Mass Spectrometry-Based Analysis of Photochemical Cross-linking between Nitrile Imine and Trp, Phe, and Tyr Amino Acid Residues in Gas-Phase Ions

Sizhong Shen

A Thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Applied Chemical Science and Technology

University of Washington

2025

Committee:

František Tureček

Nicholas M. Riley

Program Authorized to Offer Degree:

Department of Chemistry

University of Washington

Abstract

Mass Spectrometry-Based Analysis of Photochemical Cross-linking between Nitrile Imine and Trp, Phe, and Tyr Amino Acid Residues in Gas-Phase Ions

Sizhong Shen

Chair of the Supervisory Committee:

František Tureček

Department of Chemistry

In this study, conjugates of 2,5-diaryltetrazole and peptapeptides were photodissociated by 213 nm UV light. The nitrile imine group was generated from the tetrazole by losing a nitrogen molecule and cross-linked with carbonyl group of amide or amino acid side chain through intramolecular reactions. By analyzing the tandem mass spectra of ultraviolet photodissociation (UVPD) mass spectrometry (MS) and collision induced dissociation (CID) mass spectrometry, product structures were proposed. Theoretical calculations based on Born-Oppenheimer molecular dynamics (BOMD) were conducted for the conformational

analyses of the proposed structures, and density-functional theory (DFT) based computations were used to evaluate isomer energies. Nitrile-imine cross-linking reaction mechanisms were suggested based on calculated structures of the peptide conjugates and product ions.

ACKNOWLEDGEMENTS

First, I would like to thank my PI, Prof. František Tureček. His enthusiasm for science cheered me up at my most self-doubting time one year ago and encouraged me to continue my career in chemistry. His outstanding knowledge, insight into science, and pursuit of perfection not only influenced this thesis deeply but also made me understand the value of working as a scientist. I would also like to thank Dr. Martin Sadilek, manager of Mass Spectrometry Facility of the Department of Chemistry, for his technical support and suggestions for my works.

As for my peer workers in Tureček group, I would rather say you are my "peer mentors" that helped me greatly in my research. I would like to thank Dr. Jiahao Wan and Dr. Hongyi Zhu, who have already graduated, for training me a lot in organic synthesis, mass spectrometry, and calculation setups. I thank Chengyang for negotiating sample orders and instrument reservations. I thank Mars for his outstanding performance as a mentee, he made me realized my shortcomings and encouraged me to learn more.

Finally, I would like to thank my parents Dong and Jie for their selfless support and encouragement, both physical and emotional. I thank all my friends both in Seattle and my hometown that helped me in my life and research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

1.1 Photochemical Crosslinking and Nitrile Imines

Crobsslinking is a widely used research method for modifying molecular structures or conjugating molecules together.[1,2] With carefully designed reactions, crosslinkers attached to specific positions in a molecule will trigger inter- or intramolecular reactions with other reactive groups, forming covalent bonds that link those positions. The application of crosslinking has expanded significantly in recent years. The development of next-generation materials through crosslinking has a solid theoretical basis and it is now flourishing,[3,4] for example, the development of crosslinked hydrogels from macromolecules has attracted considerable attention. Crosslinked hydrogels formed from carboxylic acids or polysaccharides are biochemically-friendly, making them ideal media for drug delivery, especially their materials are readily accessible.[5,6] Thanks to the chemical bond formed by crosslinkers, these crosslinked hydrogels exhibit superior stability compared to physical hydrogels.[7] By adopting naturally originated materials such as silk, it is possible to produce durable hydrogels with high adaptability and bioactivity, offering new possibilities in bioengineering and the medical field.[8]

On the other hand, research about crosslinking in the biomolecular field, such as studies on peptides and nucleotides, is also of great interest. [9,10] Among those latest works, studies

involving photochemical crosslinkers stand out as particularly noteworthy.[11] Photochemical crosslinkers usually stay unreactive until exposed to specific conditions, such as UV light with a certain wavelength. Through photolysis reaction, the stable crosslinker group, also known as a tag, is converted into a highly reactive intermediate that initiates covalent bond

