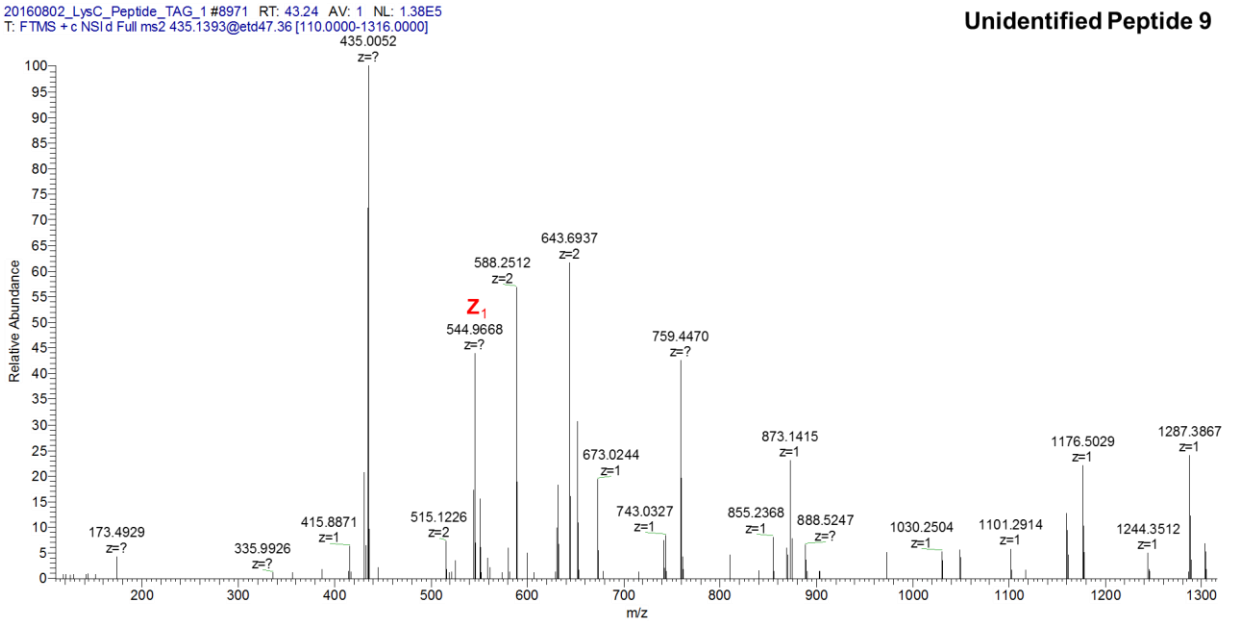


Figure 5.87. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 9, Undetected by Database but Detected Using In-House Developed Script.

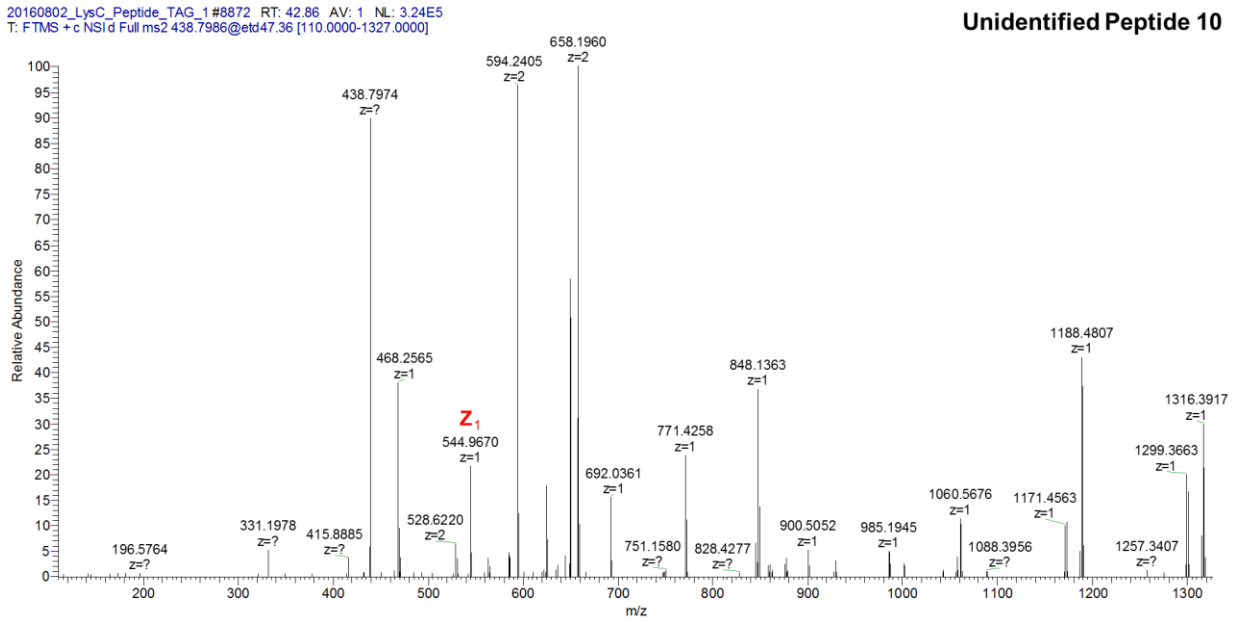


Figure 5.88. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 10, Undetected by Database but Detected Using In-House Developed Script.

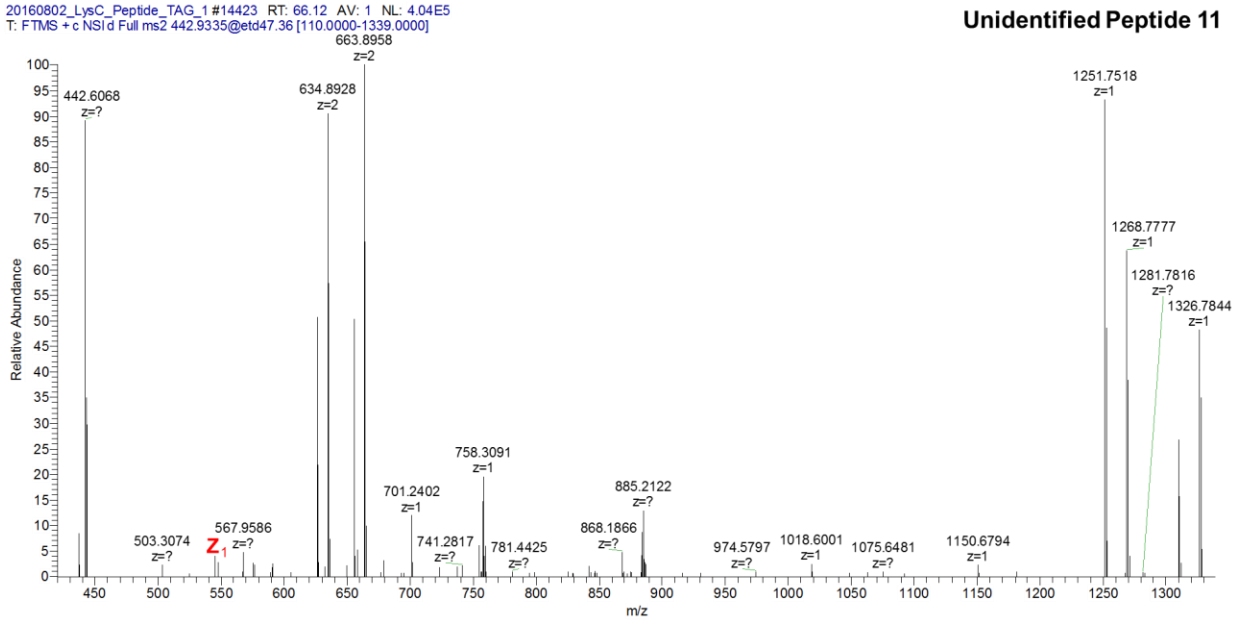


Figure 5.89. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 11, Undetected by Database but Detected Using In-House Developed Script.

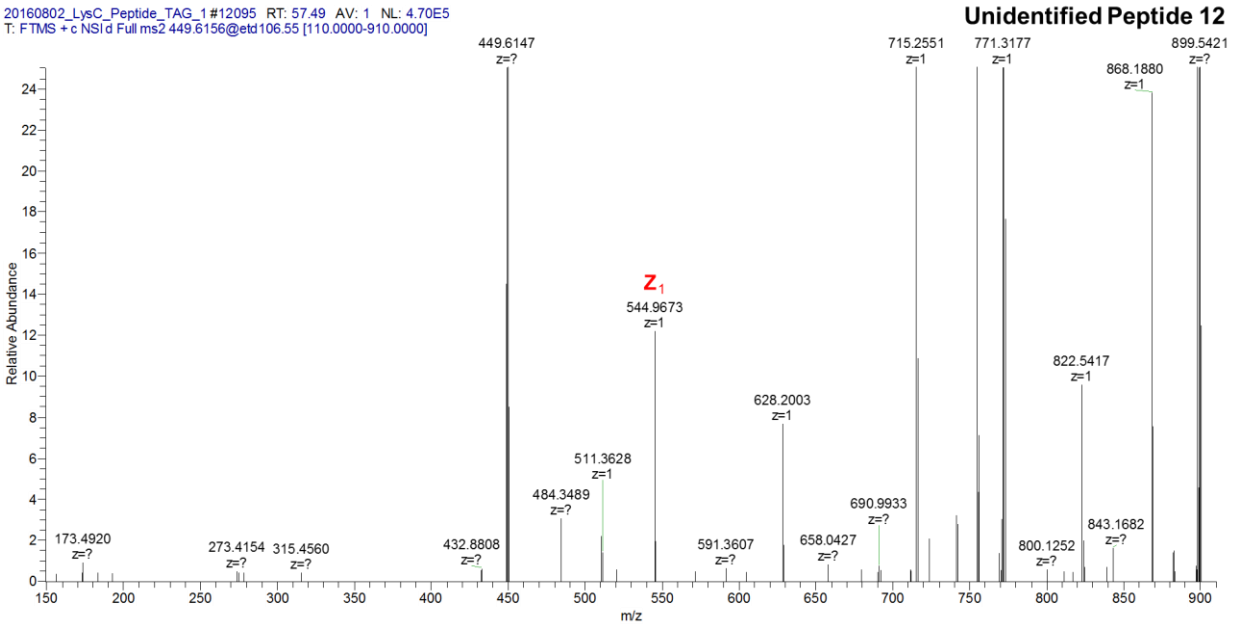


Figure 5.90. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 12, Undetected by Database but Detected Using In-House Developed Script.

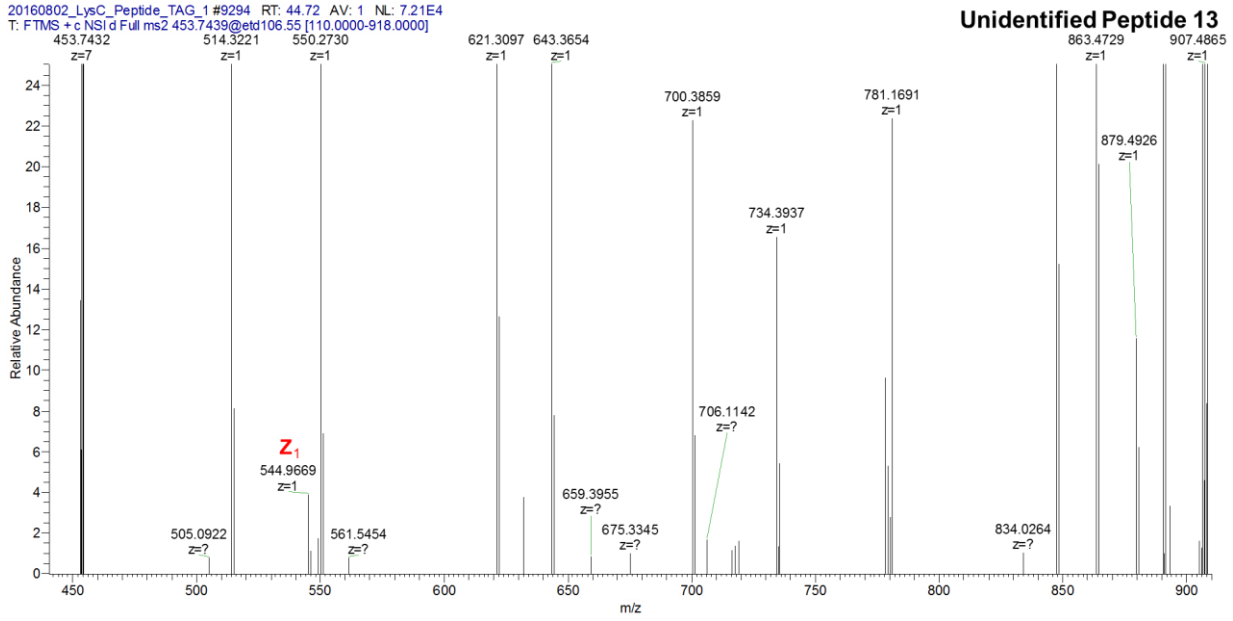


Figure 5.91. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 13, Undetected by Database but Detected Using In-House Developed Script.

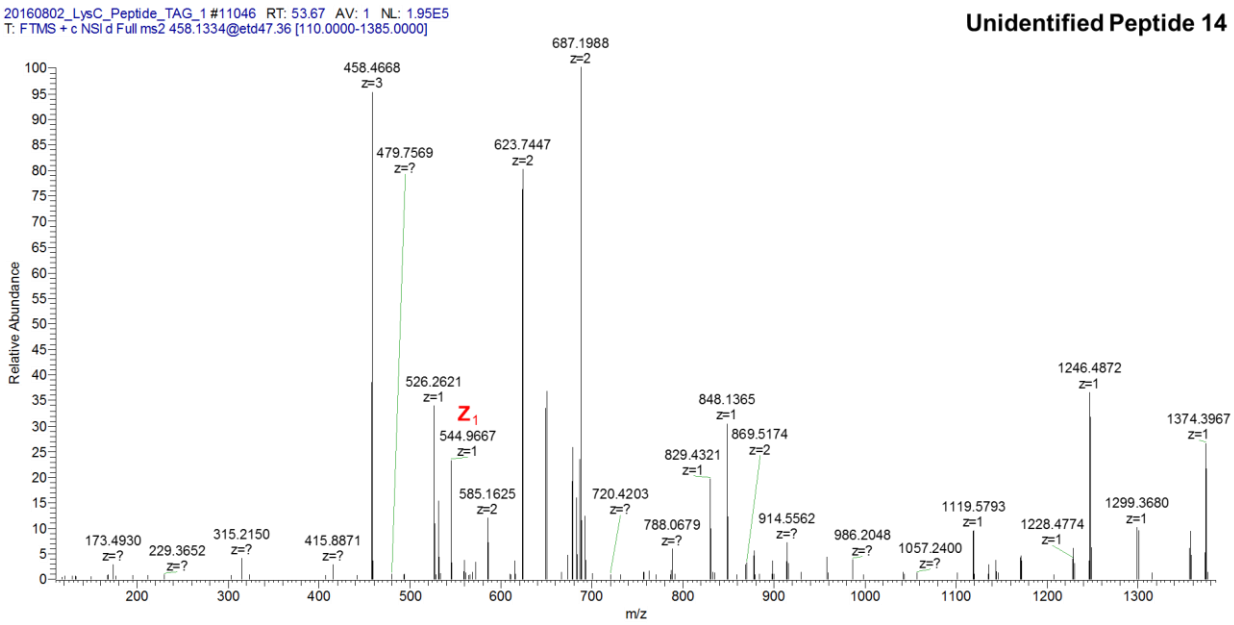


Figure 5.92. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 14, Undetected by Database but Detected Using In-House Developed Script.

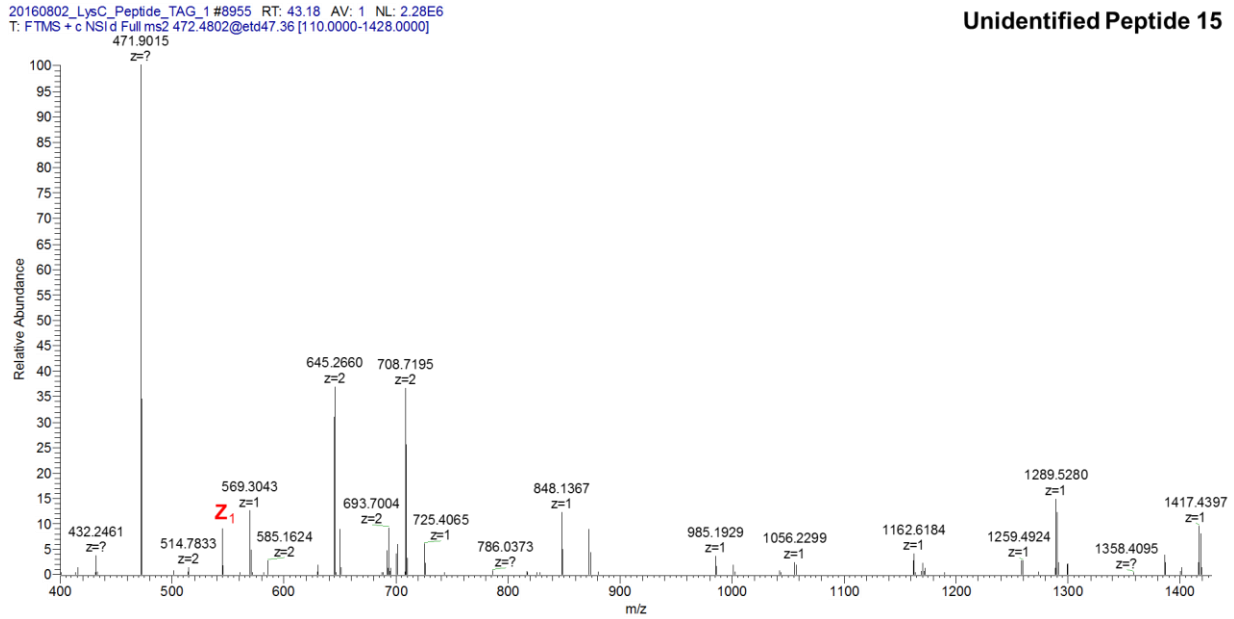


Figure 5.93. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 15, Undetected by Database but Detected Using In-House Developed Script.

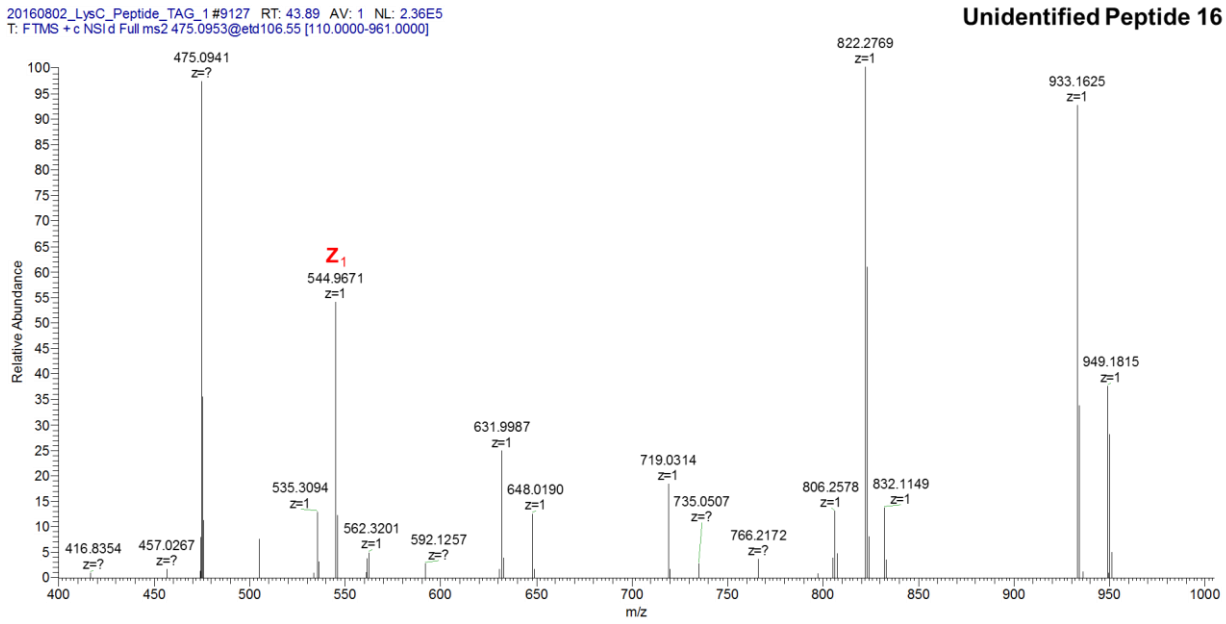


Figure 5.94. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 16, Undetected by Database but Detected Using In-House Developed Script.

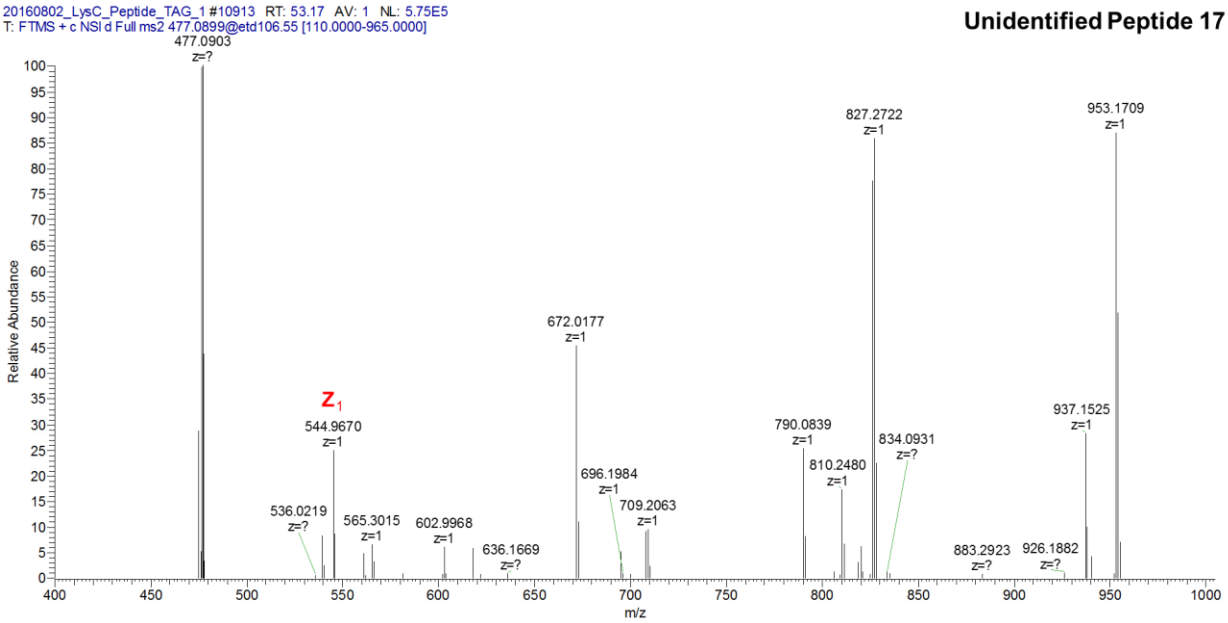


Figure 5.95. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 17, Undetected by Database but Detected Using In-House Developed Script.

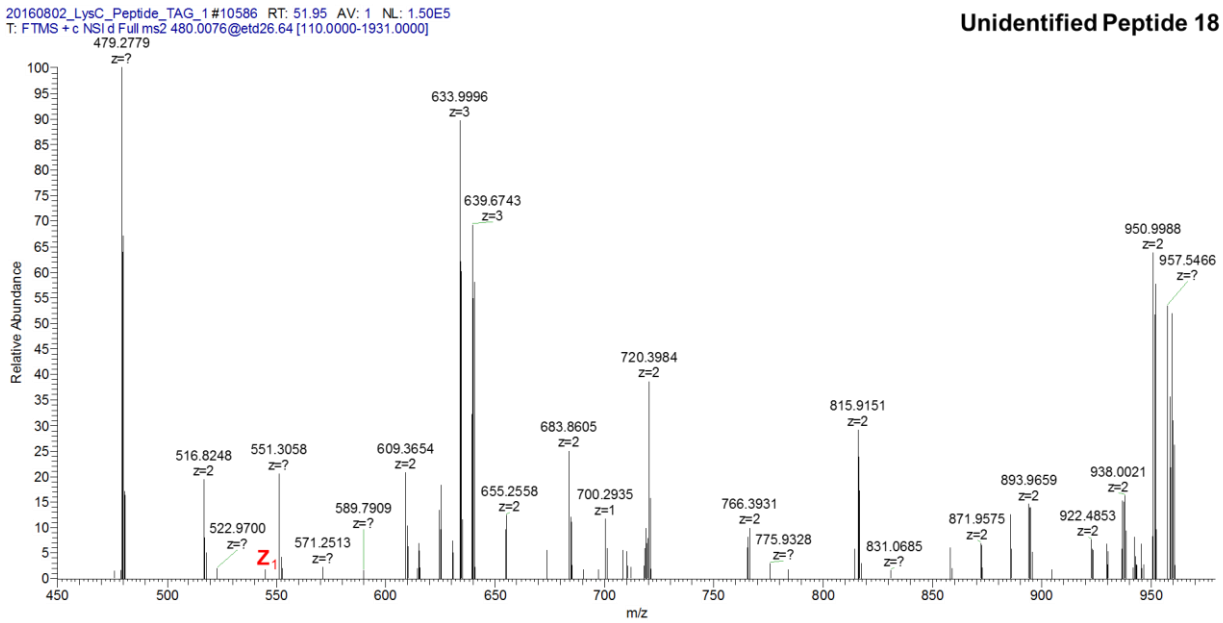


Figure 5.96. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 18, Undetected by Database but Detected Using In-House Developed Script.

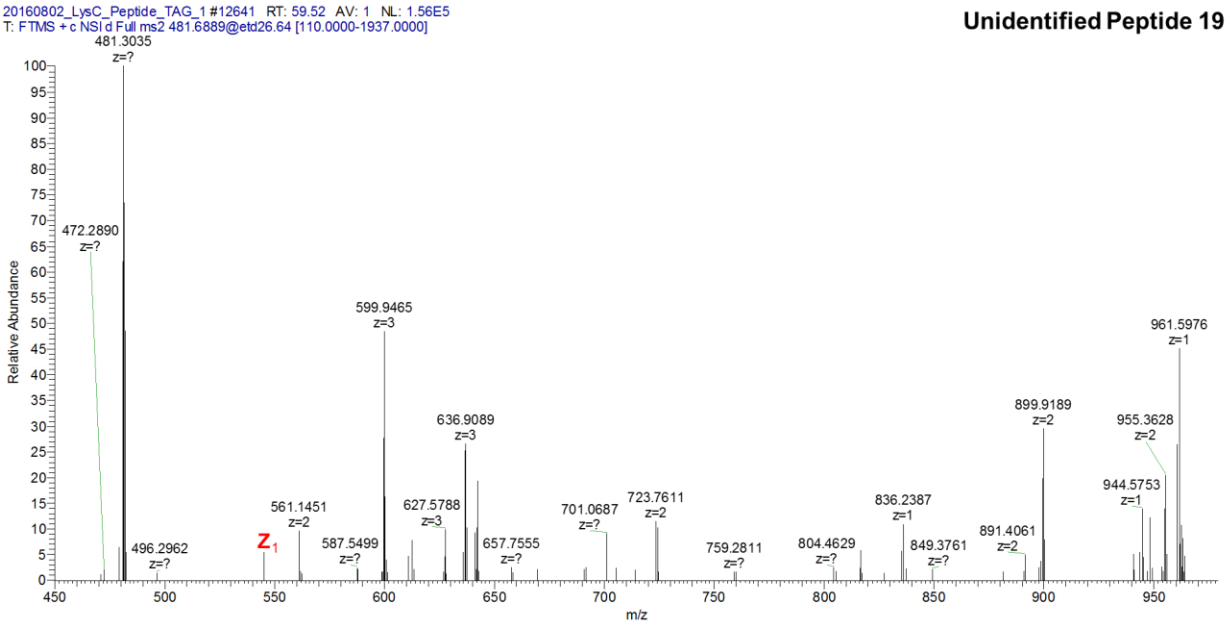


Figure 5.97. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 19, Undetected by Database but Detected Using In-House Developed Script.

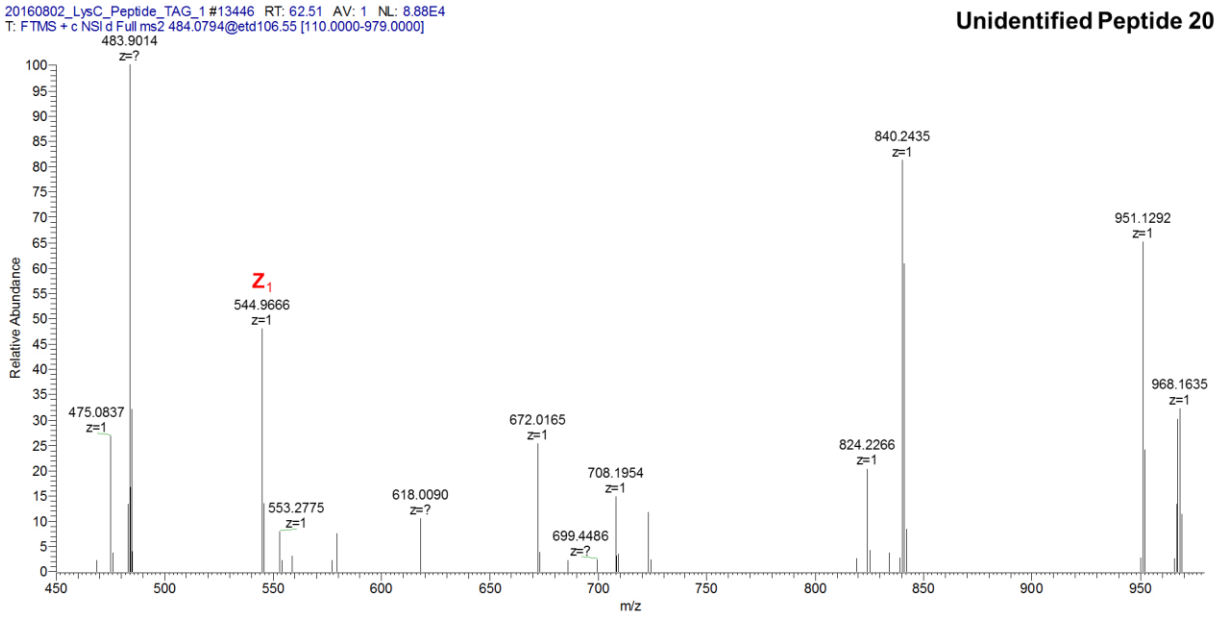


Figure 5.98. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 20, Undetected by Database but Detected Using In-House Developed Script.

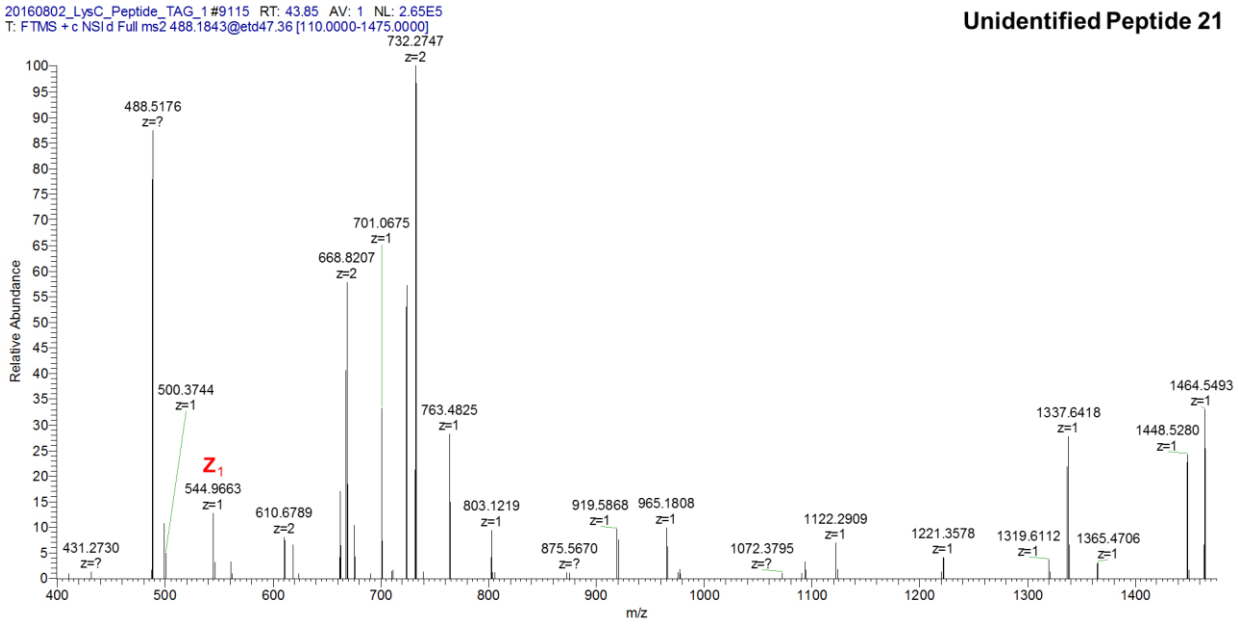


Figure 5.99. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 21, Undetected by Database but Detected Using In-House Developed Script.

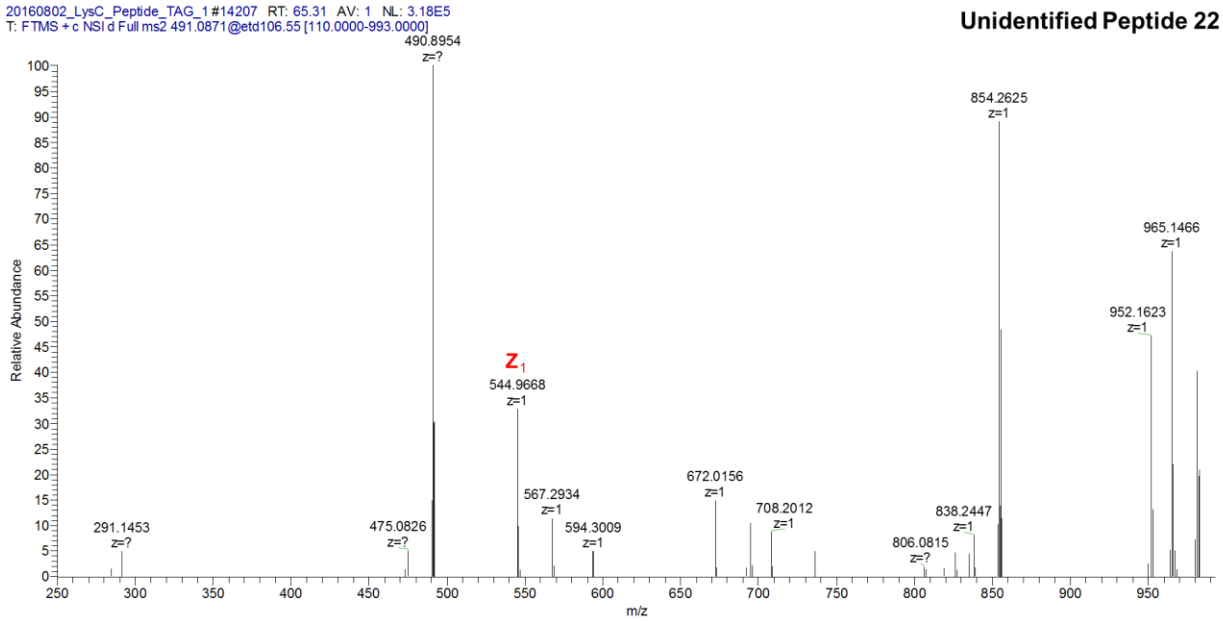


Figure 5.100. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 22, Undetected by Database but Detected Using In-House Developed Script.

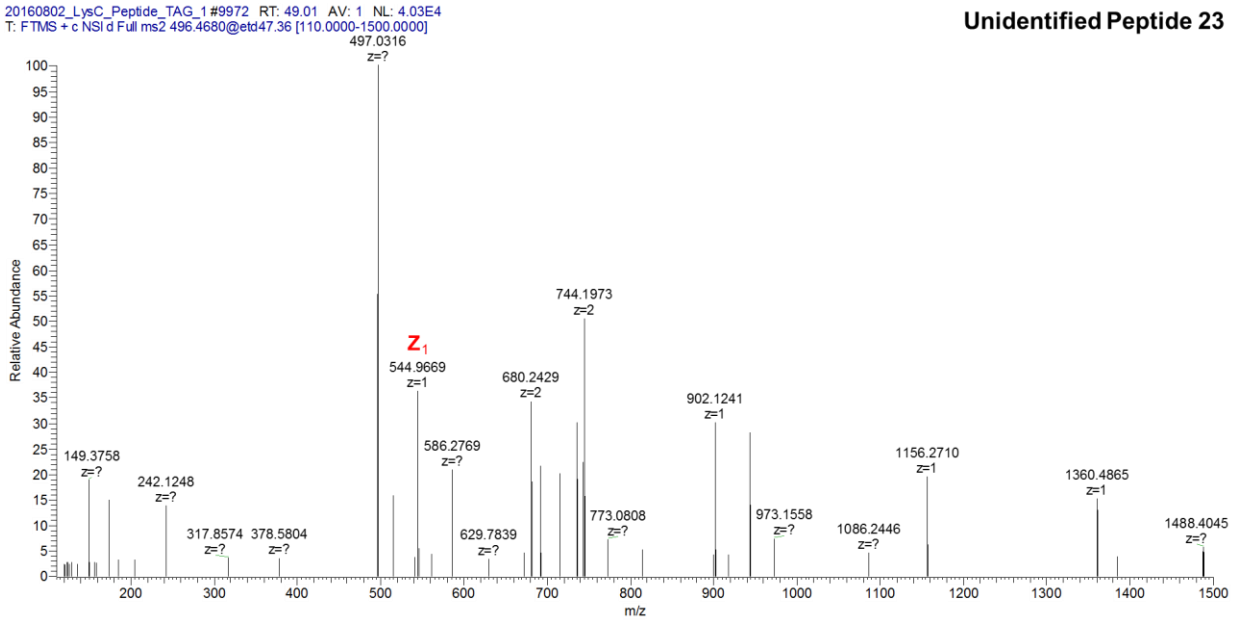


Figure 5.101. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 23, Undetected by Database but Detected Using In-House Developed Script.

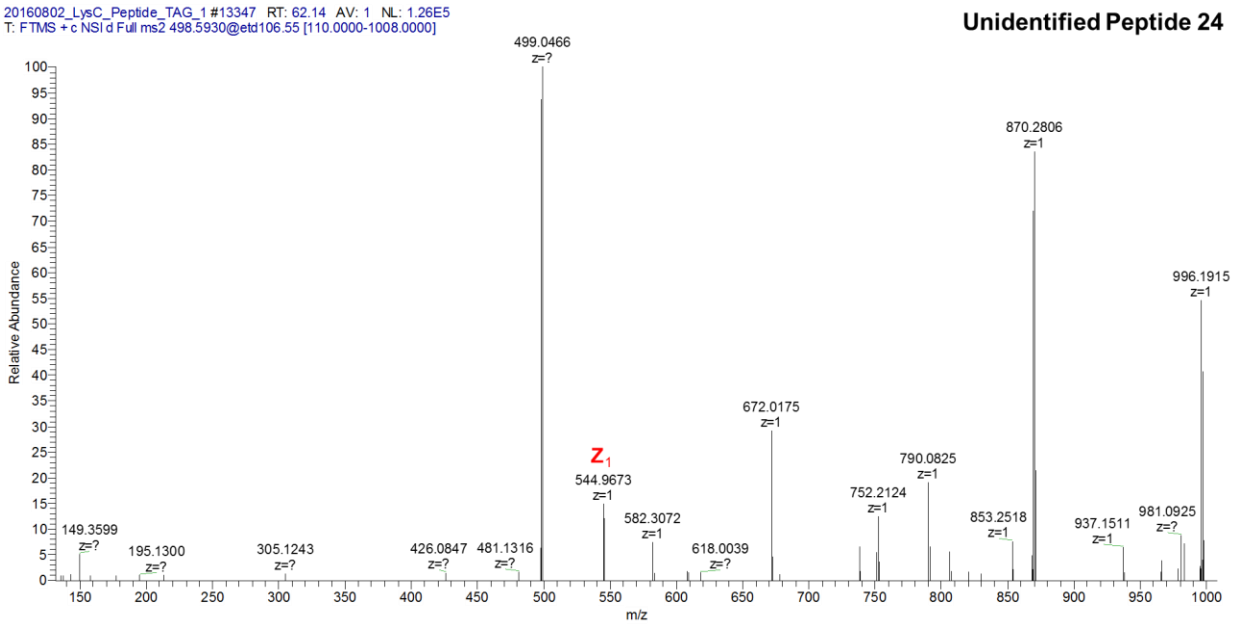


Figure 5.102. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 24, Undetected by Database but Detected Using In-House Developed Script.

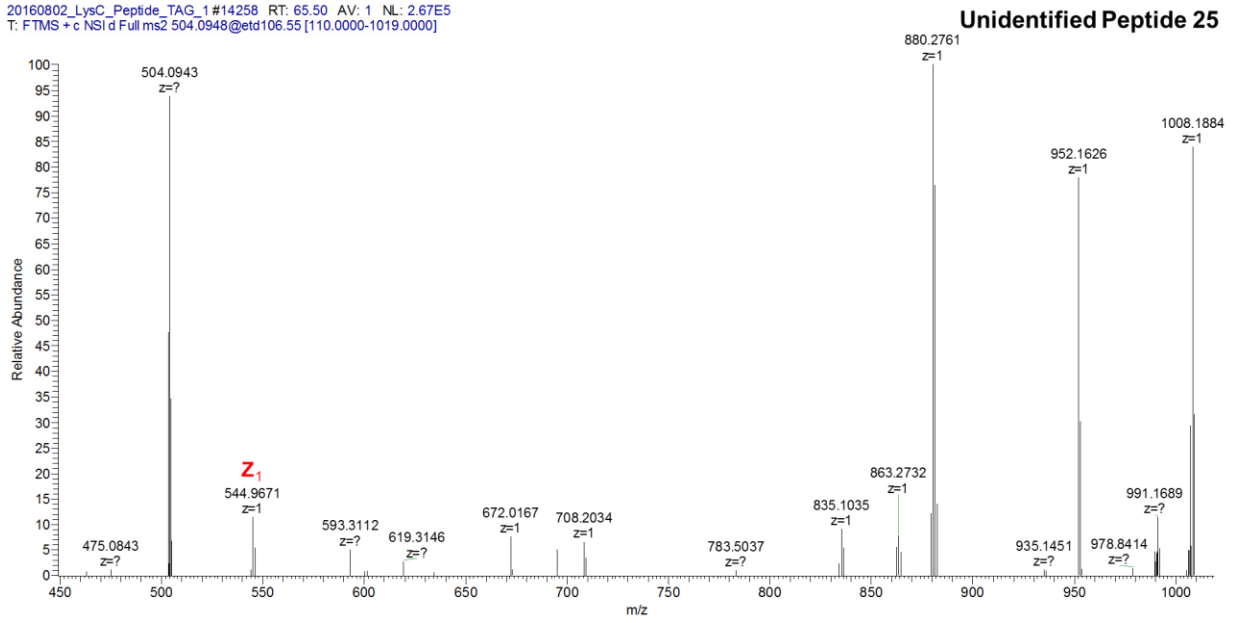


Figure 5.103. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 25, Undetected by Database but Detected Using In-House Developed Script.

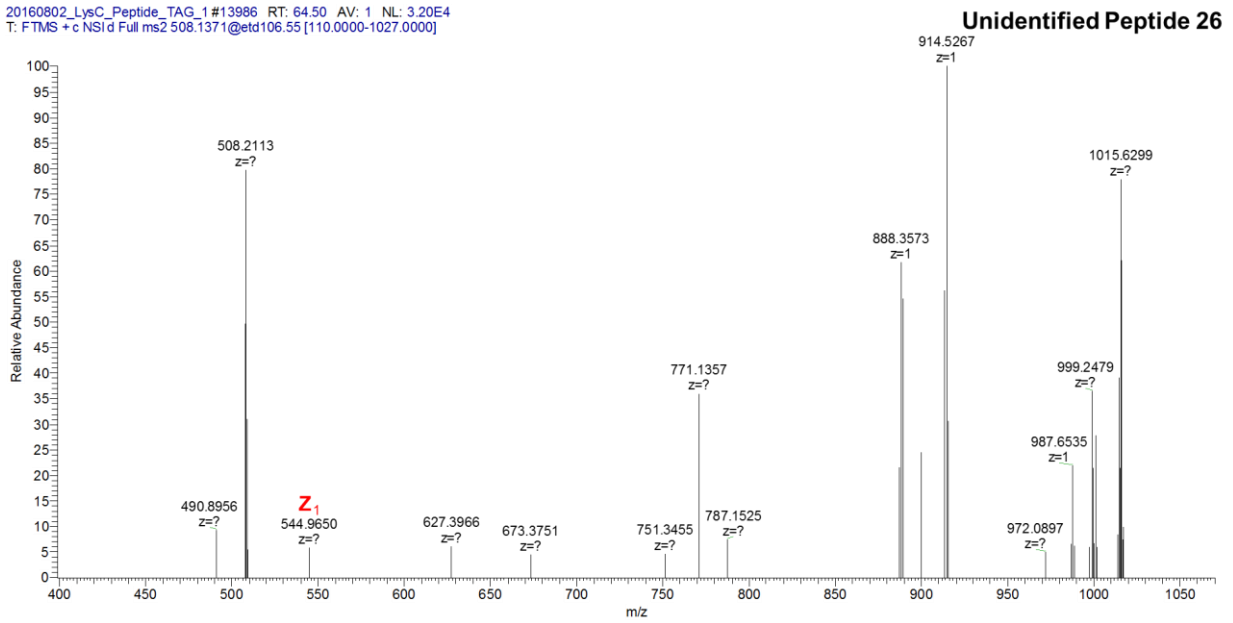


Figure 5.104. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 26, Undetected by Database but Detected Using In-House Developed Script.

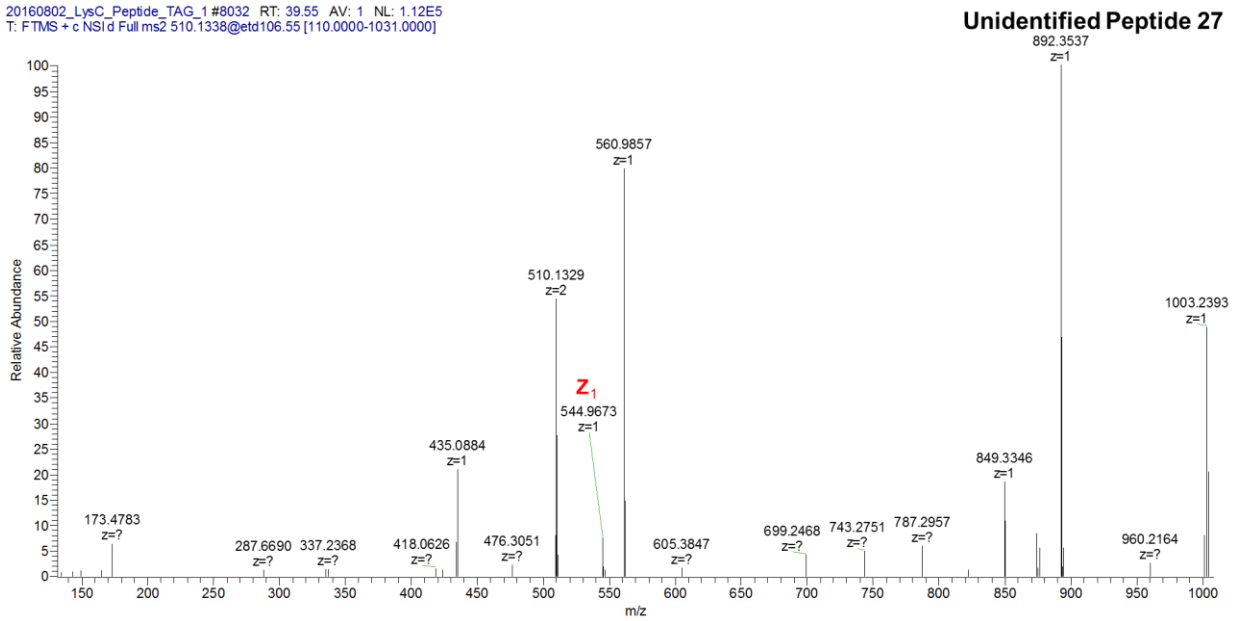


Figure 5.105. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 27, Undetected by Database but Detected Using In-House Developed Script.

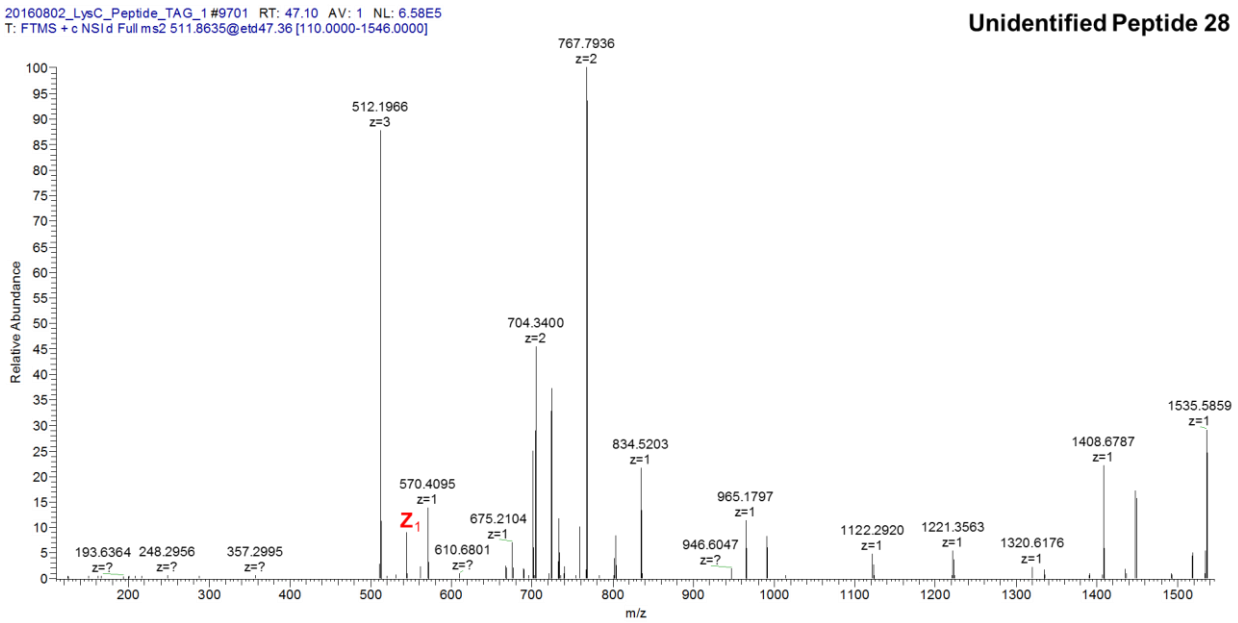


Figure 5.106. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 28, Undetected by Database but Detected Using In-House Developed Script.

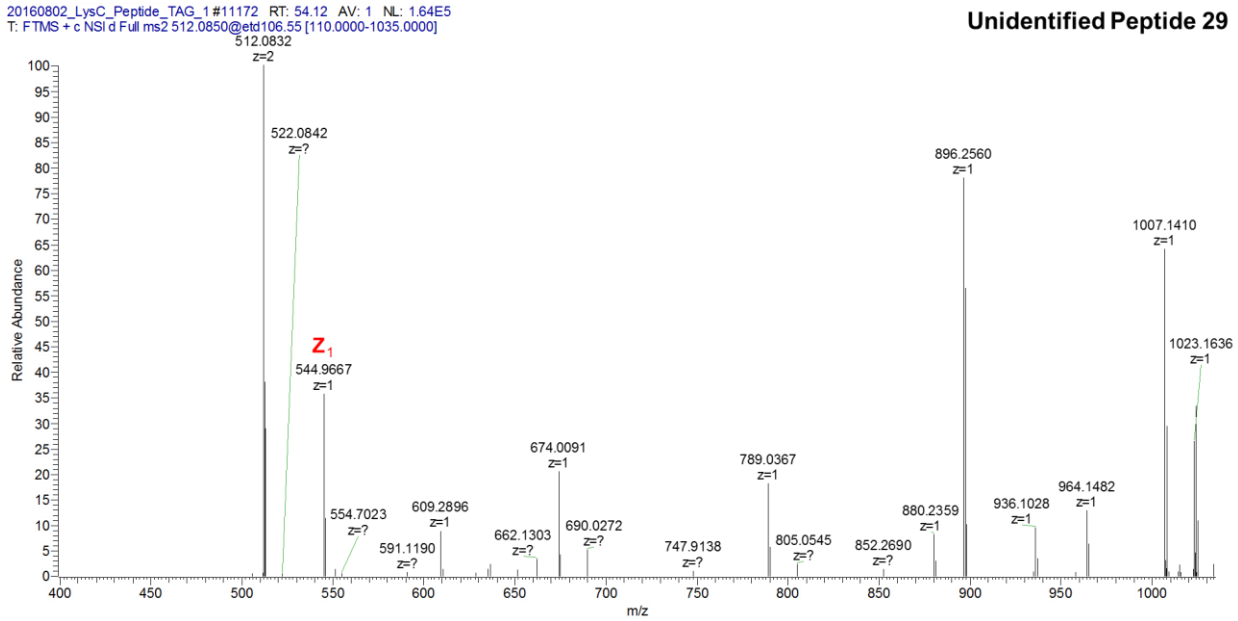


Figure 5.107. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 29, Undetected by Database but Detected Using In-House Developed Script.

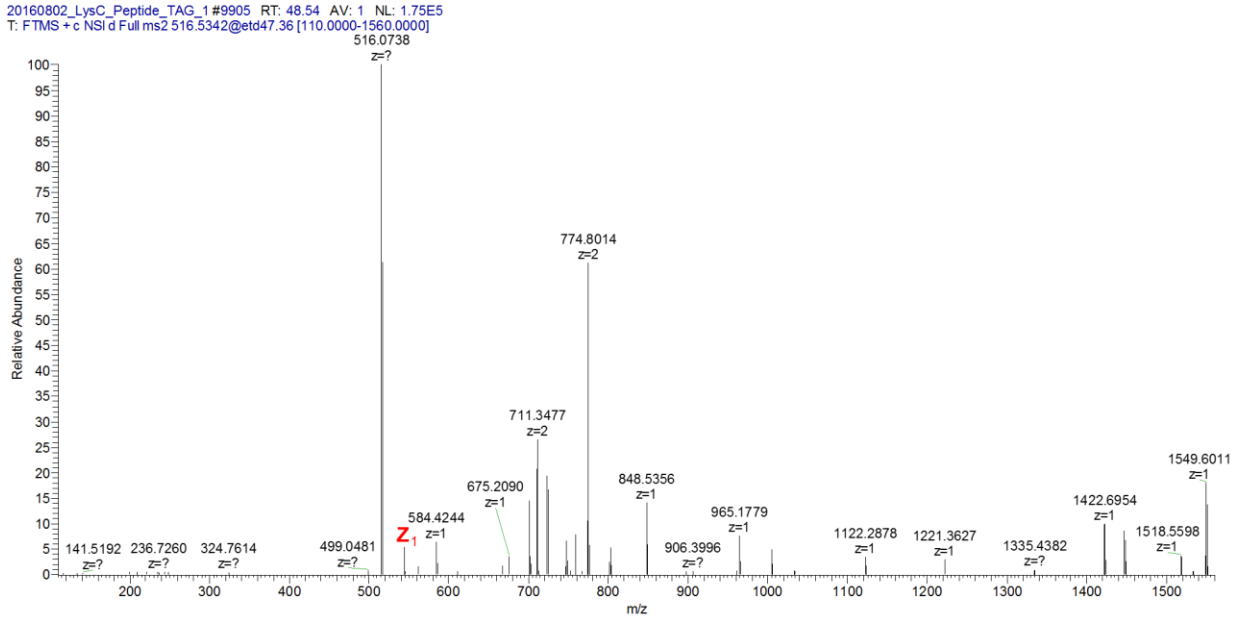


Figure 5.108. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 30, Undetected by Database but Detected Using In-House Developed Script.

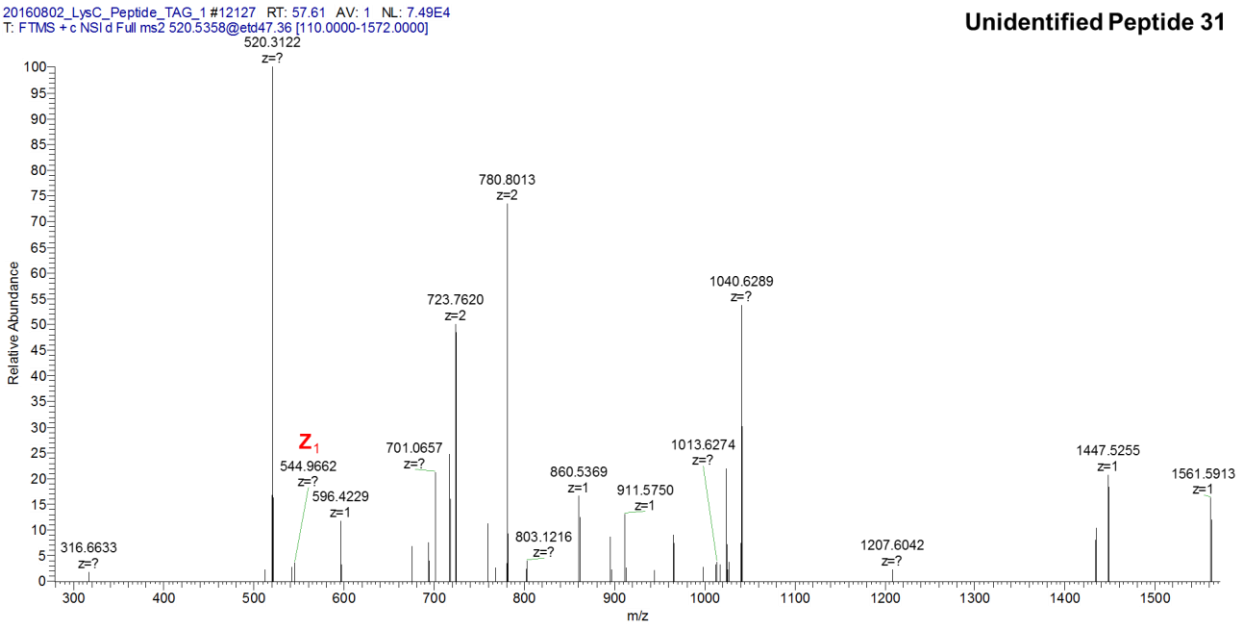


Figure 5.109. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 31, Undetected by Database but Detected Using In-House Developed Script.

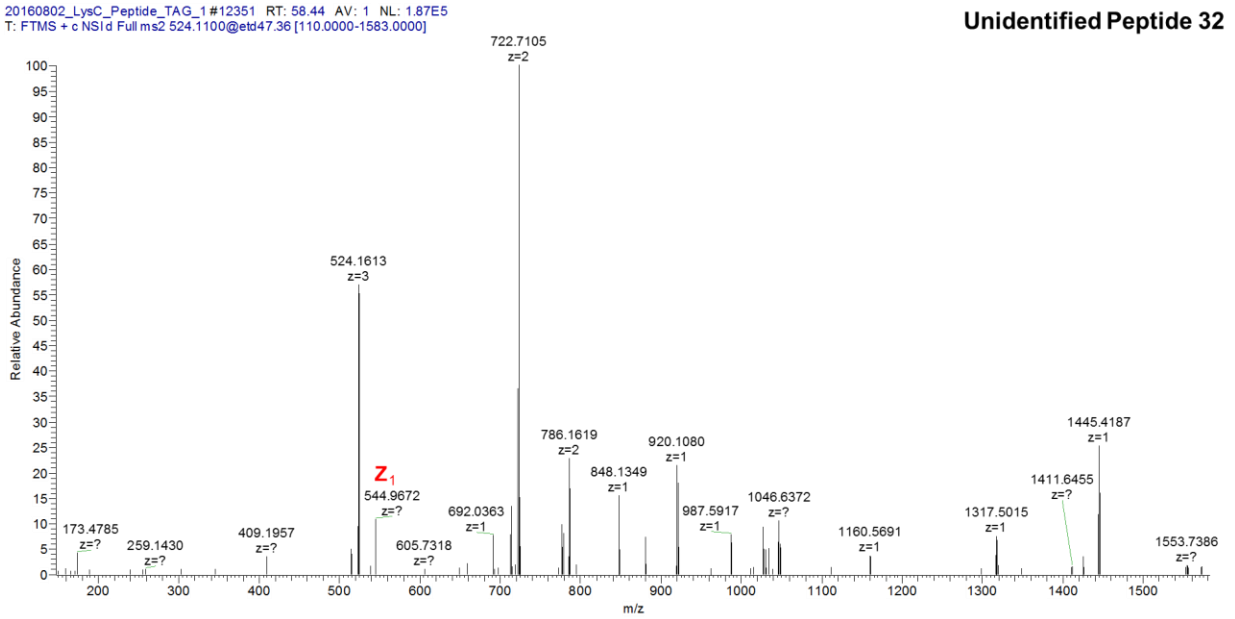


Figure 5.110. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 32, Undetected by Database but Detected Using In-House Developed Script.

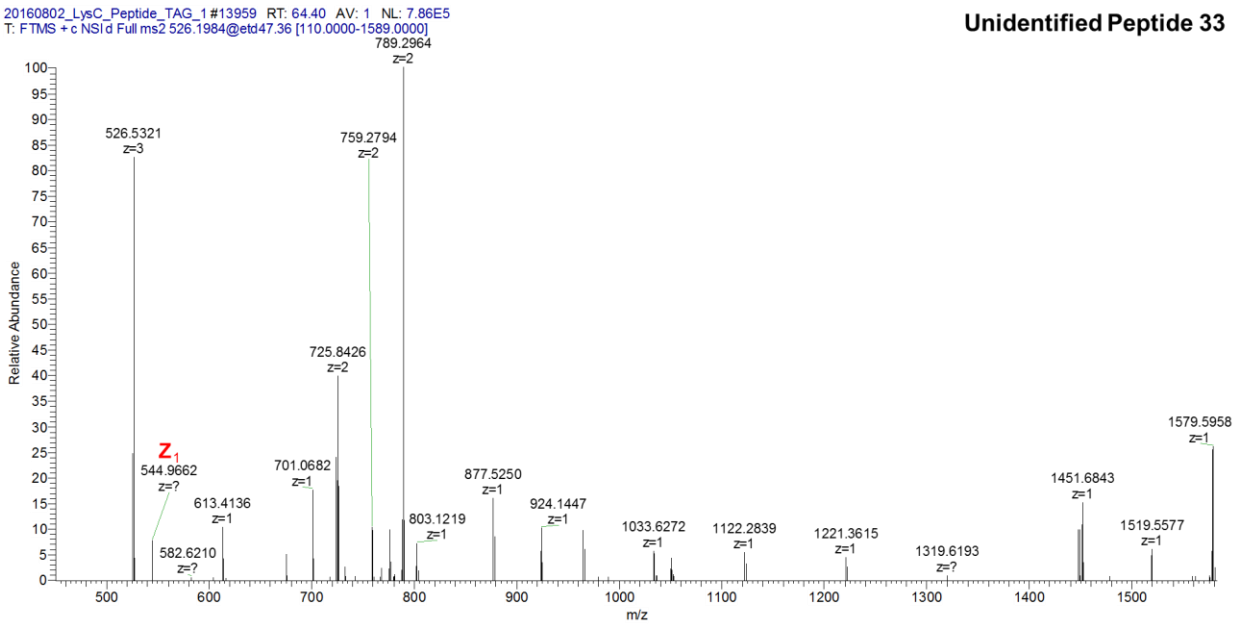


Figure 5.111. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 33, Undetected by Database but Detected Using In-House Developed Script.

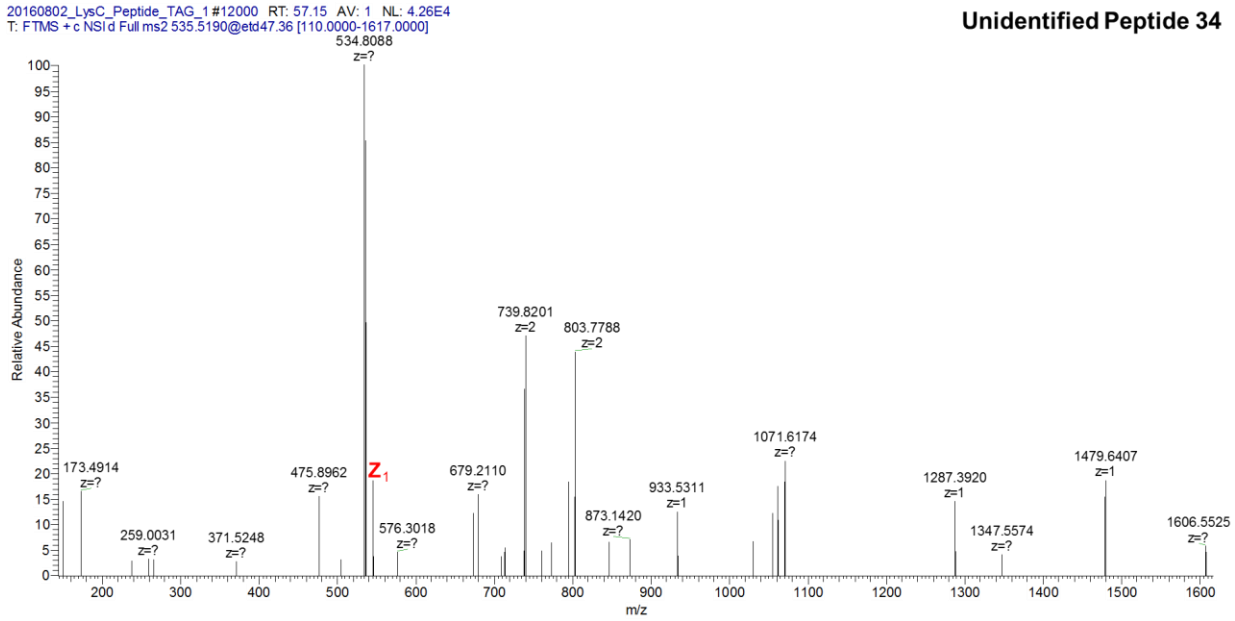


Figure 5.112. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 34, Undetected by Database but Detected Using In-House Developed Script.

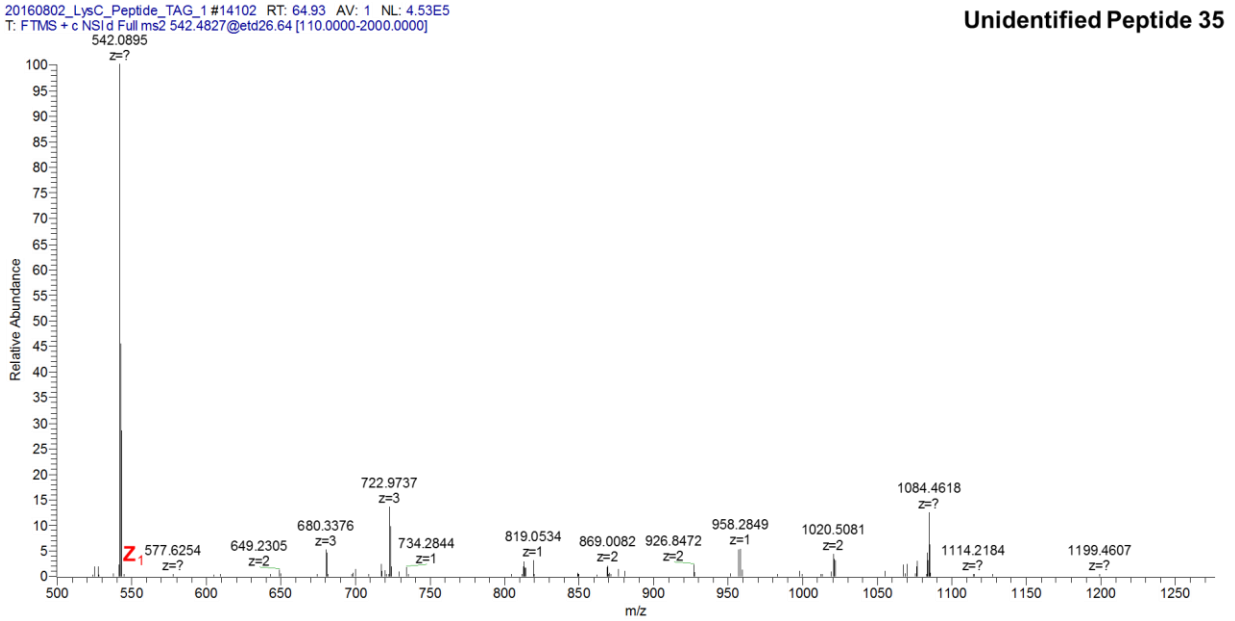


Figure 5.113. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 35, Undetected by Database but Detected Using In-House Developed Script.

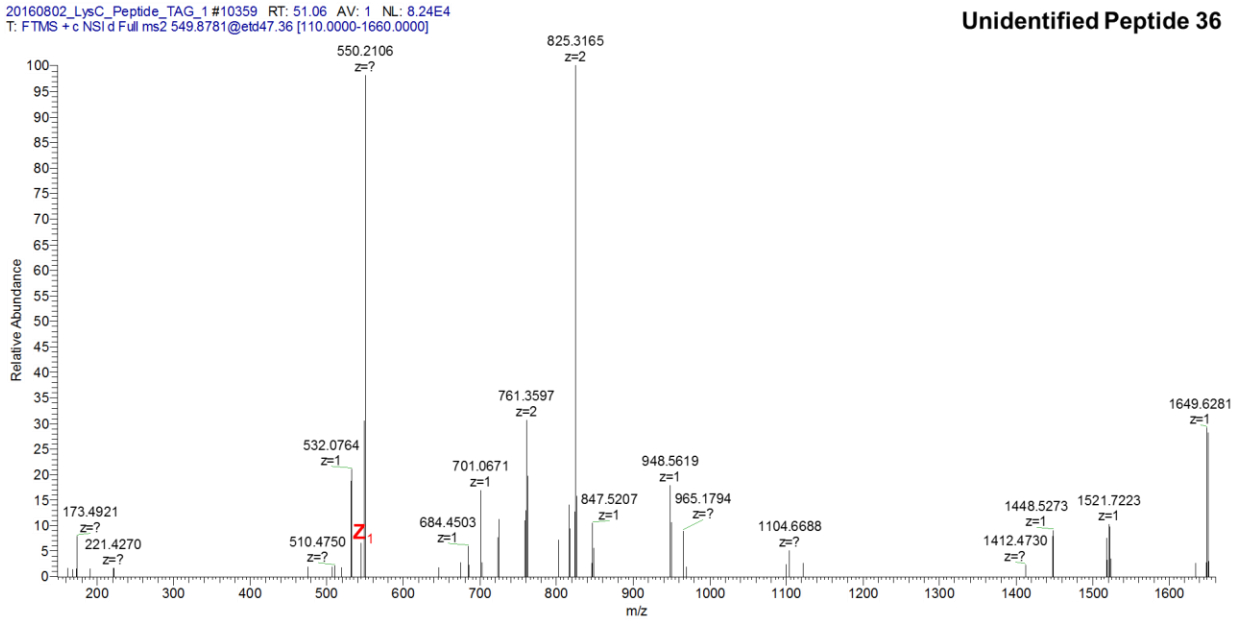


Figure 5.114. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 36, Undetected by Database but Detected Using In-House Developed Script.

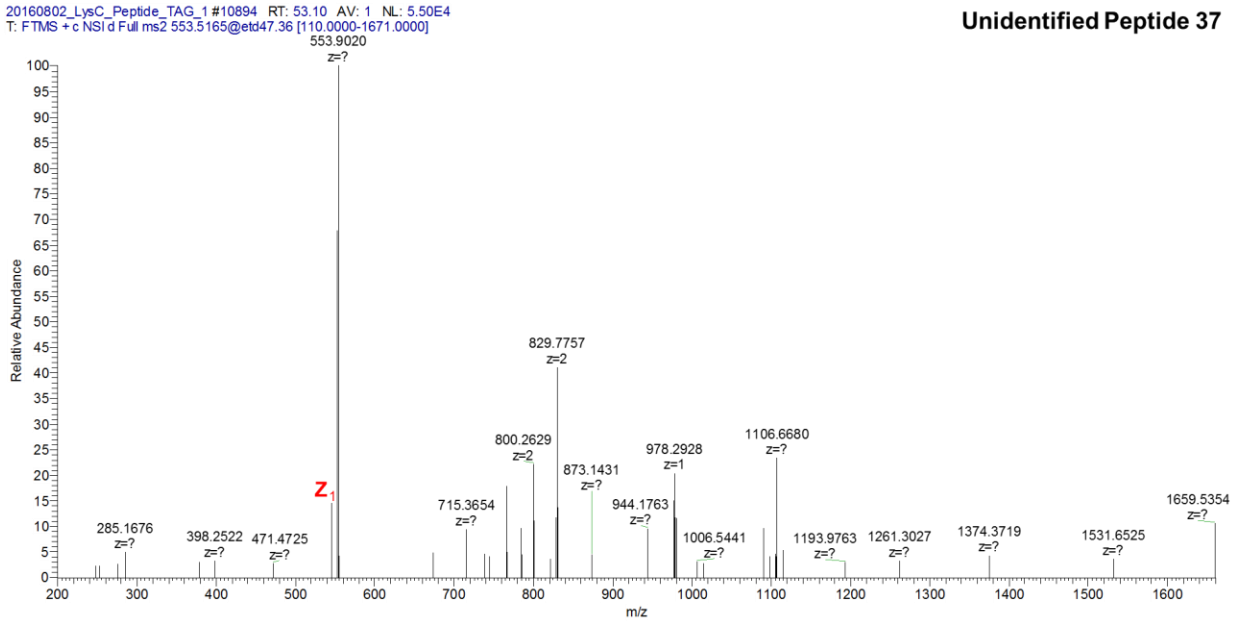


Figure 5.115. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 37, Undetected by Database but Detected Using In-House Developed Script.

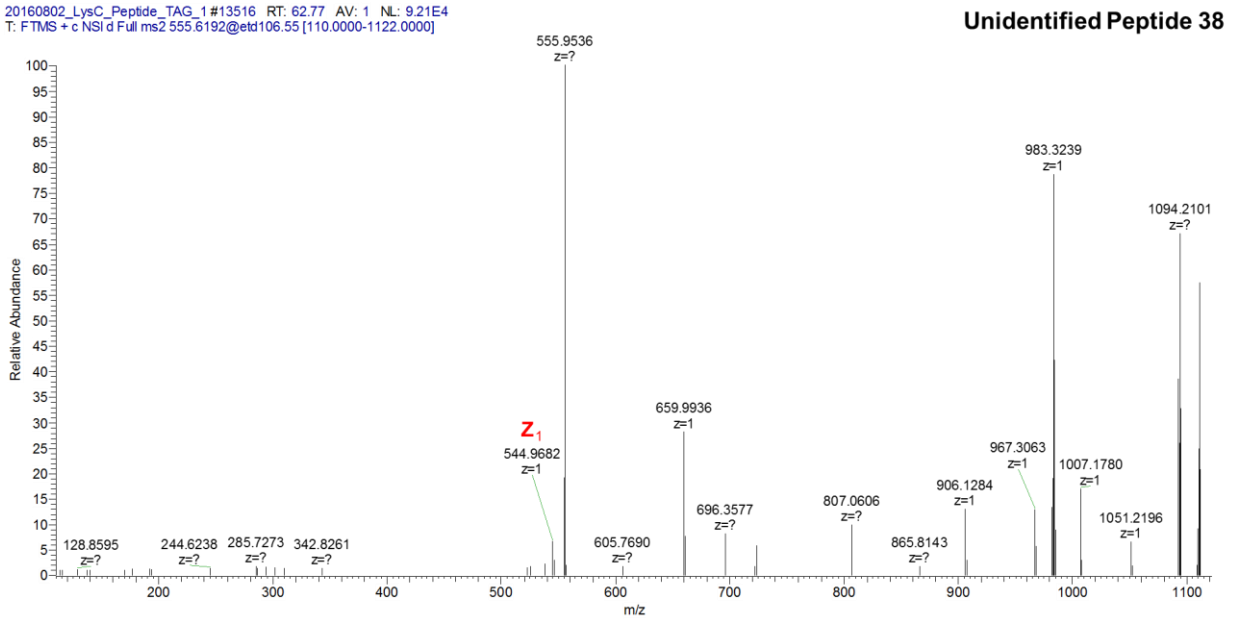


Figure 5.116. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 38, Undetected by Database but Detected Using In-House Developed Script.

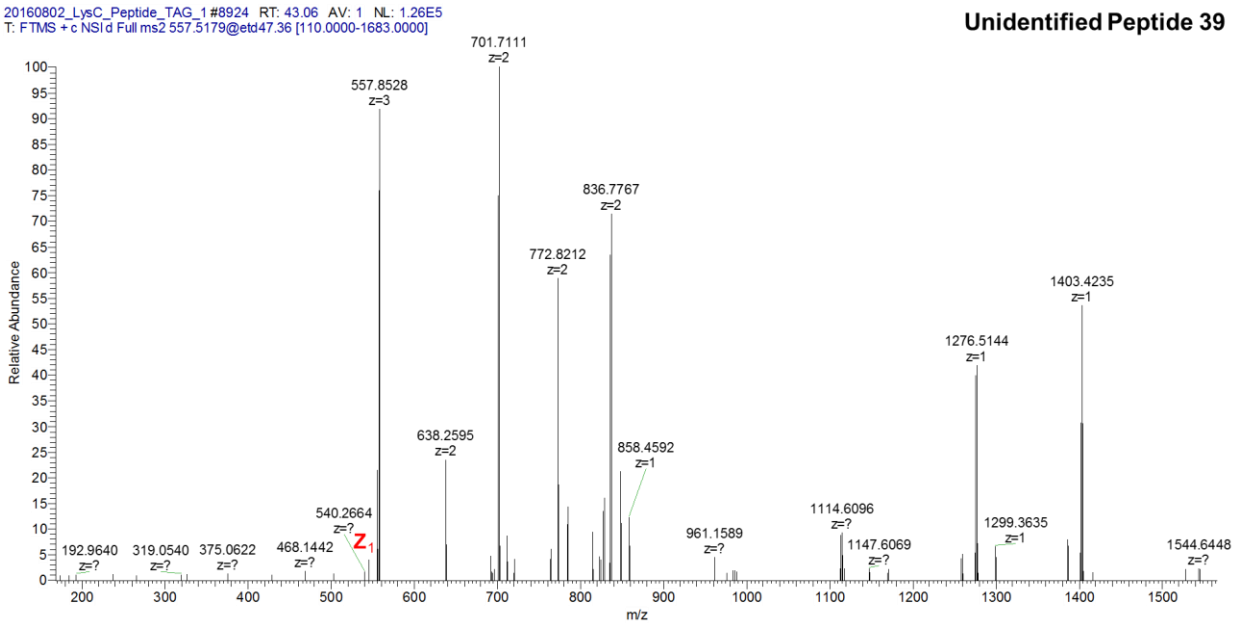


Figure 5.117. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 39, Undetected by Database but Detected Using In-House Developed Script.

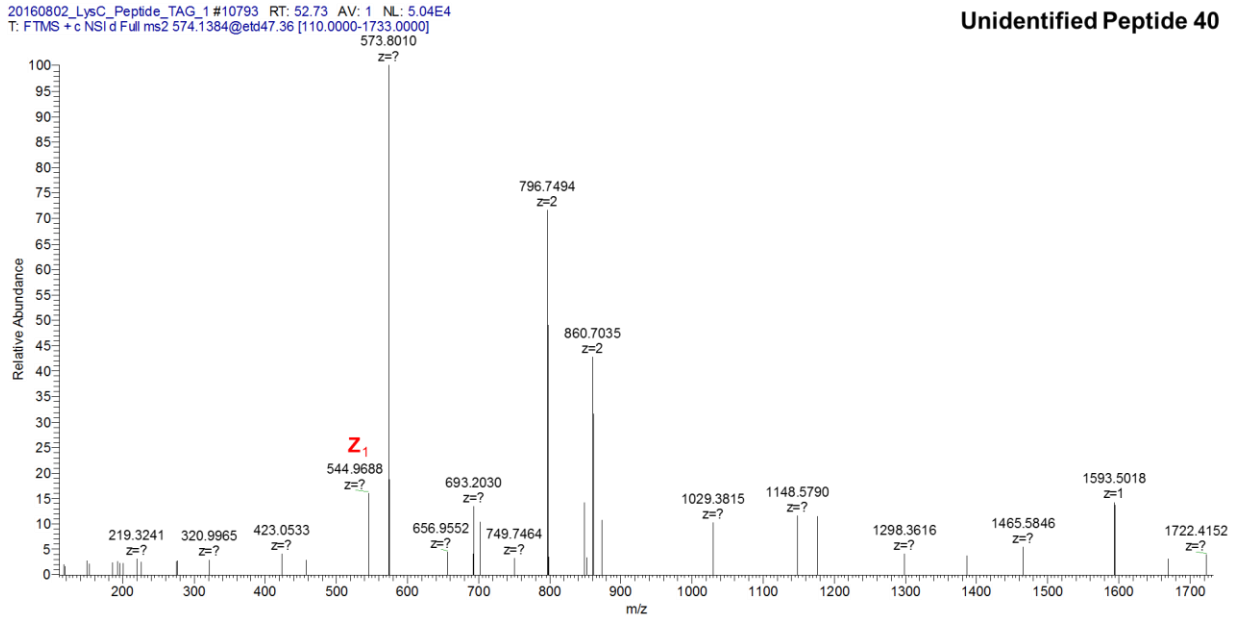


Figure 5.118. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 40, Undetected by Database but Detected Using In-House Developed Script.

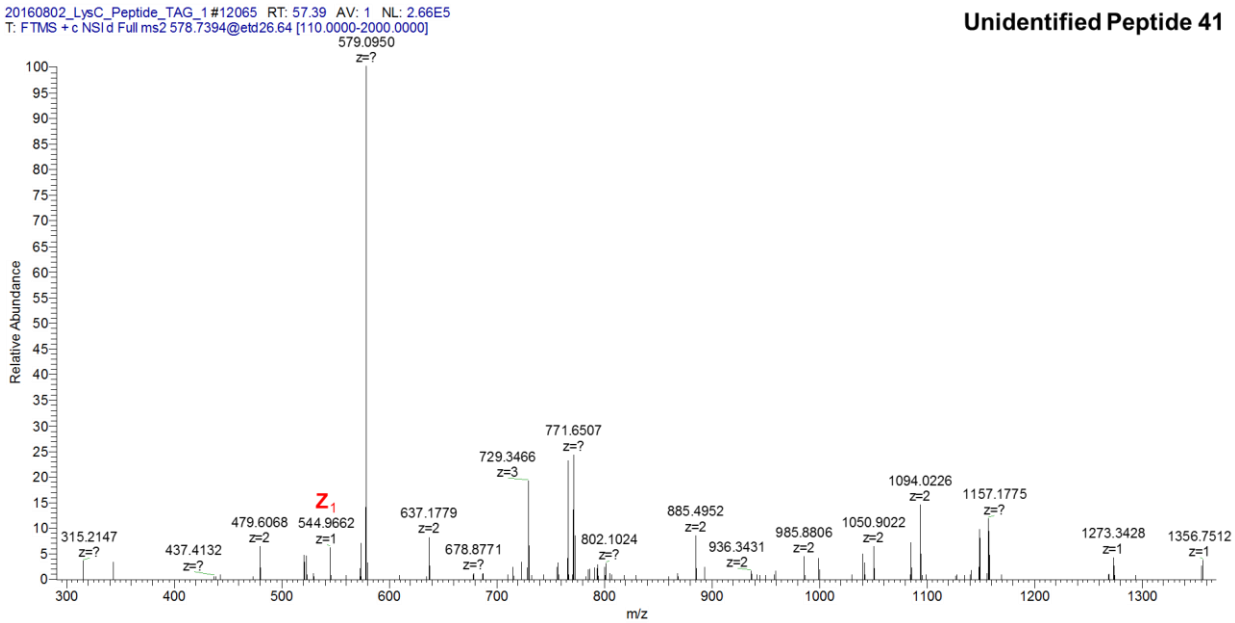


Figure 5.119. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 41, Undetected by Database but Detected Using In-House Developed Script.

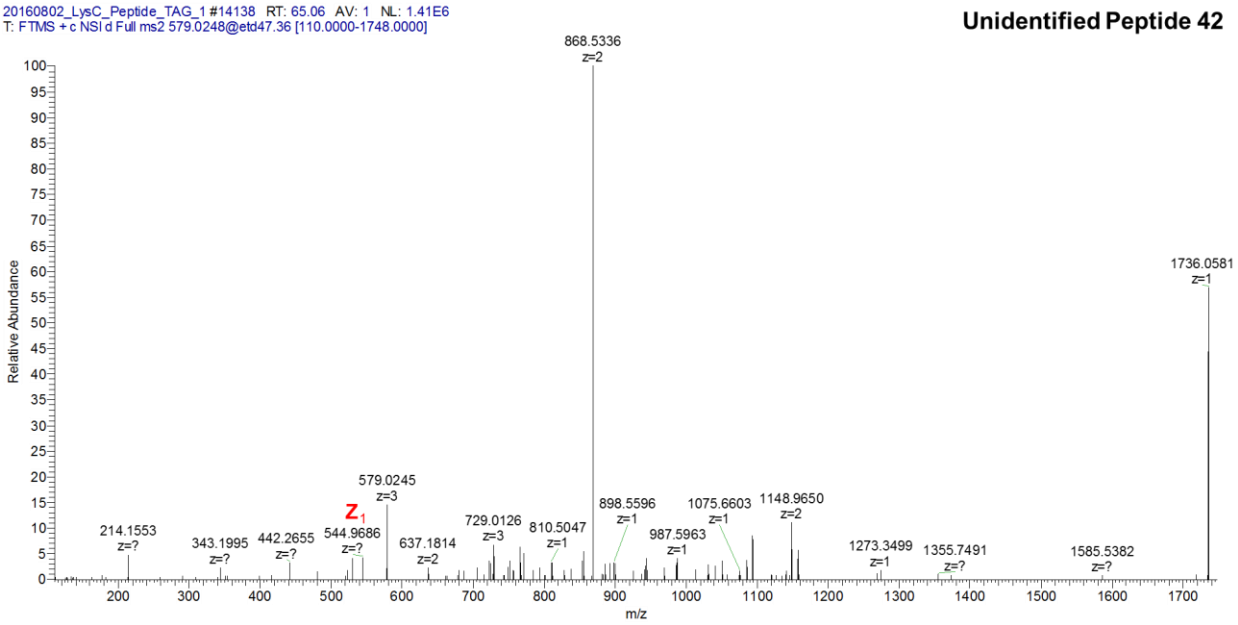


Figure 5.120. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 42, Undetected by Database but Detected Using In-House Developed Script.

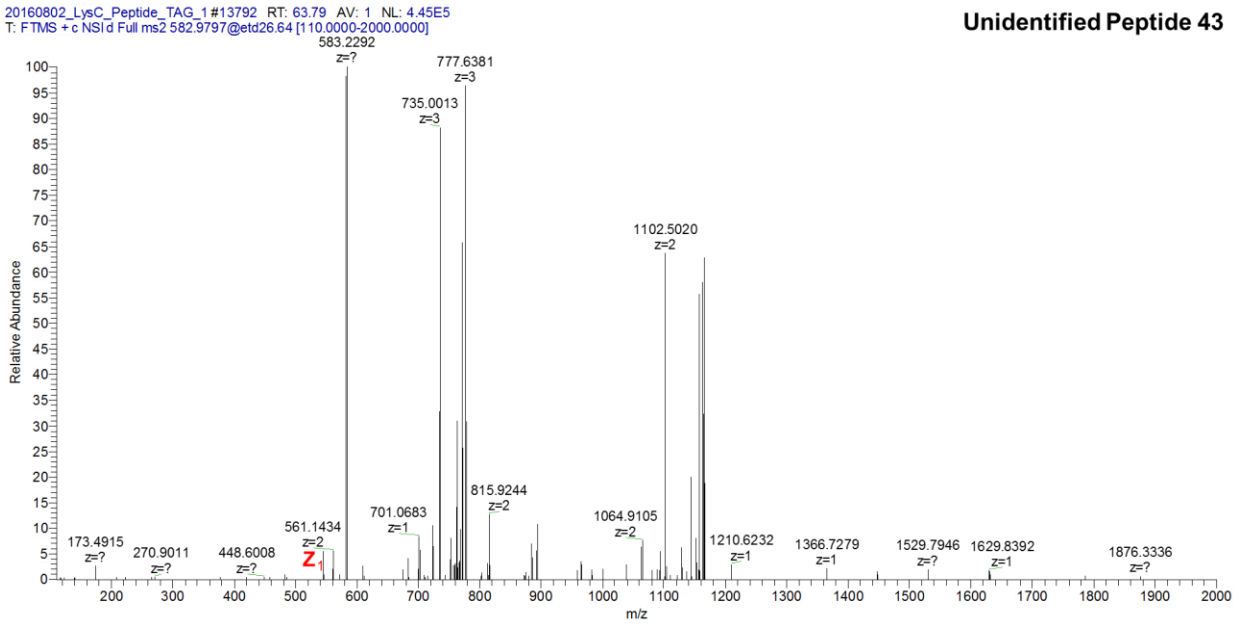


Figure 5.121. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 43, Undetected by Database but Detected Using In-House Developed Script.

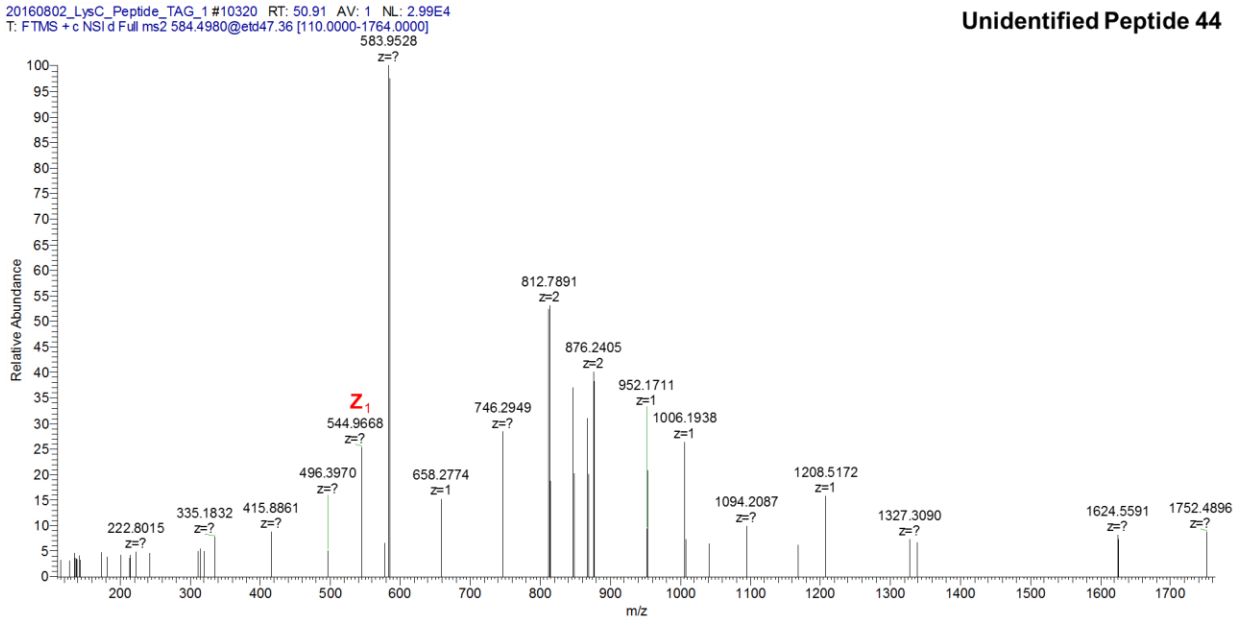


Figure 5.122. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 44, Undetected by Database but Detected Using In-House Developed Script.

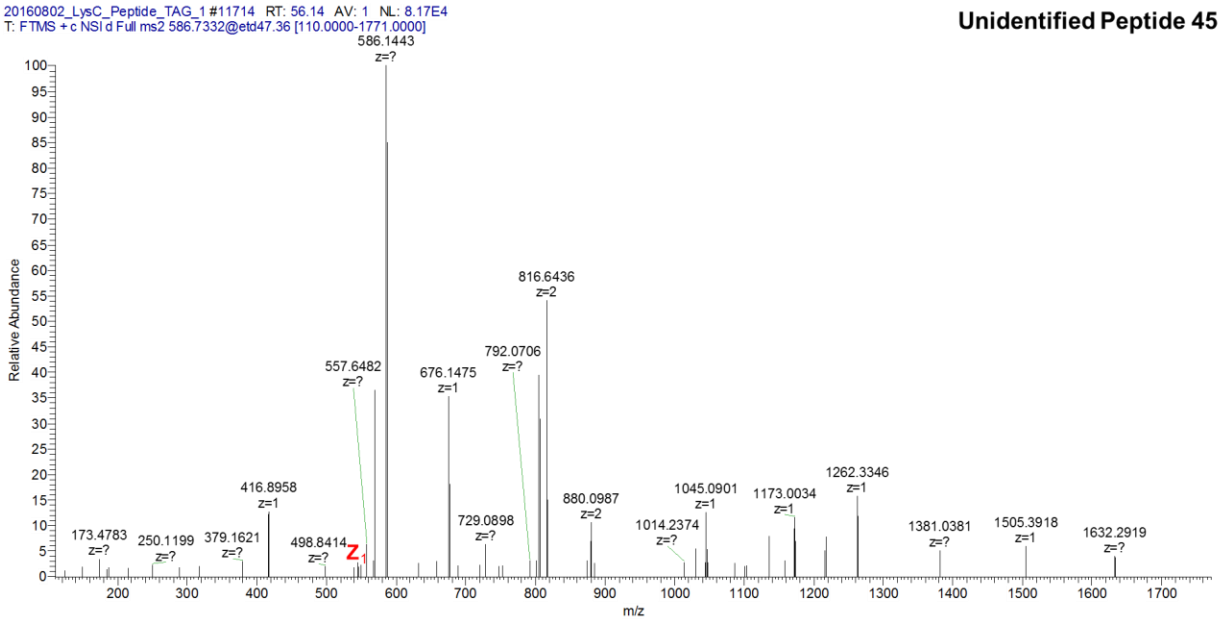


Figure 5.123. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 45, Undetected by Database but Detected Using In-House Developed Script.

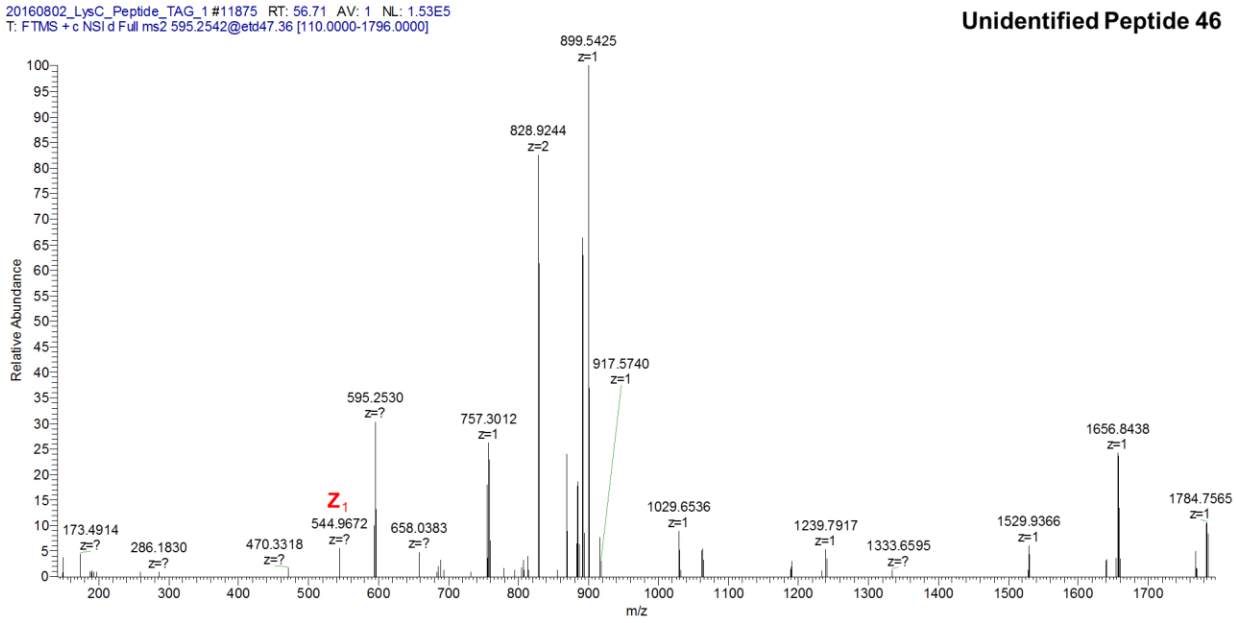


Figure 5.124. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 46, Undetected by Database but Detected Using In-House Developed Script.

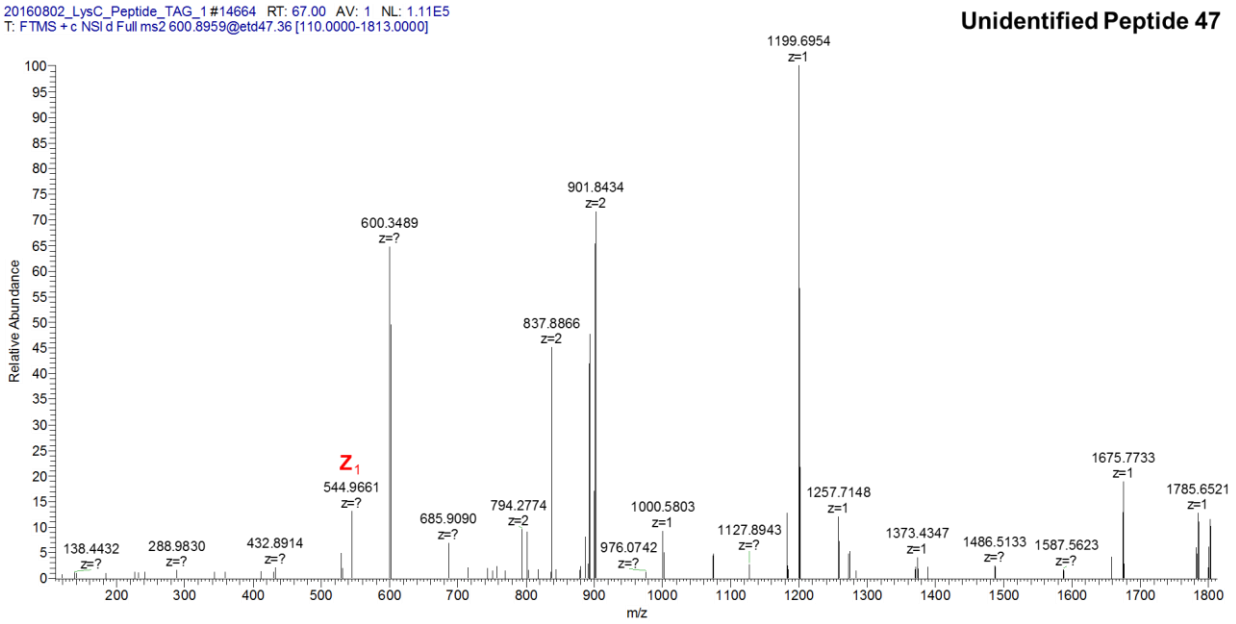


Figure 5.125. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 47, Undetected by Database but Detected Using In-House Developed Script.

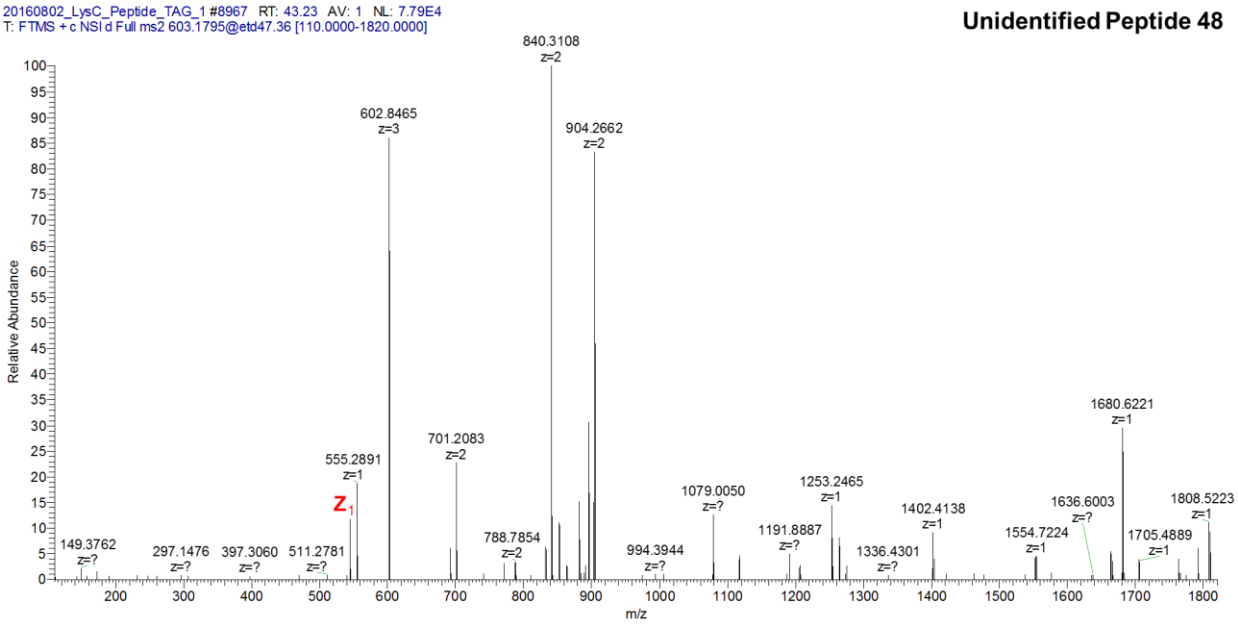


Figure 5.126. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 48, Undetected by Database but Detected Using In-House Developed Script.

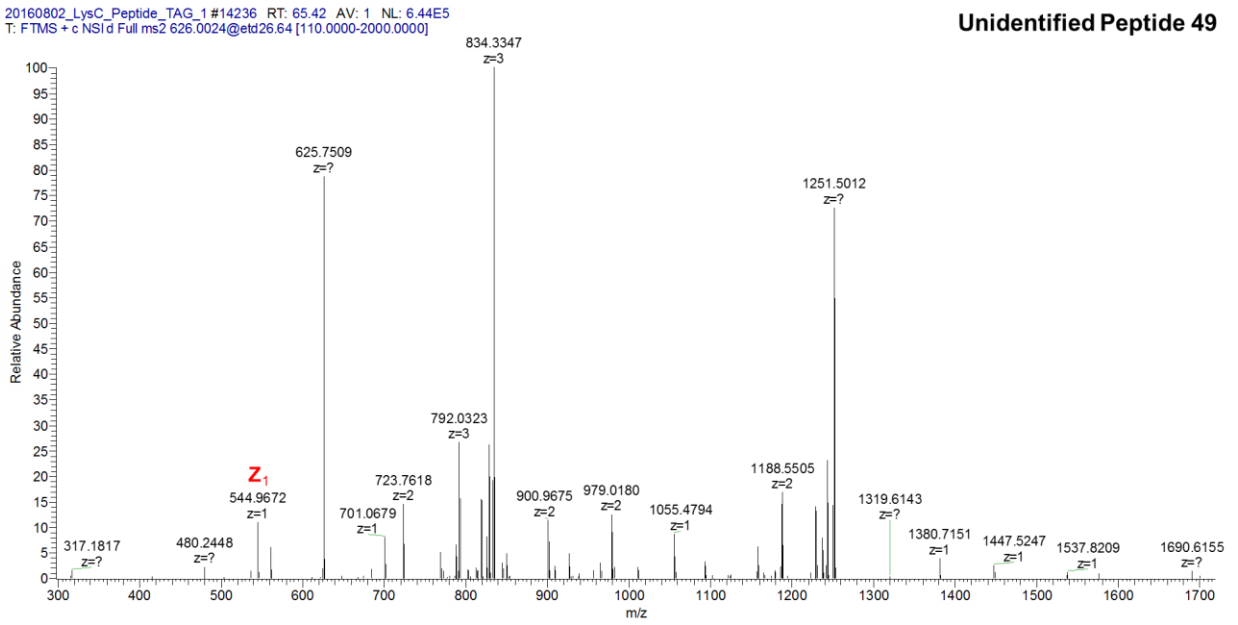


Figure 5.127. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 49, Undetected by Database but Detected Using In-House Developed Script.

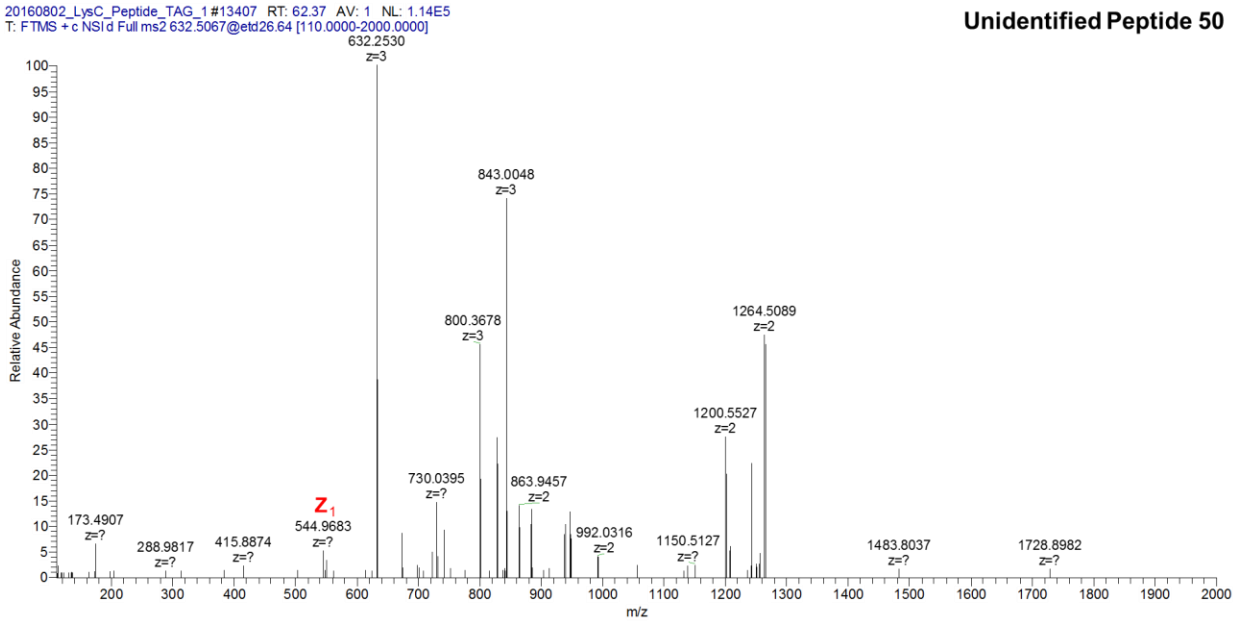


Figure 5.128. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 50, Undetected by Database but Detected Using In-House Developed Script.

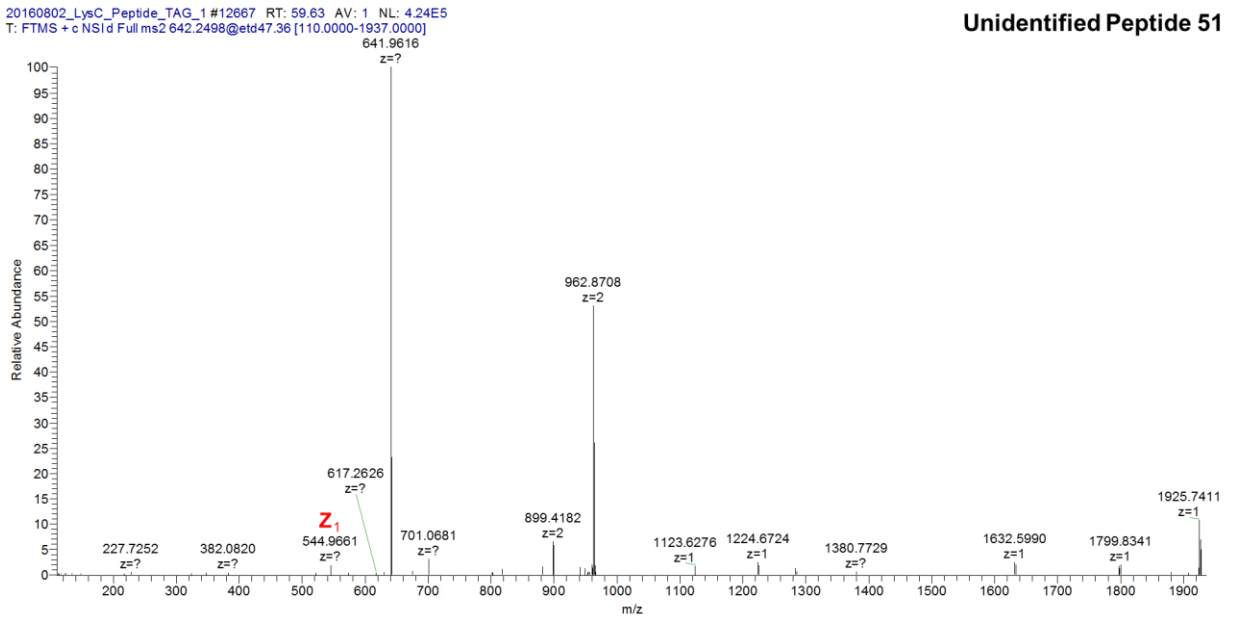


Figure 5.129. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 51, Undetected by Database but Detected Using In-House Developed Script.

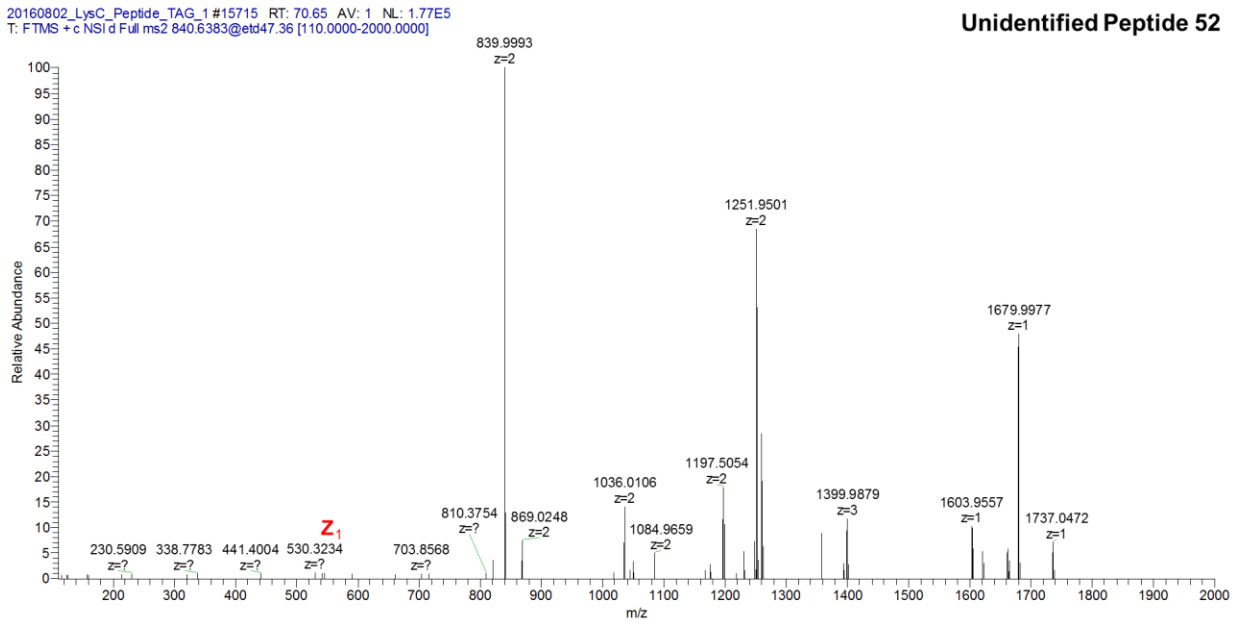


Figure 5.130. ETD-MS2 of LysC-Digested Unidentified Tagged Peptide 52, Undetected by Database but Detected Using In-House Developed Script.

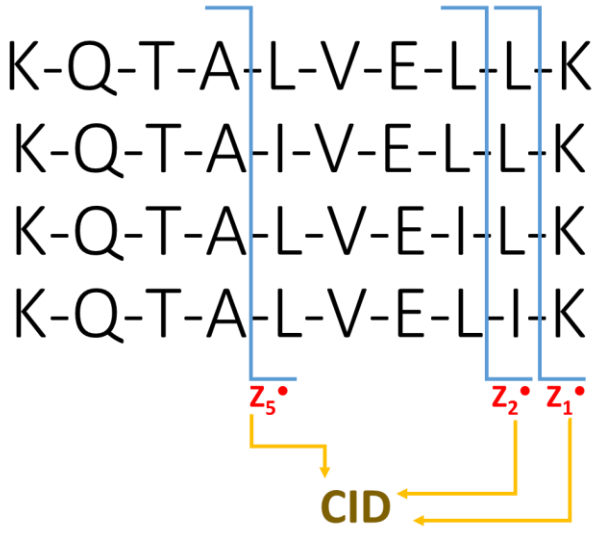


Figure 5.131. ETD-MS2-CID-MS3 Process for Leucine/Isoleucine Differentiation.

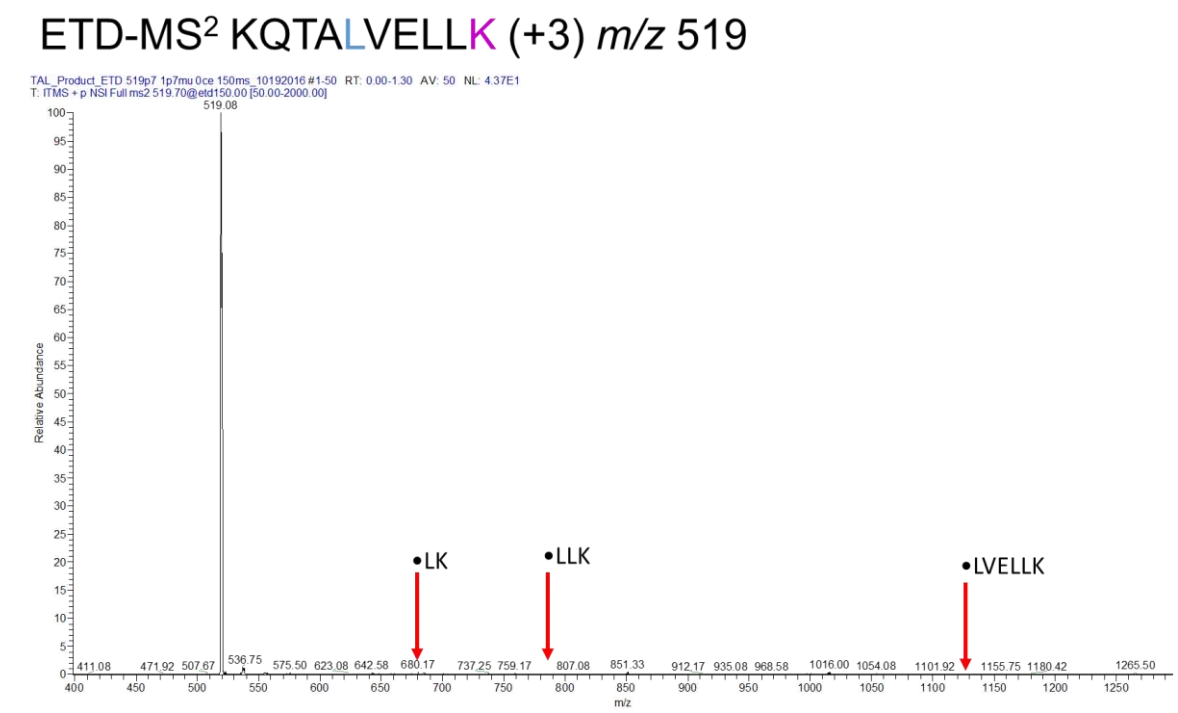


Figure 5.132. ETD-MS2 of triply charged KQTALVELLK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

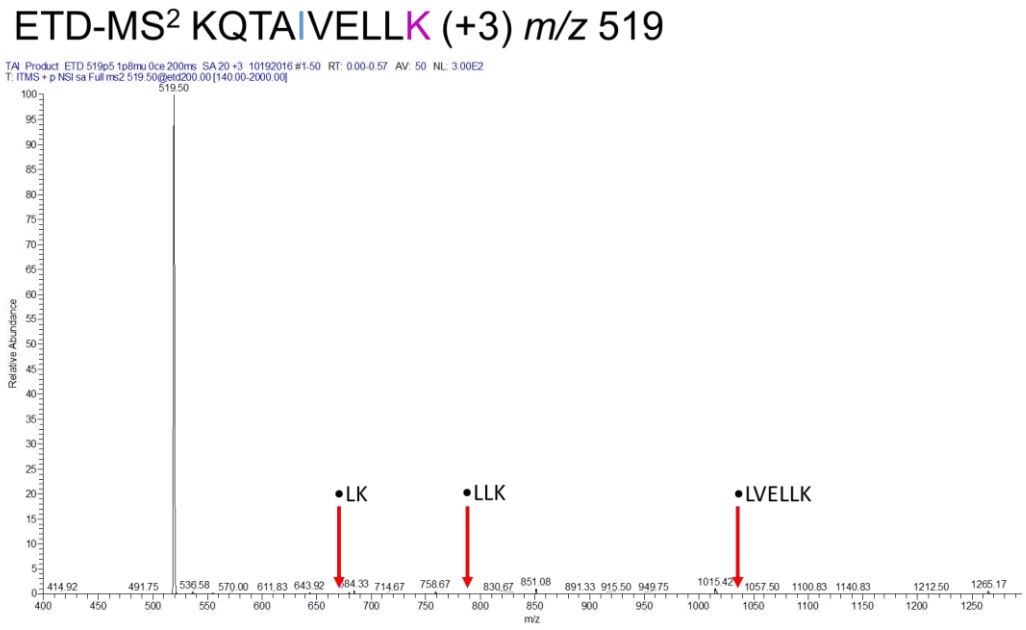


Figure 5.133. ETD-MS2 of triply charged KQTAIVELLK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

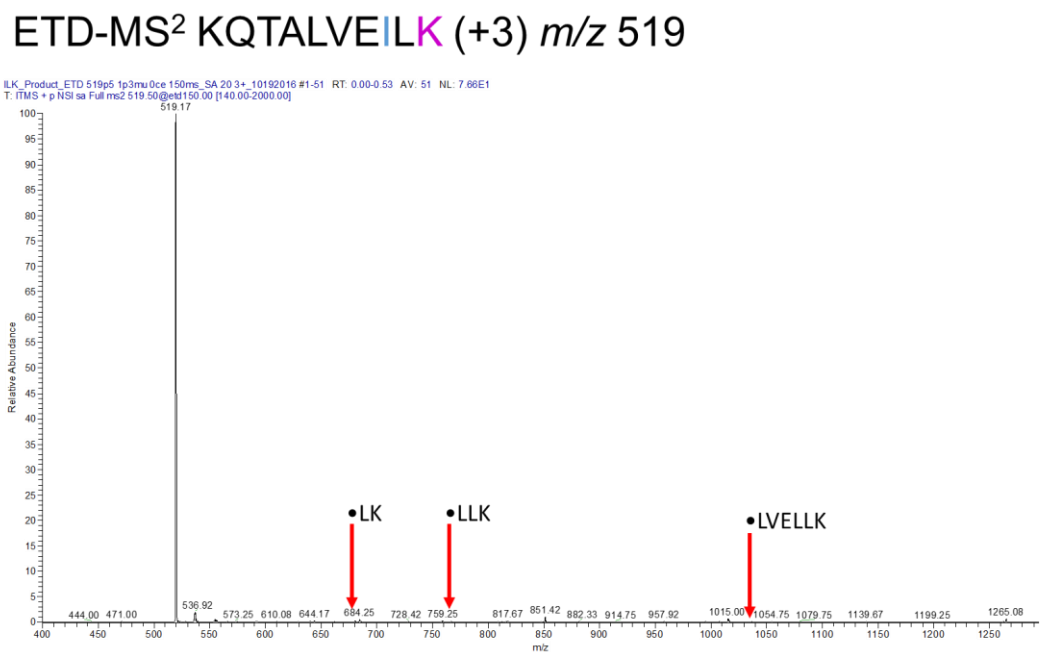


Figure 5.134. ETD-MS2 of triply charged KQTALVEILK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

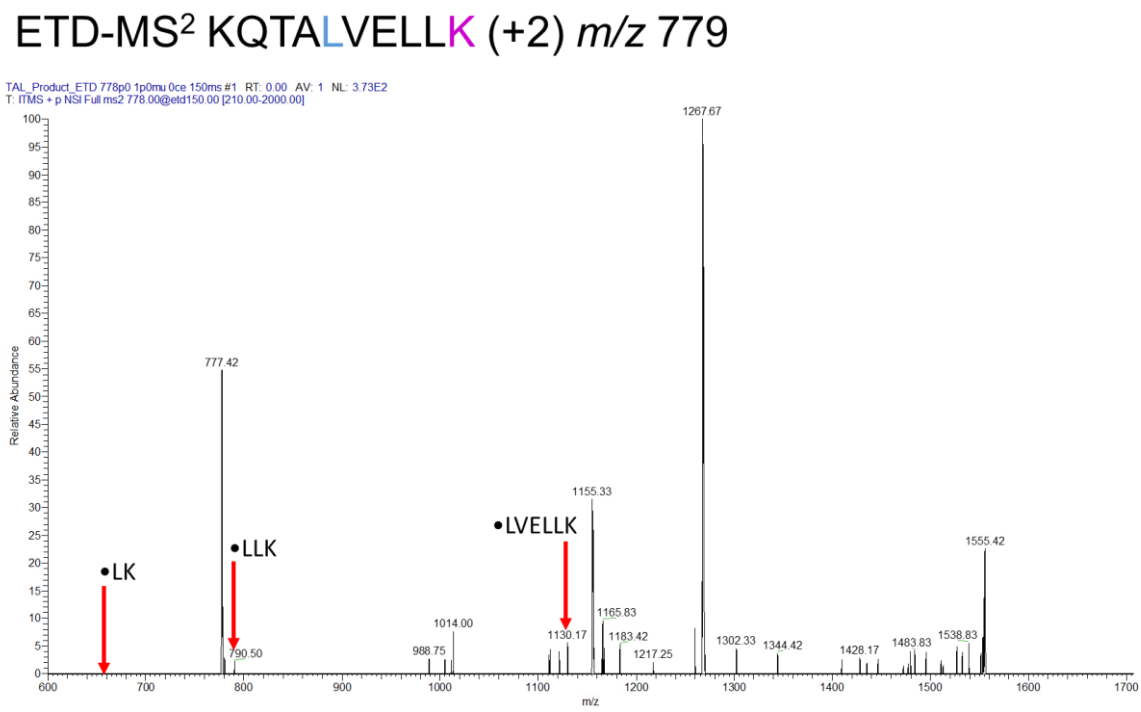


Figure 5.135. ETD-MS2 of doubly charged KQTALVELLK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

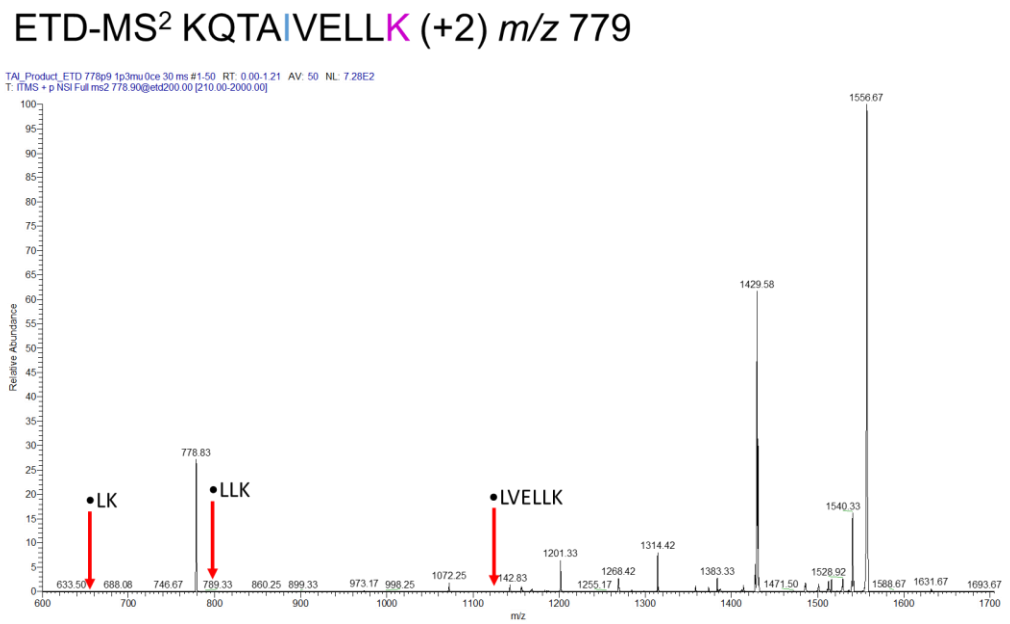


Figure 5.136. ETD-MS2 of doubly charged KQTAIVELLK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

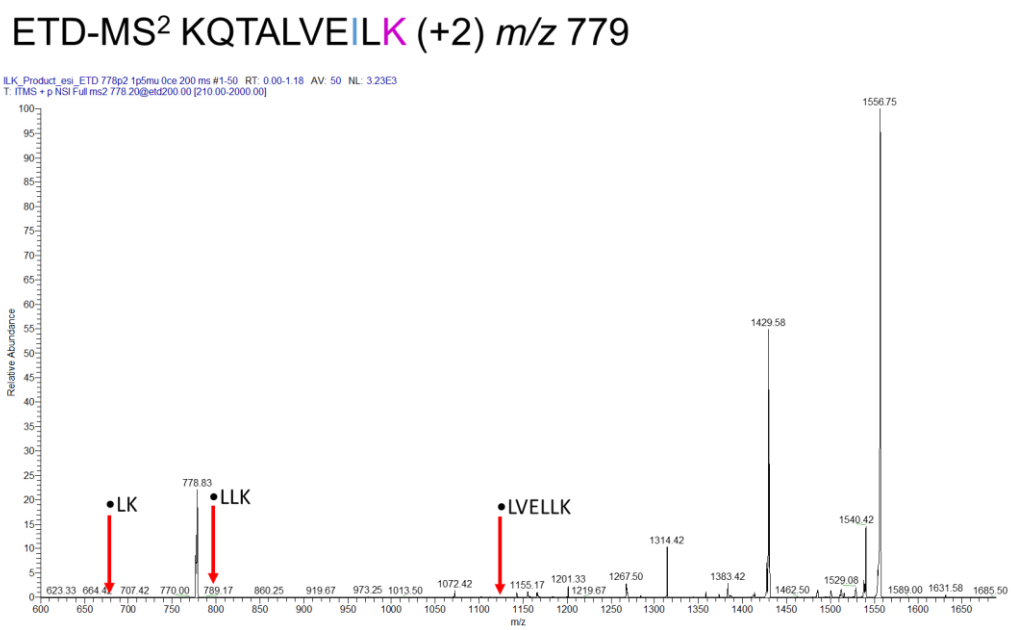


Figure 5.137. ETD-MS2 of doubly charged KQTALVEILK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

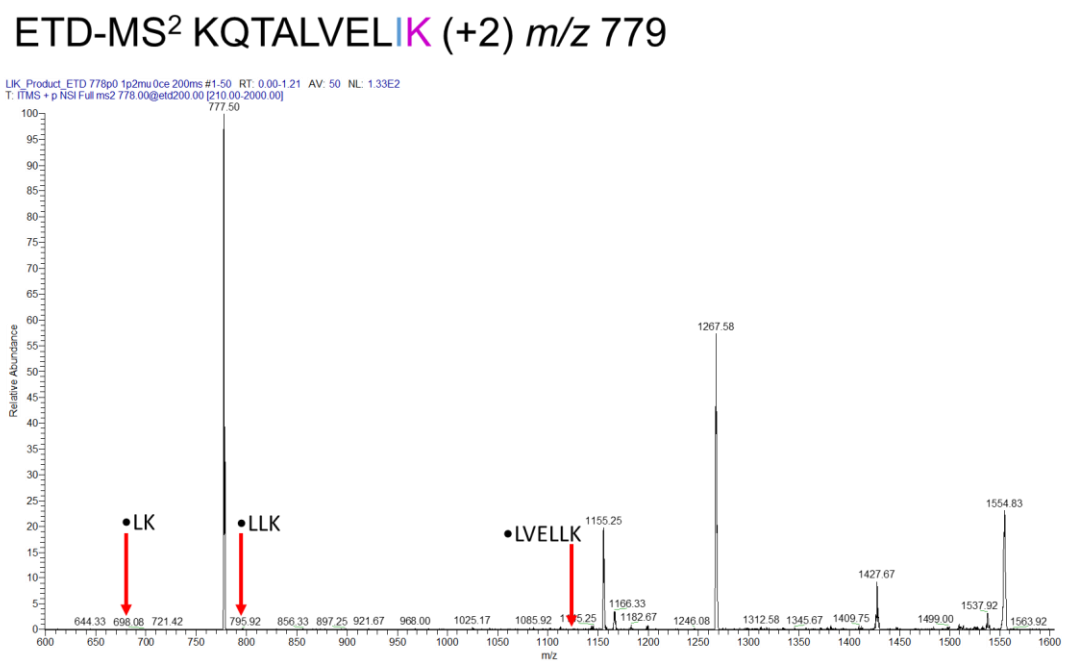


Figure 5.138. ETD-MS2 of doubly charged KQTALVEILK for the generation of z-radical ions. Note that the red arrows indicate the *m/z* that needed selection for Leu and Ile differentiation by CID.

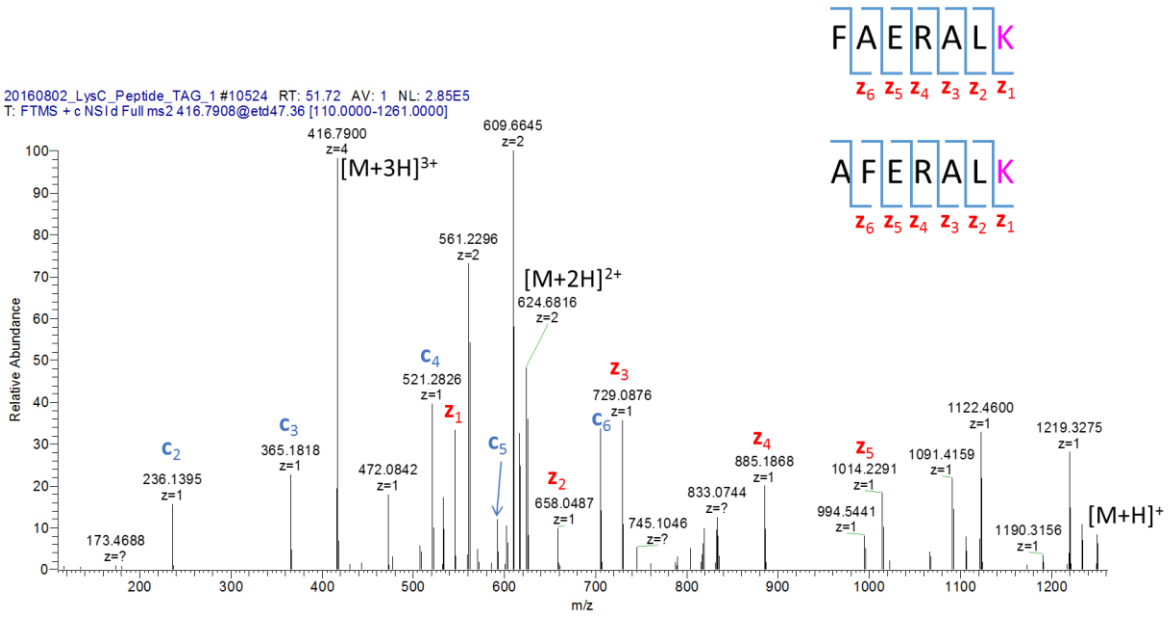


Figure 5.139. ETD-MS2 of triply charged "unidentified peptide 4" from Lys-C digest (**SI Figure 5.82**). Identified by de novo as either FAERALK or AFERALK.

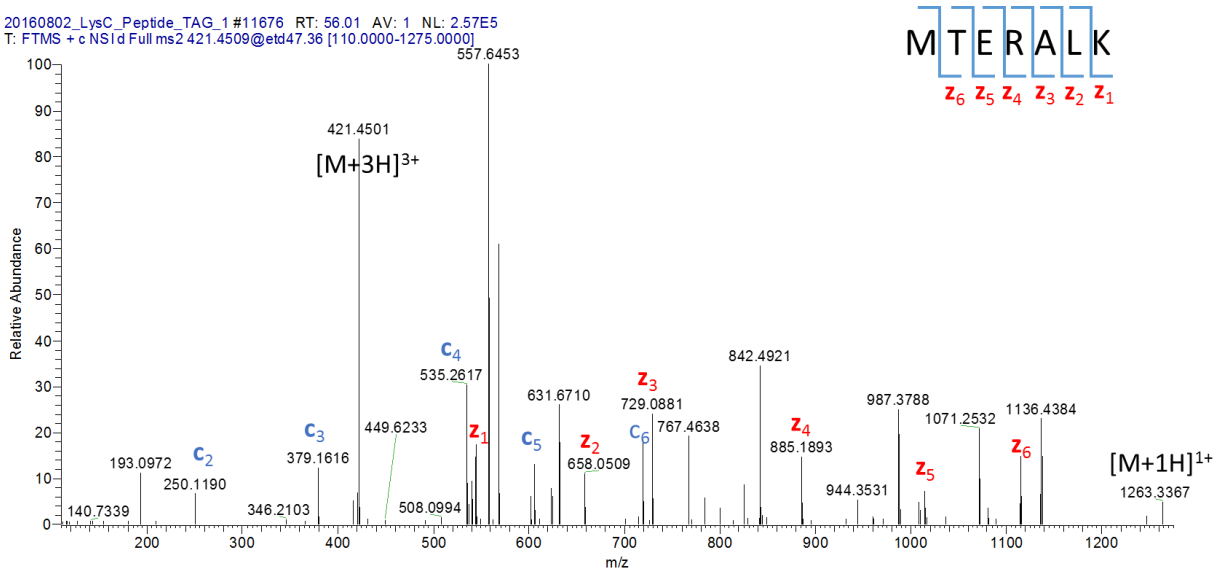


Figure 5.140. ETD-MS2 of triply charged "unidentified peptide 5" from Lys-C digest (**SI Figure 5.83**). Identified by de novo as MTERALK.

```
*****************************************************************************
*                                                                           *
*  This program is free software; you can redistribute it and/or modify     *
*  it under the terms of the GNU Library or "Lesser" General Public         *
*  License (LGPL) as published by the Free Software Foundation;             *
*  either version 2 of the License, or (at your option) any later           *
*  version.                                                                 *
*                                                                           *
*****************************************************************************/
/****
     Date:  07/29/2016
  Purpose:  Given a list of MS1 peaks, see if any above threshold and if any selected for MS2
Copyright:  Jimmy Eng, Copyright

****/

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <time.h>
#include <math.h>
#include <vector>
#include "ramp.h"

#define SIZE_FILE       512

struct OptionsStruct
{
  double dTolMS1; // tolerance used for presence of MS1 peak above threshold
```

```
  double dInten;  // intensity threshold

  // these aren't command line options but I'm putting them here anyways
  double dMinMass;
  double dMaxMass;
  int iCount;  // this is count of MS1 input masses
} pOptions;

struct PeaksStruct
{
  double dMS1;     // MS1 m/z
  int bPresent;    // present above threshold?
  int bMS2;        // selected for MS2
  std::vector<int> vMS1;
  std::vector<int> vMS2;
} *pMasses;

void FIND_PEAKS(char *szXMLFile, struct OptionsStruct pOptions);
void SET_OPTION(char *arg, struct OptionsStruct *pOptions);
void USAGE(char *argv0, struct OptionsStruct pOptions);


int main(int argc, char **argv)
{
  int i;
  int iNumArg;
  int iStartArgc;
  char *arg;

  iStartArgc = 1;
  iNumArg = 1;
  pOptions.dTolMS1 = 20;
  pOptions.dInten = 1E5;

  arg = argv[iNumArg];

  // processing arguments
  while (iNumArg < argc)
  {
    if (arg[0] == '-')
      SET_OPTION(arg, &pOptions);
    else
      break;

    iStartArgc++;
    arg = argv[++iNumArg];
  }

  // no input files specified
  if (argc-iStartArgc < 1)
  {
    USAGE(argv[0], pOptions);
  }

  FIND_PEAKS(argv[iStartArgc], pOptions);
```

```
  return(0);

}


void USAGE(char *argv0, struct OptionsStruct pOptions)
{
  printf(" USAGE:\n");
  printf("        %s input.mzXML\n", argv0);
  printf("\n");

  exit(EXIT_FAILURE);
}


void SET_OPTION(char *arg, struct OptionsStruct *pOptions)
{
  switch (arg[1])
  {

    case 'i':
      if (strlen(arg)>1)
      {
        double dTol;
        sscanf(arg+2, "%lf", &(pOptions->dInten));
        if (pOptions->dInten < 0)
        {
          printf(" Error reading tolerance from command line option.\n");
          printf("    parameter:  %s\n\n", arg);
          printf("   read value:  %0.1E\n", pOptions->dInten);
          exit(1);
        }
      }
      else
      {
        printf(" Error - no mass tolerance specified with -m parameter.\n");
        exit(1);
      }
      break;
    case 'm':
      if (strlen(arg)>1)
      {
        double dTol;
        sscanf(arg+2, "%lf", &(pOptions->dTolMS1));
        if (pOptions->dTolMS1 < 0 || pOptions->dTolMS1 > 5000.0)
        {
          printf(" Error reading tolerance from command line option.\n");
          printf("    parameter:  %s\n\n", arg);
          printf("   read value:  %0.1f\n", pOptions->dTolMS1);
          exit(1);
        }
      }
      else
      {
```

```
        printf(" Error - no mass tolerance specified with -m parameter.\n");
        exit(1);
      }
      break;
    default:
      break;
  }

  arg[0] = '\0';
}


void FIND_PEAKS(char *szXMLFile, struct OptionsStruct pOptions)
{
  int  i,
       iAnalysisFirstScan,
       iAnalysisLastScan;
  long  ctScan = 0;

  RAMPFILE  *pFI;

  ramp_fileoffset_t  indexOffset;
  ramp_fileoffset_t  *pScanIndex;

  if ( (pFI = rampOpenFile( szXMLFile )) == NULL)
  {
    printf( "could not open input file %s\n", szXMLFile);
    exit(1);
  }

  // Read the offset of the index
  indexOffset = getIndexOffset( pFI );

  // Read the scan index into a vector, get LastScan
  pScanIndex = readIndex( pFI , indexOffset, &iAnalysisLastScan );
  iAnalysisFirstScan = 1;

#define DMASS 544.9667

  for (ctScan=iAnalysisFirstScan; ctScan<=iAnalysisLastScan; ctScan++)
  {
    char  szInput[500];
    RAMPREAL *pPeaks;
    struct ScanHeaderStruct scanHeader;

    szInput[0]='\0';

    // read scan header
    readHeader(pFI, pScanIndex[ctScan], &scanHeader);

    printf("%d\t%d\t", scanHeader.msLevel, ctScan);

    if (scanHeader.peaksCount>0 && scanHeader.msLevel==2)
    {
      int iloop;
```

```
      int n = 0;
      int iCount = 0;
      // read scan peaks
      pPeaks = readPeaks( pFI, pScanIndex[ctScan]);
      double dMaxInten = 0.0;
      for (iloop=scanHeader.peaksCount; iloop-->0; )
      {
        double dMass=pPeaks[n++];
        double dInten=pPeaks[n++];

        if (fabs(1E6*(dMass - DMASS)/DMASS) <= pOptions.dTolMS1)
          if (dInten > dMaxInten)
            dMaxInten = dInten;
      }
      free(pPeaks);

      if (dMaxInten != 0.0)
        printf("%0.2E\n", dMaxInten);
      else
        printf("0.0\n");
    }
    else
    {
      printf("0.0\n");
    }
  }
  rampCloseFile(pFI);
  // print results
  for (i=0; i<pOptions.iCount; i++)
  {
    // original input masses
    printf("%f\t", pMasses[i].dMS1);

    // matched MS1 scans
    for(int ii=0; ii<pMasses[i].vMS1.size(); ii++)
    {
      if (ii!=0)
        printf(",");
      printf("%d", pMasses[i].vMS1[ii]);
    }
    printf("\t");
    // matched MS2 scans
    for(int ii=0; ii<pMasses[i].vMS2.size(); ii++)
    {
      if (ii!=0)
        printf(",");
      printf("%d", pMasses[i].vMS2[ii]);
    }
    printf("\n");
  }
}
```

Scheme 5.1. z1 ion (m/z 544.9667) Extraction Script

Complete Bibliography

1. Hoffman, W.D. & Jackson, G.P. Forensic mass spectrometry. *Annu. Rev. Anal. Chem*. **8,** 419–440 (2015).
2. Anacleto, J. F., Ramaley, L., Benoit, F. M., Boyd, R. K. & Quilliam, M. A. Comparison of liquid chromatography mass spectrometry interfaces for the analysis of polycyclic aromatic compounds. *Anal. Chem.* **67,** 4145–4154 (1995).
3. Titato, G.M. & Lanças, F. M. Optimization and validation of HPLC–UV–DAD and HPLC_APCI_MS methodologies for determination of selected PAHs in water samples. *J. Chromatogr. Sci.* **44,** 35–40 (2006).
4. Hernández, F., Sancho, J. V., Ibáñez, M., Abad, E., Portolés, T. & Mattioli, L. Curremt use of high-resolution mass spectrometry in the environmental sciences. *Anal. Bioanal. Chem.* **403,** 1251–1264 (2012).
5. Gelb, M. H., Scott, C. R. & Tureček, F. Newborn screening for lysosomal storage diseases. *Clin. Chem.* **61,** 335–346 (2015).
6. Baker, M. Mass spectrometry for biologists. *Nat. Methods* **7,** 157–161 (2010).
7. Becker, J. S. & Jakubowski, N. The synergy of elemental and biomolecular mass spectrometry: new analytical strategies in life sciences. *Chem. Soc. Rev.* **38,** 1969–83 (2009).
8. Yates III, J. R. A century of mass spectrometry: from atoms to proteomes. *Nat. Methods* **8,** (2011).
9. Griffiths, J. A Brief History of Mass Spectrometry. *Anal. Chem.* **80,** 5678–5683 (2008).
10. Münzenberg, G. Development of mass spectrometers from Thomson and Aston to present. *Int. J. Mass Spectrom.* **349,** 9–18 (2013).
11. Squires, G. Francis Aston and the mass spectrograph. *J. Chem. Soc. Dalt. Trans.* **23,** 3893–3899 (1998).
12. Nier, A. O. The development of a high resolution mass spectrometer: a reminiscence. *J. Am. Soc. Mass Spectrom.* **2,** 447–452 (1991).
13. Beynon, J. Mass spectrometry and its application to organic chemistry. 1st ed. Elsevier Pub. Co. Amsterdam: New York. (**1960**).

14. Djerassi, C. & Sample, S. D. Mass spectrometric fragmentation of nitrophenylhydrazones. *Nature* **208,** 1314 (1965).

15. Djerassi, C. & Fenselau, C. Mass spectrometry in structural and stereochemical problems. LXXXVI.1. The hydrogen-transfer reactions in butyl propionate, benzoate and phthalate 2,3. *J. Am. Chem. Soc.* **87,** 5756–5762 (1965).

16. Budzikiewicz, H., Djerassi, C. & Williams, D. H. Mass Spectrometry of Organic Compounds. 1st ed. Holden-Day: San Francisco. (**1967**).

17. McLuckey, S. A. & Goeringer, D. E. Slow heating methods in tandem mass spectrometry. *J. Mass Spectrom.* **32,** 461–474 (1997).

18. Hunt, D. F., Hignite, C. E. & Biemann, K. Structure elucidation of dinucleotides by mass spectrometry. *Biochem. Biophys. Res. Commun.* 33, 378–383 (1968).

19. Biemann, K. & Scoble, H. A. Characterization by tandem mass spectrometry of structural modifications in proteins. *Science.* **237,** 992–998 (1987).

20. Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Bio. Mass Spectrom.* **16,** 99–111 (1988).

21. Scoble, H. A., Biller, J. E. & Biemann, K. A graphics display-oriented strategy for the amino acid sequencing of peptides by tandem mass spectrometry. *Fresenius' Zeitschrift für Anal. Chemie* **327,** 239–245 (1987)

22. Hertz, H. S., Hites, R. A. & Biemann, K. Identification of mass spectra by computer-searching a file of known spectra. *Anal. Chem.* **43,** 681–691 (1971).

23. Field, F. H. Chemical ionization mass spectrometry. *Acc. Chem. Res.* **1,** 42–49 (1968).

24. Munson, M, S, B., and F. F. H. Physical and Inorganic Chemistry. *J. Am. Chem. Soc.* **88,** 2621–2630 (1966).

25. Biemann, K., Cone, C., Webster, B. R. & Arsenault, G. P. Determination of the amino acid sequence in oligopeptides by computer interpretation of their high-resolution mass spectra. *J. Am. Chem. Soc.* **88,** 5598–5606 (1966).

26. Gross, J.H. Mass spectrometry: A textbook. 2nd ed. Springer. Verlag Berlin Heidelberg (**2011**).

27. Gohlke, R. S. Time-of-flight mass spectrometry and gas-liquid partition chromatography. *Anal. Chem.* **31,** 535–541 (1959).

28. Arsenault, G. P., Dolhun, J. J. & Biemann, K. Gas chromatography–chemical ionization mass spectrometry. *J. Chem. Soc. D Chem. Commun.* **22,** 1542b–1543 (1970).

29. Arsenault, G. P., Dolhun, J. J. & Biemann, K. Alternate or simultaneous electron impact-chemical ionization mass spectrometry of gas chromatographic effluent. *Anal. Chem.* **43,** 1720–1722 (1971).

30. Yost, R. A. & Enke, C. G. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Anal. Chem.* **51,** 1251–1264 (1979).

31. Yost, R. A. & Enke, C. G. Selected ion fragmentation with a tandem quadrupole mass spectrometer. *J. Am. Chem. Soc.* **100,** 2274–2275 (1978).

32. Perchalski, R. J., Yost, R. A. & Wilder, B. J. Structural elucidation of drug metabolites by triple-quadrupole mass spectrometry. *Anal. Chem.* **54,** 1466–1471 (1982).

33. Hunt, D. F., Shabanowitz, J. & Giordani, A. B. Collision activated decompositions in mixture analysis with a triple quadrupole mass spectrometer. *Anal. Chem.* **52,** 386–390 (1980).

34. Sauter, A. D., Betowski, L. D. & Ballard, J. M. Comparison of priority pollutant response factors for triple and single quadrupole mass spectrometers. *Anal. Chem.* **55,** 116–119 (1983).

35. Hunt, D. F., Shabanowitz, J., Michael Harvey, T. & Coates, M. L. Analysis of organics in the environment by functional group using a triple quadrupole mass spectrometer. *J. Chromatogr. A* **271,** 93–105 (1983).

36. Barber, M. et al. Fast-atom-bombardment mass spectra of enkephalins. *Biochem. J.* **197,** 401–404 (1981).

37. Fenn, J. B. Electrospray ionization mass spectrometry: How it all began. *J. Biomol. Tech.* **13,** 101–118 (2002).

38. Morris, H. R. et al. Fast atom bombardment: A new mass spectrometric method for peptide sequence analysis. *Biochem. Biophys. Res. Commun.* **101,** 623–631 (1981).

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

40. 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.* **2,** 151–153 (1988).

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

42. Koster, C., Castoro, J. A. & Wilkins, C. L. High-resolution matrix-assisted laser desorption/ionization of biomolecules by Fourier transform mass spectrometry. *J. Am. Chem. Soc.* **114,** 7572–7574 (1992).

43. Lewis, J. K., Wei, J. & Siuzdak, G. Matrix-assisted laser desorption/ionization mass spectrometry in peptide and protein analysis. Encyclopedia of Analytical Chemistry. R.A. Meyers ed. pp5880–5894. (**2000**)

44. Siuzdak, G. The emergence of mass spectrometry in biochemical research. *Proc. Natl. Acad. Sci.* **91,** 11290–11297 (1994).

45. Fenn, J. B. Electrospray: wings for molecular elephants (Nobel lecture). *Angew*. *Chem. Int. Ed.* **42,** 3871–3894 (2003).

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

47. Yamashita, M. & Fenn, J. B. Negative ion production with the electrospray ion source. *J. Phys. Chem.* **88,** 4671–4675 (1984).

48. Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88,** 4451–4459 (1984).

49. Whitehouse, C. M., Dreyer, R. N., Yamashita, M. & Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57,** 675–679 (1985).

50. James, P. Protein identification in the post-genome era: the rapid rise of proteomics. *Q. Rev. Biophys.* **30,** 279–331 (1997).

51. Persidis, A. Proteomics. *Nat. Biotechnol.* **16,** 393–394 (1998).

52. Anderson, N. L. & Anderson, N. G. Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis* **19,** 1853–1861 (1998).

53. Packer, N. H. & Harrison, M. J. Glycobiology and proteomics: Is mass spectrometry the Holy Grail? *Electrophoresis*. **19,** 1872–1882 (1998).

54. De Hoffmann, E. & Stroobant, V. Mass spectrometry : principles and applications. 3rd ed. John Wiley & Sons Ltd. The atrium, Southern Gate, Chichester, West Sussex, England. (2007).

55. Mitchell Wells, J. & McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* **402,** 148–185 (2005).

56. Syrstad, E. A. & Tureček, F. Toward a general mechanism of electron capture dissociation. *J. Am. Soc. Mass Spectrom.* **16,** 208–224 (2005).

57. Gunawardena, H. P., He, M., Chrisman, P. A., Pitteri, S. J., Hogan, J. M., Hodges, B. D. M., & McLuckey, S. A. Electron transfer versus proton transfer in gas-phase ion/ion reactions of polyprotonated peptides. *J. Am. Chem. Soc.* **127,** 12627–12639 (2005).

58. Kjeldsen, F., Giessing, A. M. B., Ingrell, C. R. & Jensen, O. N. Peptide sequencing and characterization of post-translational modifications by enhanced ion-charging and liquid chromatography electron-transfer dissociation tandem mass spectrometry. *Anal. Chem.* **79,** 9243–52 (2007).

59. Crizer, D. M. & Mcluckey, S. A. Electron transfer dissociation of amide nitrogen methylated polypeptide cations. *J. Am. Soc. Mass Spectrom.* **20,** 1349–1354 (2009).

60. Jensen, C.S., Wyer, J. A., Houmøller, J., Hvelplund, P. & Brønsted, S. N. Electron-capture induced dissociation of doubly charged dipeptides: on the neutral losses and N–Cα bond cleavages. *Phys. Chem. Chem. Phys.* **13,** 18373 (2011).

61. Zubarev, R. A., Kelleher, N. L. & Mclafferty, F. W. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **120,** 3265–3266 (1998).

62. Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kellerher, N. L., Kruger, N. A., Lewis, M. A., Carpernter, B. K., and McLafferty, F. W. Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* **72,** 563–573 (2000).

63. Zubarev, R. A. Reactions of polypeptide ions with electrons in the gas phase. *Mass Spectrom. Rev.* **22,** 57–77 (2003).

64. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci.* **101,** 9528–9533 (2004).

65. Pitteri, S. J., Chrisman, P. A., Hogan, J. M. & McLuckey, S. A. Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: Reactions of doubly and triply protonated peptides with $SO_2^{\bullet-}$. *Anal. Chem.* **77,** 1831–1839 (2005).

66. Brodbelt, J. S. & Wilson, J. J. Infrared multiphoton dissociation in quadrupole ion traps. *Mass Spectrom. Rev.* **28,** 390–424 (2009).

67. Bensimon, M., Rapin J. & Gaunmann, T. Comparison of infrared photodissociation in a Fourier transform mass spectrometer with metastable ion decay in a double-focusing mass spectrometer. *Int. J. Mass Spectrom. Ion Process*. **72,** 125–135 (1986).

68. Watson, C. H., Baykut, G., Mowafy, Z., Katritzky, A. R. & Eyler, J. R. Photodissociation and collisionally activated dissociation studies of N-alkylpyridinium cations formed by laser desorption. *Anal. Instrum.* **17,** 155–172 (1988).

69. Oomens, J., Tielens, A. G. G. M., Sartakov, B. G., von Helden, G. & Meijer, G. Laboratory infrared spectroscopy of cationic polycyclic aromatic hydrocarbon molecules. *Astrophys. J.* **591,** 968–985 (2003).

70. Polfer, N. C. Infrared multiple photon dissociation spectroscopy of trapped ions. *Chem. Soc. Rev.* **40**, 2211–2221 (2001).

71. Holden, D. D., Pruet, J. M. & Brodbelt, J. S. Ultraviolet photodissociation of protonated, fixed charge, and charge-reduced peptides. *Int. J. Mass Spectrom.* **390,** 81–90 (2015).

72. Shaffer, C. J., Slováková, K. & Tureček, F. Near-UV photodissociation of phosphopeptide cation-radicals. *Int. J. Mass Spectrom.* **390,** 71–80 (2015).

73. Nguyen, H. T. H., Shaffer, C. J. & Tureček, F. Probing peptide cation-radicals by near-UV photodissociation in the gas phase. Structure elucidation of histidine radical chromophores formed by electron transfer reduction. *J. Phys. Chem. B* **119,** 3948–61 (2015).

74. Viglino, E., Lai, C. K., Mu, X., Chu, I. K. & Tureček, F. Ground and excited-electronic-state dissociations of hydrogen-rich and hydrogen-deficient tyrosine peptide cation radicals. *J. Am. Soc. Mass Spectrom.* **27,** 1454–1467 (2016).

75. Wilson, J. J., Kirkovits, G. J., Sessler, J. L. & Brodbelt, J. S. Photodissociation of non-covalent peptide-crown ether complexes. *J. Am. Soc. Mass Spectrom.* **19,** 257–60 (2008).

76. Oh, H. Bin & Moon, B. Radical-driven peptide backbone dissociation tandem mass spectrometry. *Mass Spectrom. Rev.* **34,** 116–32 (2015).

77. Halim, M. A., Girod, M., MacAleese, L., Lemoine, J., Antoine, R., and Dugourd, P. 213 nm Ultraviolet photodissociation on peptide anions: Radical-directed fragmentation patterns. *J. Am. Soc. Mass Spectrom. Soc.* **27,** 474–486 (2016).

78. Tureček, F. Benchmarking electronic excitation energies and transitions in peptide radicals. *J. Phys. Chem. A* **119,** 10101–10111 (2015).

79. Ly, T. & Julian, R. R. Residue-specific radical-directed dissociation of whole proteins in the gas phase. *J. Am. Chem. Soc.* **130,** 351–358 (2008).

80. Viglino, E., Shaffer, C. J. & Tureček, F. UV/Vis Action spectroscopy and structures of tyrosine peptide cation radicals in the gas phase. *Angew. Chemie Int. Ed.* **55,** 7469–7473 (2016).

81. Nguyen, H. T. H., Shaffer, C. J., Pepin, R. & Tureček, F. UV Action spectroscopy of gas-phase peptide radicals. *J. Phys. Chem. Lett.* **6,** 4722–7 (2015).

82. Shaffer, C. J., Pepin, R. & Tureček, F. Combining UV photodissociation action spectroscopy with electron transfer dissociation for structure analysis of gas-phase peptide cation-radicals. *J. Mass Spectrom.* **50,** 1438–42 (2015).

83. Sleno, L. & Volmer, D. A. Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **39,** 1091–112 (2004).

84. Lin, C., Cournoyer, J. J. & O'Connor, P. B. Use of a double resonance electron capture dissociation experiment to probe fragment intermediate lifetimes. *J. Am. Soc. Mass Spectrom.* **17,** 1605–1615 (2006).

85. Li, X., Cournoyer, J. J., Lin, C. & O'Connor, P. B. The effect of fixed charge modifications on electron capture dissociation. *J. Am. Soc. Mass Spectrom.* **19,** 1514–26 (2008).

86. Perry, R. H., Cooks, R. G. & Noll, R. J. Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrom. Rev.* **27,** 661–699 (2008).

87. Chiva, C. & Sabidó, E. HCD-only fragmentation method balances peptide identification and quantitation of TMT-labeled samples in hybrid linear ion trap/orbitrap mass spectrometers. *J. Proteomics* **96,** 263–270 (2014).

88. Köcher, T., Pichler, P., Schutzbier, M., Stingl, C., Kaul, A., Teucher, N., Hasenfuss, G., Penninger, J. M., & Mechtler, K. High precision quantitative proteomics using iTRAQ on an LTQ Orbitrap: A new mass spectrometric method combining the benefits of all. *J. Proteome Res.* **8,** 4743–4752 (2009).

89. Dayon, L., Pasquarello, C., Hoogland, C., Sanchez, J.C. & Scherl, A. Combining low- and high-energy tandem mass spectra for optimized peptide quantification with isobaric tags. *J. Proteomics* **73,** 769–777 (2010).

90. Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. R. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc.* **11,** 993–1006 (2016).

91. Zhang, D., Liu, H., Zhang, S., Chen, X., Li, S., Zhang, C., Hu, X., Bi, K., Chen, X., & Jiang, Y. An effective method for de novo peptide sequencing based on phosphorylation strategy and mass spectrometry. *Talanta* **84,** 614–622 (2011).

92. Brancia, F. L., Oliver, S. G. & Gaskell, S. J. Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun. Mass Spectrom.* **14,** 2070–2073 (2000).

93. Beardsley, R. L. & Reilly, J. P. Optimization of guanidination procedures for MALDI mass mapping. *Anal. Chem.* **74,** 1884–1890 (2002).

94. Warwood, S., Mohammed, S., Cristea, I. M., Evans, C., Whetton, A. D., & Gaskell, Simon. Guanidination chemistry for qualitative and quantitative proteomics. *Rapid Commun. Mass Spectrom.* **20,** 3245–3256 (2006).

95. Shen, Y., Tolić, N., Xie, F., Zhao, R., Purvine, S. O., Schepmoes, A. A., Moore, R. J., Anderson, G. A. & Smith, R. D. Effectiveness of CID, HCD, and ETD with FT MS/MS for degradomic-peptidomic analysis: Comparison of peptide identification methods. *J. Proteome Res.* **10,** 3929–3943 (2011).

96. Zee, B. M. & Garcia, B. A. Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem.* **52,** 147–63 (2012).

97. Shi, Y., Bajrami, B. & Yao, X. Passive and active fragment ion mass defect labeling: Distinct proteomics potential of iodine-based reagents. *Anal. Chem.* **81,** 6438–6448 (2009).

98. Hennrich, M. L., Boersema, P. J., van den Toorn, H., Mischerikow, N., Heck, A. J. R. & Mohammed, S. Effect of chemical modifications on peptide fragmentation behavior upon electron transfer induced dissociation. *Anal. Chem.* **81,** 7814–7822 (2009).

99. Hall, M. P., Ashrafi, S. Obegi, I., Petesch, R., Peterson, J. N. & Schneider, L. V. 'Mass defect' tags for biomolecular mass spectrometry. *J. Mass Spectrom.* **38,** 809–816 (2003).

100. Bajrami, B., Shi, Y., Lapierre, P. & Yao, X. Shifting unoccupied spectral space in mass spectrum of peptide fragment ions. *J. Am. Soc. Mass Spectrom.* **20,** 2124–2134 (2009).

101. Marek, A., Pepin, R., Peng, B., Laszlo, K. J., Bush, M. F. & Tureček, F. Electron transfer dissociation of photolabeled peptides. Backbone cleavages compete with diazirine ring rearrangements. *J. Am. Soc. Mass Spectrom.* **24,** 1641–1653 (2013).

102. Shaffer, C. J., Marek, A., Pepin, R., Slováková, K. & Turecek, F. Combining UV photodissociation with electron transfer for peptide structure analysis. *J. Mass Spectrom.* **50,** 470–475 (2015).

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

104. Tureček, F. & Julian, R. R. Peptide radicals and cation radicals in the gas phase. *Chem. Rev.* **113,** 6691–6733 (2013).

105. Chu, I. K., Siu, C-K., Lau, J. K-C., Tang, W. K., Mu, X., Lai, C. K., Guo, X., Wang, X., Li, N., Xia, Y., Kong, X., Oh, H. B., Ryzhov, V., Tureček, F., Hopkinson, A. C. & Siu, K. W. M. Proposed nomenclature for peptide ion fragmentation. *Int. J. Mass Spectrom.* **390,** 24–27 (2015).

106. McLafferty, F. W. & Tureček, F. Interpretation of mass spectra. 4th ed. Univerity Science Books. Sausalito, California. 369pp. (**1993**).

107. Chu, I. K., Zhao, J., Xu, M., Siu, S. O., Hopkinson, A. C. & Siu, K. W. M. Are the radical centers in peptide radical cations mobile? The generation, tautomerism, and dissociation of isomeric α-

carbon-centered triglycine radical cations in the gas phase. *J. Am. Chem. Soc.* **130,** 7862–7872 (2008).

108. Madsen, J. A., Cheng, R. R., Kaoud, T. S., Dalby, K. N., Dmitrii, E. M. & Brodbelt, J. S. Charge-site-dependent dissociation of hydrogen-rich radical peptide cations upon vacuum UV photoexcitation. *Chemistry* **18,** 5374–83 (2012).

109. Han, H., Xia, Y. & McLuckey, S. A. Ion trap collisional activation of c and z• ions formed via gas-phase ion/ion electron-transfer dissociation. *J. Proteome Res.* **6,** 3062–9 (2007).

110. Zhao, J., Song, T., Xu, M., Quan, Q., Siu, K. W. M., Hopkinson, A. C. & Chu, I. K. Intramolecular hydrogen atom migration along the backbone of cationic and neutral radical tripeptides and subsequent radical-induced dissociations. *Phys. Chem. Chem. Phys.* **14,** 8723 (2012).

111. Ly, T., Yin, S., Loo, J. A. & Julian, R. R. Electron-induced dissociation of protonated peptides yields backbone fragmentation consistent with a hydrogen-deficient radical. *Rapid Commun. Mass Spectrom.* **23,** 2099–2101 (2009).

112. Ly, T. & Julian, R. R. Ultraviolet photodissociation: Developments towards applications for mass-spectrometry-based proteomics. *Angew. Chemie Int. Ed.* **48,** 7130–7137 (2009).

113. Hodyss, R., Cox, H. A., & Beauchamp, J. L. Bioconjugates for tunable peptide fragmentation: Free radical initiated peptide sequencing (FRIPS). *J. Am. Chem. Soc.* **127,** 12436–12437 (2005).

114. Laskin, J., Yang, Z., Ng, C. M. D. & Chu, I. K. Fragmentatio of α-radical cations of arginine-containing peptides. *J. Am. Soc. Mass Spectrom.* **21,** 511–521 (2010).

115. Barlow, C. K., Wee, S., McFayden, W. D. & O'Hair, A. J. Designing copper(II) ternary complexes to generate radical cations of peptides in the gas phase: Role of the auxiliary ligand. *Dalt. Trans.* **101,** 3199 (2004).

116. Chu, I. K., Rodriquez, C. F., Lau, T.-C., Hopkinson, A. C. & Siu, K. W. M. Molecular Radical Cations of Oligopeptides. *J. Phys. Chem. B* **104,** 3393–3397 (2000).

117. Chu, I. K., Rodriquez, C. F., Hopkinson, A. C., Siu, K. W. & Lau, T. C. Formation of molecular radical cations of enkephalin derivatives via collision-induced dissociation of electrospray-generated copper (II) complex ions of amines and peptides. *J. Am. Soc. Mass Spectrom.* **12,** 1114–9 (2001).

118. Chu, I. K. & Lam, C. N. W. Generation of peptide radical dications via low-energy collision-induced dissociation of [CuII(terpy)(M + H)]$^{\bullet 3+}$. *J. Am. Soc. Mass Spectrom.* **16,** 1795–1804 (2005).

119. Tureček, F. Copper-biomolecule complexes in the gas phase. The ternary way. *Mass Spectrom. Rev.* **26,** 563–582 (2007).

120. Siu, C-K., Ke-YuYong, Guo, Y., Hopkinson, A. C. & Siu, K. W. M. Dissociations of copper(II)-containing complexes of aromatic amino acids: radical cations of tryptophan, tyrosine, and phenylalanine. *Phys. Chem. Chem. Phys.* **10,** 5908 (2008).

121. Hopkinson, A. C. Radical cations of amino acids and peptides: Structures and stabilities. *Mass Spectrom. Rev.* **28,** 655–671 (2009).

122. Yu, L., Tan, Y., Tsai, Y., Goodlett, D. R. & Polfer, N. C. On the relevance of peptide sequence permutations in shotgun proteomics studies. *J. Proteome Res.* **10,** 2409–2416 (2011).

123. Tureček, F. Renaissance of cation-radicals in mass spectrometry. *Mass Spectrom. (Tokyo, Japan)* **2,** S0003 (2013).

124. Barlow, C. K., Wright, A., Easton, C. J. & O'Hair, A. J. Gas-phase ion-molecule reactions using regioselectively generated radical cations to model oxidative damage and probe radical sites in peptides. *Org. Biomol. Chem.* **9,** 3733 (2011).

125. Piatkivskyi, A., Osburn, S., Jaderberg, K., Grzetic, J., Steill, J. D., Oomens, J., Zhao, J., Lau, J, K-C., Verkerk, U. H., Hopkinson, A. C., Siu, K. W. M. & Ryzhov, V. Structure and reactivity of the distonic and aromatic radical cations of tryptophan. *J. Am. Soc. Mass Spectrom.* **24,** 513–23 (2013).

126. Wang, E. H., Combe, P. C. & Schug, K. A. Multiple reaction monitoring for direct quantitation of intact proteins using a triple quadrupole mass spectrometer. *J. Am. Soc. Mass Spectrom.* **27,** 886–896 (2016).

127. Moskovets, E. V. & Ivanov, A. R. Comparative studies of peak intensities and chromatographic separation of proteolytic digests, PTMs, and intact proteins obtained by nanoLC-ESI MS analysis at room and elevated temperatures. *Anal. Bioanal. Chem.* **408,** 3953–3968 (2016).

128. Fang, H., Xiao, K., Li, Y., Yu, F., Liu, Y., Xue, B. & Tian, Z. Intact protein quantitation using pseudoisobaric dimethyl labeling. *Anal. Chem.* **88,** 7198–7205 (2016).

129. Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E. P., Shabanowitz, J. & Hunt, D. F. The utility of ETD mass spectrometry in proteomic analysis. *Biochim. Biophys. Acta - Proteins Proteomics* **1764,** 1811–1822 (2006).

130. Zimnicka, M., Moss, C. L., Chung, T. W., Hui, R. & Tureček, F. Tunable charge tags for electron-based methods of peptide sequencing: Design and applications. *J. Am. Soc. Mass Spectrom.* **23,** 608–620 (2012).

131. Miyashita, M., Hanai, Y., Awane, H., Yoshikawa, T. & Miyagawa, H. Improving peptide fragmentation by N-terminal derivatization with high proton affinity. *Rapid Commun. Mass Spectrom.* **25,** 1130–1140 (2011).

132. Frison, G., van der Rest, G., Tureček, F., Besson, T., Lemaine, J., Maître, P. & Chamot-Rooke, J. Structure of electron-capture dissociation fragments from charge-tagged peptides probed by tunable infrared multiple photon dissociation. *J. Am. Chem. Soc.* **130,** 14916–14917 (2008).

133. Roth, K. D. W., Huang, Z.-H., Sadagopan, N. & Watson, J. T. Charge derivatization of peptides for analysis by mass spectrometry. *Mass Spectrom. Rev.* **17,** 255–274 (1998).

134. Vath, J. E. & Biemann, K. Microderivatization of peptides by placing a fixed positive charge at the N-terminus to modify high energy collision fragmentation. *Int. J. Mass Spectrom. Ion Process.* **100,** 287–299 (1990).

135. Chung, T. W., Moss, C. L., Zimnicka, M., Johnson, R. S., Moritz, R. L. & Tureček. Electron-capture and -transfer dissociation of peptides tagged with tunable fixed-charge groups: Structures and dissociation energetics. *J. Am. Soc. Mass Spectrom.* **22,** 13–30 (2011).

136. Good, D. M., Wirtala, M., McAlister, G. C. & Coon, J. J. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **6,** 1942–51 (2007).

137. Harms, M. J., Schlessman, J. L., Sue, G. R. & García-Moreno, B. Arginine residues at internal positions in a protein are always charged. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 18954–9 (2011).

138. Thurman, E. M. & Ferrer, I. The isotopic mass defect: A tool for limiting molecular formulas by accurate mass. in *Analytical and Bioanalytical Chemistry* **397,** 2807–2816 (2010).

139. Hall, M. P. & Schneider, L. V. Isotope-differentiated binding energy shift tags (IDBEST™) for improved targeted biomarker discovery and validation. *Expert Rev. Proteomics* **1,** 421–431 (2004).

140. Sleno, L. The use of mass defect in modern mass spectrometry. *J. Mass Spectrom.* **47,** 226–236 (2012).

141. Hernandez, H., Niehauser, S., Boltz, S. A., Gawandi, V., Phillips, R. S. & Amster, I. J. Mass defect labeling of cysteine for improving peptide assignment in shotgun proteomic analyses. *Anal. Chem.* **78,** 3417–3423 (2006).

142. Toumi, M. L. & Desaire, H. Improving mass defect filters for human proteins. *J. Proteome Res.* **9,** 5492–5 (2010).

143. Nefedov, A. V., Mitra, I., Brasier, A. R. & Sadygov, R. G. Examining troughs in the mass distribution of all theoretically possible tryptic peptides. *J. Proteome Res.* **10,** 4150–4157 (2011).

144. Mann, M. Useful tables of possible and probable peptide masses. *Annual Conference on Mass Spectrometry and Allied Topics* May 1-11 (American Society of Mass Spectrometry, 1995).

145. Mitra, I., Nefedov, A. V., Brasier, A. R. & Sadygov, R. G. Improved mass defect model for theoretical tryptic peptides. *Anal. Chem.* **84,** 3026–3032 (2012).

146. Taguchi, V. Y., Nieckarz, R. J., Clement, R. E., Krolik, S. & Williams, R. Dioxin analysis by gas chromatography-Fourier transform ion cyclotron resonance mass spectrometry (GC-FTICRMS). *J. Am. Soc. Mass Spectrom.* **21,** 1918–1921 (2010).

147. Sleighter, R. L. & Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **42,** 559–574 (2007).

148. Zhang, H., Zhu, M., Ray, K. L., Ma, L. & Zhang, D. Mass defect profiles of biological matrices and the general applicability of mass defect filtering for metabolite detection. *Rapid Commun. Mass Spectrom.* **22,** 2082–2088 (2008).

149. LeBlanc, A., Shiao, T. C., Roy, R. & Sleno, L. Improved detection of reactive metabolites with a bromine-containing glutathione analog using mass defect and isotope pattern matching. *Rapid Commun. Mass Spectrom.* **24,** 1241–1250 (2010).

150. Hughes, C., Ma, B. & Lajoie, G. A. De novo sequencing methods in proteomics. *Methods Mol. Biol.* **604,** 105–21 (2010).

151. Lu, M., Wang, H., Wang, Z., Li, X.-F. & Le, X. C. Identification of reactive cysteines in a protein using arsenic labeling and collision-induced dissociation tandem mass spectrometry. *J. Proteome Res.* **7,** 3080–3090 (2008).

152. Lu, M., Li, X.-F., Le, X. C., Weinfeld, M. & Wang, H. Identification and characterization of cysteinyl exposure in proteins by selective mercury labeling and nano-electrospray ionization quadrupole time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **24,** 1523–1532 (2010).

153. Whetstone,P. A., Butlin, N. G., Corneillie, T. M. & Meares, C. F. Element-coded affinity tags for peptides and proteins. *Bioconjug. Chem.* **15,** 3–6 (2004).

154. Lee, S., Young, N. L., Whetstone, P. A., Cheal, S. M., Benner, W. H., Lebrilla, C. B. & Meares, C. F. Method to site-specifically identify and quantitate carbonyl end products of protein oxidation using oxidation-dependent element coded affinity tags (O-ECAT) and nanoLiquid chromatography Fourier transform mass spectrometry. *J. Proteome Res.* **5,** 539-47 (2006).

155. Petrotchenko, E. V. & Borchers, C. H. Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom. Rev.* **29,** 862–876 (2010).

156. Yao, X., Diego, P., Ramos, A. A. & Shi, Y. Averagine-scaling analysis and fragment ion mass defect labeling in peptide mass spectrometry. *Anal. Chem.* **80,** 7383–7391 (2008).

157.

158. Kim, S., Bandeira, N. & Pevzner, P. A. Spectral profiles, a novel representation of tandem mass spectra and their applications for *de novo* peptide sequencing and identification. *Mol. Cell. Proteomics* **8,** 1391–1400 (2009).

159. Medzihradszky, K. F. & Chalkley, R. J. Lessons in *de novo* peptide sequencing by tandem mass spectrometry. *Mass Spectrometry Reviews* **34,** 43–63 (2015).

160. Frank, A. M., Savitski, M. M., Nielsen, M. L., Zubarev, R. A. & Pevzner, P. A. *De Novo* peptide sequencing and identification with precision mass spectrometry. *J. Proteome Res.* **6,** 114–123 (2007).

161. Kim, S., Gupta, N., Bandeira, N. & Pevzner, P. A. Spectral dictionaries: Integrating *de novo* peptide sequencing with database search of tandem mass spectra. *Mol. Cell. Proteomics* **8,** 53–69 (2009).

162. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A. & Lajoie, G. PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17,** 2337–2342 (2003).

163. Jeong, K., Kim, S. & Pevzner, P. A. UniNovo: a universal tool for *de novo* peptide sequencing. *Bioinformatics* **29,** 1953–1962 (2013).

164. Grotemeyer, J., Boesl, U., Walter, K. & Schlag, E. W. A general soft ionization method for mass spectrometry: Resonance-enhanced multi-photon ionization of biomolecules. *Org. Mass Spectrom.* **21,** 645–653 (1986).

165. Grotemeyer, J. & Schlag, E. W. Peptides investigated by laser desorption—multiphoton ionization mass spectrometry. *Org. Mass Spectrom.* **23,** 388–396 (1988).

166. Grotemeyer, J. & Schlag, E. W. Biomolecules in the gas phase: multiphoton ionization mass spectrometry. *Acc. Chem. Res.* **22,** 399–406 (1989).

167. Hopkinson, A. C. & Siu, K. W. M. Principles of mass spectrometry applied to biomolecules. 1st Ed. John Wiley & Sons, Inc. Hoboken, New Jersey. pp 301–335. (2006).

168. Chu, I. K. & Laskin, J. Formation of peptide radical ions through dissociative electron transfer in ternary metal-ligand-peptide complexes. *Eur. J. Mass Spectrom.* **17,** 543–56 (2011).

169. Chakraborty, T., Holm, A. I. S., Hvelplund, P., Brønstead, N., Poully, J. C., Worm, E. S. & Williams, E. R. On the survival of peptide cations after electron capture: Role of internal hydrogen bonding and microsolvation. *J. Am. Soc. Mass Spectrom.* **17,** 1675–1680 (2006).

170. Holm, A. I. S., Hvelplund, P., Kadhane, U., Larsen, M. K., Liu, B., Brønstead, S. N., Panja, S., Pedersen, J. M., Skrydstrup, T., Støchel, K., Williams, E. R. & Worm, E. S. On the mechanism of electron-capture-induced dissociation of peptide dications from ^{15}N-labeling and crown-ether complexation. *J. Phys. Chem. A* **111,** 9641–3 (2007).

171. Tureček, F. Yao, C., Fung, Y. M. E., Hayakawa, S., Hashimoto, M. & Matsubara, H. Histidine-containing radicals in the gas phase. *J. Phys. Chem. B* **113,** 7347–7366 (2009).

172. Hayakawa, S., Hashimoto, M., Matsubara, H. & Tureček, F. Dissecting the proline effect: Dissociations of proline radicals formed by electron transfer to protonated Pro-Gly and Gly-Pro dipeptides in the gas phase. J. Am. Chem. Soc. **129,** 7936–7949 (2007).

173. Byskov, C. S. & Nielsen, S. B. On the formation, stability, and dissociation of peptide radicals after femtosecond electron transfer from alkali metal atoms. *Int. J. Mass Spectrom.* **390,** 2–13 (2015).

174. Sun, Q., Nelson, H., Ly, T., Stoltz, B. M. & Julian, R. R. Side chain chemistry mediates backbone fragmentation in hydrogen deficient peptide radicals. *J. Proteome Res.* **8,** 958–966 (2008).

175. Chung, T. W. & Tureček, F. Backbone and side-chain specific dissociations of z ions from non-tryptic peptides. *J. Am. Soc. Mass Spectrom.* **21,** 1279–1295 (2010).

176. Moss, C. L., Chung, T. W., Wyer, J. A., Brønstead, S. N., Hvelplund, P. & Tureček, F. Dipole-guided electron capture causes abnormal dissociations of phosphorylated pentapeptides. *J. Am. Soc. Mass Spectrom.* **22,** 731–751 (2011).

177. Chung, T. W. & Tureček, F. Proper and improper aminoketyl radicals in electron-based peptide dissociations. *Int. J. Mass Spectrom.* **301,** 55–61 (2011).

178. Xu, M., Song, T., Quan, Q., Hao, Q., Fang, D-C., Siu, C-K. & Chu, I. K. Effect of the N-terminal basic residue on facile Cα–C bond cleavages of aromatic-containing peptide radical cations. *Phys. Chem. Chem. Phys.* **13,** 5888 (2011).

179. Ledvina, A.R., Chung, T. W., Hui, R., Coon, J. J. & Tureček, F. Cascade dissociations of peptide cation-radicals. Part 2. Infrared multiphoton dissociation and mechanistic studies of z-ions from pentapeptides. *J. Am. Soc. Mass Spectrom.* **23,** 1351–1363 (2012).

180. Hao, Q., Song, T., Ng, D. C. M., Quan, Q., Siu, C-K. & Chu, I. Arginine-facilitated isomerization: Radical-induced dissociation of aliphatic radical cationic glycylarginyl(iso)leucine tripeptides. *J. Phys. Chem. B* **116,** 7627–7634 (2012).

181. Zhang, X. & Julian, R. R. Exploring radical migration pathways in peptides with positional isomers, deuterium labeling, and molecular dynamics simulations. *J. Am. Soc. Mass Spectrom.* **24,** 524–533 (2013).

182. Thomas, D. A., Sohn, C. H., Gao, J. & Beauchamp, J. L. Hydrogen bonding constrains free radical reaction dynamics at serine and threonine residues in peptides. *J. Phys. Chem. A* **118,** 8380–8392 (2014). *Int. J. Mass. Spectrom.* **390, 56-62** (2015).

183. Xu, M., Tang, W-K., Mu, X., Ling, Y., Siu, C-K., Laskin, J. & Chu, I.K. α-Radical-induced CO_2 loss from the aspartic acid side chain of the collisional induced tripeptide aspartylglycylarginine radical cation.

184. Mu, X., Song, T., Xu, M., Lai, C-K., Siu, C-K., Laskin, J. & Chu, I. K. Discovery and mechanistic studies of facile N-terminal Cα–C bond cleavages in the dissociation of tyrosine-containing peptide radical cations. *J. Phys. Chem. B* **118,** 4273–4281(2014).

185. Frisch, M. J., Trucks, G. W., Shlegel, H.B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J. & Fox, G. J. Gaussian 09, Revision A.02. *Gaussian 09, Revision A.02* (2009).

186. Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **98,** 1372 (1993).

187. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06 functionals and 12 other functionals. *Theor. Chem. Acc.* **119,** 525–525 (2008).

188. Chai, J-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10,** 6615 (2008).

189. Møller, C. & Plesset, M. S. A note on an approximation treatment for many-electron systems. *Phys. Rev.* **46,** 618–622 (1934).

190. Ehlerding, A., Jensen, C. S., Wyer, J. A., Holm, A. I. S., Jørgensen, P., Kadhane, U., Larsen, M. K., Panja, S., Poully, J. C., Worm, E. S., Zettergren, H., Hvelplund, P. & Brønstead, S. N. Influence of temperature and crown ether complex formation on the charge partitioning between z and c fragments formed after electron capture by small peptide dications. *Int. J. Mass Spectrom.* **282,** 21–27 (2009).

191. Gerbaux, P. & Tureček, F. Protonated carbonic acid and the trihydroxymethyl Radical in the Gas Phase. A Neutralization-Reionization Mass Spectrometric and ab Initio/RRKM Study. *J. Phys. Chem. A* **106,** 5938–5950 (2002).

192. Hao, C., Seymour, J. L. & Tureček, F. Electron super-rich radicals in the gas phase. A neutralization-reionization mass spectrometric and ab Initio/RRKM study of diaminohydroxymethyl and triaminomethyl radicals. *J. Phys. Chem. A* **111,** 8829–8843 (2007).

193. Pepin, R., Ek, F. & Ek, T. Kinetic ion thermometers for electron transfer dissociation. *J. Phys. Chem. B* **119,** 2818–2826 (2015).

194. Julian, R. R. & Beauchamp, J. L. Site-specific sequestering and stabilization of charge in peptides by supramolecular adduct formation with 18-crown-6 ether by way of electrospray ionization. *Int. J. Mass. Spectrom.* **210-211,** 613–623 (2001).

195. Chen, Y. & Rodgers, M. T. Structural and energetic effects in the molecular recognition of amino acids by 18-Crown-6. *J. Am. Chem. Soc.***134,** 5863–5875 (2012).

196. Holm, A. I. S., Larsen, M. K., Panja, S., Hvelplund, P., Brønstead, S. N., Leib, R. D., Donald, W. A., Williams, E. R., Hao, C. & Tureček, F. Electron capture, femtosecond electron transfer and theory: A study of noncovalent crown ether 1,n-diammonium alkane complexes. *Int. J. Mass Spectrom.* **276,** 116–126 (2008).

197. Sobczyk, M., Anusiewicz, I., Berdys-Kochanska, J., Sawicka, A., Skurski, P. & Simons, J. Coulomb-assisted dissociative electron attachment: Application to a model peptide. *J. Phys. Chem. A* **109,** 250–258 (2005).

198. Yao, C. & Tureček, F. Hypervalent ammonium radicals. Competitive N–C and N–H bond dissociations in methyl ammonium and ethyl ammonium. *Phys. Chem. Chem. Phys.* **7,** 912–920 (2005).

199. Chen, X. & Tureček, F. The arginine anomaly: Arginine radicals are poor hydrogen atom donors in electron transfer induced dissociations. (2006).

200. Moss, C. L., Liang, W., Li, X. & Tureček, F. The early life of a peptide cation-radical. Ground and excited-state trajectories of electron-based peptide dissociations during the first 330 Femtoseconds. *J. Am. Soc. Mass Spectrom* **23,** 446–459 (2012).

201. Hayakawa, S., Matsubara, H., Panja, S., Hvelplund, P., Brønsted, S. N., Chen, X. & Tureček, F. Experimental evidence for an inverse hydrogen migration in arginine radicals. *J. Am. Chem. Soc.* **130,** 7645-7654 (2008).

202. Whittaker, M. M., DeVito, V. L., Asher, S. A. & Whittaker, J. W. Resonance Raman evidence for tyrosine involvement in the radical site of galactose oxidase. *J. Biol. Chem.* **264,** 7104–6 (1989).

203. Sjoberg, B. M. & Reichard, P. Nature of the free radical in ribonucleotide reductase from *Escherichia coli*. *J. Biol. Chem.* **252,** 536-543 (1977).

204. Voicescu, M., El Khoury, Y., Martel, D., Heinrich, M. & Hellwig, P. Spectroscopic analysis of tyrosine derivatives: On the role of the tyrosine-histidine covalent linkage in cytochrome C oxidase. *J. Phys. Chem. B* **113,** 13429–13436 (2009).

205. Moore, G. F., Hambourger, M., Gervaldo, M., Poluektov, O. G., Rajh, T., Gust, D., Moore, T. A. & Moore, A. L., A bioinspired construct that mimics the proton coupled electron transfer between P680 radical cation and the Tyrz-his190 pair of photosystem II. *J. Am. Chem. Soc.* **130,** 10466–10467 (2008).

206. Stubbe, J. & Donk, W. a Van Der. Protein Radicals in Enzyme Catalysis. *Chem. Rev.* **98,** 705–762 (1998).

207. Bernini, C., Arezzini, E., Basosi, R. & Sinicropi, A. In silico spectroscopy of tryptophan and tyrosine radicals involved in the long-range electron transfer of cytochrome c peroxidase. *J. Phys. Chem. B* **118,** 9525–9537 (2014).

208. Nara, S. J., Valgimigli, L., Pedulli, G. F. & Pratt, D. A. Tyrosine analogues for probing proton-coupled electron transfer processes in peptides and proteins. *J. Am. Chem. Soc.* **132,** 863–872 (2010).

209. Svistunenko, D. A. & Jones, G. A. Tyrosyl radicals in proteins: A comparison of empirical and density functional calculated EPR parameters. *Phys. Chem. Chem. Phys.* **11,** 6600–6613 (2009).

210. Ivancich, A., Mattioli, T. A. & Un, S. Effect of protein microenvironment on tyrosyl radicals. A high-field (285 GHz) EPR, resonance raman, and hybrid density functional study. *J. Am. Chem. Soc.* **121,** 5743–5753 (1999).

211. Faraggi, M., DeFelippis, M. R. & Klapper, M. H. Long-range electron transfer between tyrosine and tryptophan in peptides. *J. Am. Chem. Soc.* **111,** 5141–5145 (1989).

212. Turecek, F. Transient intermediates of chemical reactions by neutralization-reionization mass spectrometry. *Top. Curr. Chem.* **225,** 77–129 (2003).

213. Brunet, C., Antoine, R., Allouche, A-R., Dugourd, P., Canon, F., Giuliani, A. & Nahon, L. Gas phase photo-formation and vacuum UV photofragmentation spectroscopy of tryptophan and tyrosine radical-containing peptides. *J. Phys. Chem. A* **115,** 8933–8939 (2011).

214. Bagheri-Majdi, E., Ke, Y., Orlova, G., Chu, I. K., Hopkinson, A. C. & Siu, K. W. M. Copper-mediated peptide radical ions in the gas phase. *J. Phys. Chem. B* **108,** 11170-11181 (2004).

215. Bellina, B., Compagnon, I., Houver, S., Maître, P., Allouche, A-R., Antoine, R. & Dugourd, P. Spectroscopic signatures of peptides containing tryptophan radical cations. *Angew. Chem. Int. Ed. Engl.* **50,** 11430–2 (2011).

216. Polfer, N. C. & Dugourd, P. Laser photodissociation and spectroscopy of mass-separated biomolecular ions. *Laser Photodissociation Spectrosc. Mass-separated Biomol. Ions* **83,** 1–21 (2013).

217. Janz, J. M., Ren, Y., Looby, R., Kazmi, M. A., Sachdev, P., Grunbeck, A., Haggis, L., Chinnapen, D., Lin, A. Y., Seibert, C., McMurry, T., Carlson, K. E., Muir, T. W., Hunt, S. & Sakmar, T. P. Direct interaction between an allosteric agonist pepducin and the chemokine receptor CXCR4. *J. Am. Chem. Soc.* **2011**, *133*, 15878-15881.

218. Coste, J., LeNguyen, D. & Castro, B. PyBOP: A new peptide coupling reagent devoid of toxic by-product. *Tetrahedron Lett.* **1990**, *31*, 205-208.

219. Moss, C. L., Chamot-Rooke, J., Brown, J., Campuzano, I., Richardson, K., Williams, J., Bush, M. F., B. Bythell, Paizs, B. & Tureček, F. Assigning structures to gas-phase peptide cations and cation-radicals. An infrared multiphoton dissociation, ion mobility, electron transfer and computational study of a histidine peptide ion. *J. Phys Chem. B.* **2012**, *116*, 3445-3456.

220. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. & Schulten K. Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **2005**, *26*, 1781–1802.

221. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D. & Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–616 (1998).

222. Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314,** 141–151 (1999).

223. Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **10,** 209–220 (1989).

224. Furche, F. & Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **117,** 7433 (2002).

225. Schlegel, H. B. Potential energy curves using unrestricted Møller–Plesset perturbation theory with spin annihilation. *J. Chem. Phys.* **4530,** 1–6 (1986).

226. Mayer, I. The spin-projected extended Hartree-Fock method. *Adv. Quantum Chem.* **12,** 189–262 (1980).

227. McWeeny, R. & Diercksen, G. Self-consistent perturbation theory. II. Extension to open shells. *J. Chem. Phys.* **49,** 4852–4856 (1968).

228. Parkinson, C. J., Mayer, P. M. & Radom, L. An assessment of theoretical procedures for the calculation of reliable radical stabilization energies. *J. Chem. Soc. Perkin Trans. 2* **43,** 2305–2313 (1999).

229. Gregersen, J. & Tureček, F. Mass-spectrometric and computational study of tryptophan radicals (Trp + H)• produced by collisional electron transfer to protonated tryptophan in the gas phase. *Phys. Chem. Chem. Phys.* **12,** 13434–47 (2010).

230. Chu, I. K., Lam, C. N. W. & Siu, S. O. Facile generation of tripeptide radical cations in vacuo via intramolecular electron transfer in CuII tripeptide complexes containing sterically encumbered terpyridine ligands. *J. Am. Soc. Mass Spectrom.* **16,** 763–771 (2005).

231. Marek, A. & Tureček, F. Collision-induced dissociation of diazirine-labeled peptide ions. Evidence for Brønsted-acid assisted elimination of nitrogen. *J. Am. Soc. Mass Spectrom.* **25,** 778–789 (2014).

232. Dunbar, R. C. Infrared radiative cooling of isolated polyatomic molecules. *J. Chem. Phys.* **90,** 7369 (1989).

233. Ichou, F., Schwarzenberg, A., Lesage, D., Alves, S., Junot, C., Machuron-Mandard, X. & Tabet, J. C. Comparison of the activation time effects and the internal energy distributions for the CID, PQD and HCD excitation modes. *J. Mass Spectrom.* **49,** 498–508 (2014).

234. Tolmachev, A. V., Vilkov, A. N., Bogdanov, B., Păsa-Tolić, L., Masselon, C. D. & Smith, R. D. Collisional activation of ions in RF ion traps and ion guides: The effective ion temperature treatment. *J. Am. Soc. Mass Spectrom.* **15,** 1616–1628 (2004).

235. Taylor, J. A. & Johnson, R. S. Sequence database searches via *de novo* peptide sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.***11,** 1067–1075 (1997).

236. Shui, W., Liu, Y., Fan, H., Bao, H., Liang, S., Yang, P. & Chen, X. Enhancing TOF/TOF-based *de novo* sequencing capability for high throughput protein identification with amino acid-coded mass tagging. *J. Proteome Res.* **4,** 83–90 (2005).

237. Cagney, G. & Emili, A. *De novo* peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nature*. **20,** 163–170 (2002).

238. Cao, X. & Nesvizhskii, A. I. Improved sequence tag generation method for peptide identification in tandem mass spectrometry. *J. Proteome Res.* **7,** 4422–4434 (2008).

239. Kwon, K-H., Kim, M., Kim, J. Y., Kim, K. W., Kim, S. I. Park, Y. M. & Yoo J. S. Efficiency improvement of peptide identification for an organism without complete genome sequence, using expressed sequence tag database and tandem mass spectral data. *Proteomics* **3,** 2305–2309 (2003).

240. Wang, X., Li, Y., Wang, H., Tan, H & Peng, J. JUMP: A tag-based database search tool for peptide identification with high sensitivity and accuracy. *Mol. Cell Proteomics* **13,** 3663–3673 (2014).

241. Chao, S., Green, J. R. & Smith, J. C. Evaluation of a GPGPU-based *de novo* peptide sequencing algorithm. *J. Med. Biol. Eng.* **34,** 461–468 (2014).

242. Riley, N. M., Mullen, C., Weisbrod, C. R., Sharma, S., Senko, M. W., Zabrouskov, V., Westphall, M. S., Syka, J. E. P. & Coon, J. J. Enhanced dissociation on intact proteins with high capacity electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **27,** 520–531 (2016).

243. Wang, P. & Wilson, S. R. Mass spectrometry-based protein identification by integrating *de novo* sequencing with database searching. *BMC Bioinformatics* **14,** S2–S24 (2013).

244. Yefremova, Y., Al-Majdoub, M., Opuni, K. F., Koy, C., Cui, W., Yan, Y., Gross, M. L. & Glocker, M. O. "*De novo*" amino acid sequence elucidation of protein G'e by combined "top-down" and "bottom-up" mass spectrometry. *J. Am. Mass Spectrom.* **26,** 482–492 (2015).

245. Liu, X., Dekker, L. J. M., Wu, S., Vanduijn, M. M., Luider, T. M., Tolíc, N., Kou, Q., Dvorkin, M., Alexandrova, S., Vyatkina, K., Paša-Tolíc, L. & Pevzner, P. A. *De novo* protein sequencing by combining top-down and bottom-up tandem mass spectra. *J. Proteome Res.* **13,** 3241–3248 (2014).

246. Allmer, J. Algorithms for the *de novo* sequencing of peptides from tandem mass spectra. *Expert Rev. Proteomics* **8,** 645–657 (2011).

247. Xue, C. Method development for qualitative and quantitative study in mass spectrometry. (Dissertation). ResearchWorks: University of Washington, Seattle (2014).

248. De Rossi, R. H. & Veglia, A. V. On the product distribution in the iodination of phenol. *Tetrahedron Lett.* **27,** 5963–5966 (1986).

249. Veglia, A. V. & De Rossi, R. H. Selectivity in the iodination of phenol in the presence of .beta.-cyclodextrin. *J. Org. Chem.* **53,** 5281–5287 (1988).

250. Salek, M. & Lehmann, W. D. Analysis of thyroglobulin iodination by tandem mass spectrometry using immonium ions of monoiodo- and diiodo-tyrosine. *Proteomics* **5,** 351–353 (2005).

251. Lista, L., Pezzella, A., Napolitano, A. & D'Ischia, M. Mild and efficient iodination of aromatic and heterocyclic compounds with the NaClO2/NaI/HCl system. *Tetrahedron* **64,** 234–239 (2008).

252. Emmanuvel, L., Shukla, R. K., Sudalai, A., Gurunath, S. & Sivaram, S. NaIO4/KI/NaCl: a new reagent system for iodination of activated aromatics through in situ generation of iodine monochloride. Tetrahedron Lett. 47, 4793–4796 (2006).

253. Zielinska, A. & Skulski, L. Eco-friendly oxidative iodination of various arenes with sodium percarbonate as the oxidant. *Molecules* **10,** 1307–1317 (2005).

254. Das, B., Krishnaiah, M., Venkateswarlu, K. & Reddy, V. S. A mild and simple regioselective iodination of activated aromatics with iodine and catalytic ceric ammonium nitrate. *Tetrahedron Lett.* **48,** 81–83 (2007).

255. Arotsky, J., Darby, A. C. & Hamilton, J. B. A. Iodination and iodo-compounds. Part II. The kinetics of aromatic iodination by means of the tri-iodine cation. *J. Chem. Soc. B Phys. Org.* **1,** 739–742 (1968).

256. Arotsky, J., Darby, A. C. & Hamilton, J. B. A. Iodination and iodo-compounds. Part IV. The effect of substituents and solvent composition on the rate of aromatic iodination by means of the tri-iodine cation. *J. Chem. Soc. Perkin Trans.* **2**, 595 (1973).

257. Soper, F. G. & Smith, G. F. CCCLXVII. The halogenation of phenols. Part II. Iodination. *J. Chem. Soc.* 2757–2761 (1927).

258. Adimurthy, S., Ramachandraiah, G., Ghosh, P. K. & Bedekar, A. V. A New, Environment Friendly Protocol for Iodination of Electron-Rich Aromatic Compounds. *ChemInform* **34,** (2003).

259. Gallo, R. D. C., Gebara, K. S., Muzzi, R. M. & Raminelli, C. Efficient and selective iodination of phenols promoted by iodine and hydrogen peroxide in water. *J. Braz. Chem. Soc.* **21,** 770–774 (2010).

260. Chaikovskii, V. K., Funk, A. A., Filimonov, V. D., Petrenko, T. V. & Kets, T. S. Facile iodination of aromatic compounds having electron-withdrawing substituents. Generation of the triiodide cation in the system tetra-N-iodoglycoluril-iodine-sulfuric acid. *Russ. J. Org. Chem.* **44,** 935–936 (2008).

261. Kilby, D. C. Kinetic study of the mechanism of the iodination of phenol (Dissertation). Georgia Institue of Technology (1957).

262. Gallo, R. D. C., Gebara, K. S., Muzzi, R. M. & Raminelli, C. Efficient and selective iodination of phenols promoted by iodine and hydrogen peroxide in water. *J. Braz. Chem. Soc.* **21,** 770–774 (2010).

263. Khalilzadeh, M. A., Hosseini, A., Shokrollahzadeh, M., Halvagar, M., Ahmadi, D., Mohannazadeh, F. & Tajbakhsh, M. HIO4/Al2O3 as a new system for iodination of activated aromatics and 1,3-dicarbonyl compounds. *Tetrahedron Lett.* **47,** 3525–3528 (2006).

264. Schutte, L. & Havinga, E. Base catalysis in aromatic iodination. Kinetics and mechanism of the iodination of p-cresol, 2,6-dimethylphenol, imidazole and dimethylaminobenzenesulfonic acid. *Tetrahedron* **26,** 2297–2304 (1970).

265. Masson, I. 322. Hypoiodous cations, and their action upon an organic reagent. *J. Chem. Soc.* 1708 (1938).

266. Edgar, K. J. & Falling, S. N. An efficient and selective method for the preparation of iodophenols. *J. Org. Chem.* **55,** 5287–5291 (1990).

267. Vasicek, L. & Brodbelt, J. S. Enhanced electron transfer dissociation through fixed charge derivatization of cysteines. *Anal. Chem.* **81,** 7876–7884 (2009).

268. Samgina, T. Y., Vorontsov, E. A., Gorshkov, V. A., Artemenko, K. A., Nifant'ev, I. E., Kanawati, B., Schmitt-Kopplin, P., Zubarev, R. A. & Lebedev, A. T. Novel cysteine tags for the sequencing of non-tryptic disulfide peptides of anurans: ESI-MS study of fragmentation efficiency. *J. Am. Soc. Mass Spectrom.* **22,** 2246–2255 (2011).

269. Frey, B. L., Ladror, D. T., Sondalle, S. B., Krusemark, C. J., Jue, A. L., Coon, J. J. & Lloyd, M. S. Chemical derivatization of peptide carboxyl groups for highly efficient electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **24,** 1710–1721 (2013).

270. Anderson, L. C., English, A. M., Wang, W-H., Bai, D. L., Shabanowitz, J. & Hunt, D. F. Protein derivatization and sequential ion/ion reactions to enhance sequence coverage produced by electron transfer dissociation mass spectrometry. *Int. J. Mass Spectrom.* **377,** 617–624 (2015).

271. Seymour, J. L. & Tureček, F. Distinction and quantitation of leucine-isoleucine isomers and lysine-glutamine isobars by electrospray ionization tandem mass spectrometry (MS^n, n=2,3) of copper (II)-diimine complexes. *J. Mass Spectrom.* **35,** 566–571 (2000).

272. Tao, A. W., Wu, L. & Cooks, G. R. Differentiation and quantitation of isomeric dipeptides by low-energy dissociation of copper(II)-bound complexes. *J. Am. Soc. Mass Spectrom.* **12,** 490–496 (2001).

273. Borisov, R. S., Zaikin, V. G., Vas'kovskii, B. V. & Sklyarov, L. Yu. Modification of dipeptides by alkyl chloroformate –alkanol mixtures for analysis by gas chromatography/mass spectrometry with electron and chemical ionization and collisonal activation: Differentiation of isomers. *Russ. Chem. Bull.* **55,** 2285–2290 (2006).

274. Cook, S. L., Collin, O. L. & Jackson, G. P. Metastable atom-activated dissociation mass spectrometry: leucine/isoleucine differentiation and ring cleavage of proline residues. *J. Mass Spectrom.* **44,** 1211–1223 (2009).

275. Nakamura, T., Nagaki, H., Ohki, Y. & Kinoshita, T. Differentiation of leucine and isoleucine residues in peptides by consecutive reaction mass spectrometry. *Anal. Chem.* **62,** 311–313 (1990).

276. Armirotti, A., Millo, E. & Damonte, G. How to discriminate between leucine and isoleucine by low energy ESI-TRAP MS^n. *J. Am. Soc. Mass Spectrom.* **18,** 57–63 (2007).

277. Lebedev, A. T., Damoc, E., Marakov, A. A. & Samgina, T, Yu. Discrimination of Leucine and isoleucine in peptides sequencing with Orbitrap Fusion Mass Spectrometer. *Anal. Chem.* **86,** 7017 –7022 (2014).

278. Sidoli, S., Yuan, Z-F., Lin, S., Karch, K., Wang, X. Bhanu, N., Arnaudo, A. M., Britton, L-M., Cao, X-J., Gonzales-Cope, M., Han, Y., Liu, S., Molden, R. C., Wein, S., Afjehi-Sadat, L. & Garcia, B. A. Drawbacks in the use of unconventional hydrophobic anhydrides for histone derivatization in bottom-up proteomics PTM analysis. *Proteomics*. **15,** 1459–1469 (2015).

279. Tang, H., Fang, H., Tin, E., Brasier, A. R., Sowers, L. C. & Zhang, K. Multiplexed parallel reaction monitoring targeting histone modifications on the QExactive Mass Spectrometer. *Anal. Chem.* **86,** 5526–5534 (2014).

280. Zhang, J., Chen, Y., Zhang, Z., Xing, G., Wysocka, J. & Zhao, Y. MS/MS/MS reveals false positive identification of histone serine methylation. *J. Proteome Res.* **9,** 585–594 (2010).

281. Maile, T. M., Izrael-Tomasevic, A., Cheung, T., Guler, G. D., Tindell, C., Masselot, A., Liang, J., Zhao, F., Trojer, P., Classon, M. & Arnott, D. *Mol. Cell. Proteomics*. **14,** 1148–1158 (2015).

282. Lin, S., Wein, S., Gonzales-Cope, M., Otte, G. L., Yuan, Z-F., Afjehi-Sadat, L. Maile, T., Berger, S. L., Rush, J., Lill, J. R., Arnott, D. & Garcia, B. A. *Mol. Cell. Proteomics*. **13,** 2450–2466 (2014).

283. Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J. & Hunt, D. F. *Nat. Protoc.* **2,** 933–938 (2007).