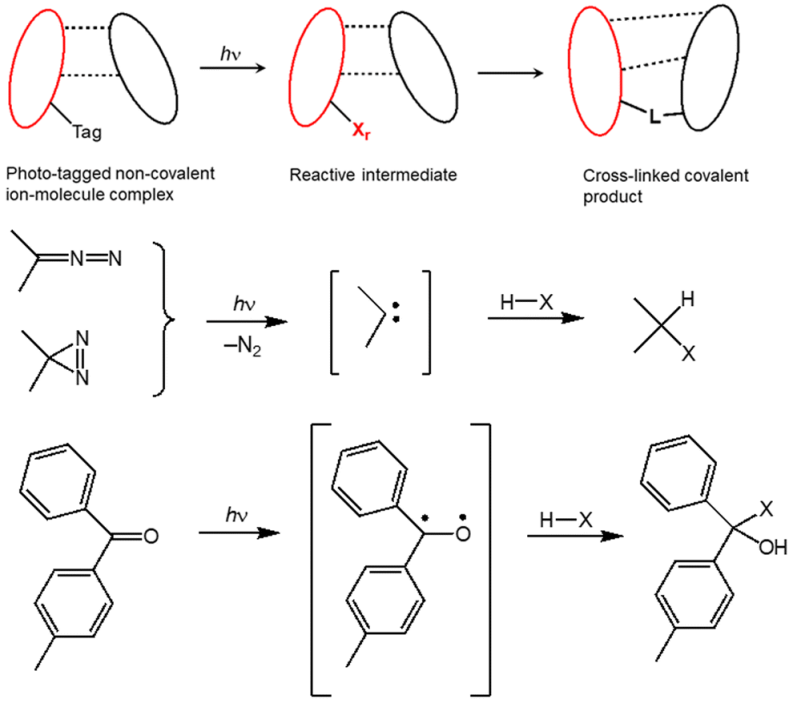

Figure 1.1. Photochemical crosslinking, diazirine and benzophenone.

formation (Figure 1.1). In this way, photochemical crosslinkers become valuable tools for cmodifying or conjugating biomolecules, such as peptides. Most common photochemical crosslinkers, like diazirine, generate reactive carbenes upon activation, while nitrenes or other intermediates may also be involved in cases with other crosslinkers. Photo-activable amino acids with diazirine or benzophenone tags (i.e., amino acids with photochemical tags attached to their side chains) have been applied in peptide research for over 30 years and are now commercially available.[12] However, as research progresses, some drawbacks of carbenes and diradical intermediates have been identified. When working with diazirine, activating takes longer because electron excitation from the ground-state π_{xy} molecular orbital to the π^*_z

first excited singlet state is inefficient, thus resulting in an extended data acquisition time. This brings drawbacks for gaseous phase studies, such as mass spectrometry.[13] Secondly, when activated with UV light at a wavelength of approximately 365 nm, the diazirine group generates a carbene intermediate that becomes highly reactive with C–H, O–H, and N–H single bonds. However, its reactivity with C–H bonds can lead to unfavorable side reactions and brings additional challenges for data analysis: the lack of selectivity makes carbene-based crosslinkers less suitable in certain applications.[14] Additionally, due to the very short half-life of the reactive intermediates (caused by their rapid self-destruction reactions), both diazirine and benzophenone have a low probability of reacting with intended targets: these intermediates may yield low crosslinking efficiency.[11,13]

As a class of synthetic heterocyclic compounds, tetrazoles are widely used in pharmaceutical research and in the design of high-performance explosives.[15,16] However, the reactive intermediate nitrile imine, generated from tetrazoles under heat or UV light, has recently gained attention and made 2,5-diaryltetrazoles novel photo-tags in photochemical crosslinking research. Nitrile imines have been applied in organic synthesis for a long time,[17] and the discovery of photodecomposition of diaryltetrazole group and the formation of nitrile imines in 1959 built the theoretical basis of the photochemical crosslinking attempts.[18] Nitrile imines are known to undergo well-known [3 + 2] cycloadditions with various dipolarophiles (Figure 1.2),[19,20,21] but recent research also pointed out its possibility of not only reacting with amide group,[21] carboxyl group,[11,22] but also with primary amine group.[23]

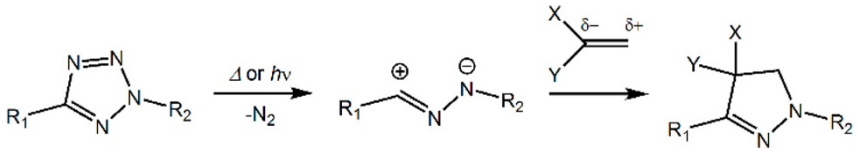

Figure 1.2. Formation of nitrile imine and [3 + 2] addition.[21]

These discoveries demonstrate the potential of nitrile imines as a universal tool for biomolecular conjugation.[21] Comparing with the tags employing carbene or nitrene intermediates, 2,5-diaryltetrazoles have some benefits, especially for experiments in the gas phase as applied in mass spectrometry. First, photoactivation of the nitrile imine takes much less time than with regular crosslinkers to react in the gas phase. Since 2,5-diaryltetrazole has a major absorption band with λ_{max} = 255nm,[24] the energy provided by photons is about 4.96 eV (479 kJ/mol). According to calculations based on M06-2X/6-311++G(2d,p), the reaction enthalpy of tetrazole dissociation by loss of N_2 is about 51 kJ/mol, which is much less than the energy provided by photons. Due to this great energy difference, the crosslinking reaction could occur on a millisecond time scale and elongated irradiation time is avoided. [13] Second, this rapid reaction considerably improves the crosslinking yield. Research done in the Turecek research group has shown that crosslinking yields as high as 69% to 80% were reached in some peptide samples, and even a 90% yield was possible.[13,22] These benefits make nitrile-imine-based crosslinkers ideal tools for biomolecular structure research.

1.2 Crosslinking Mass Spectrometry in Protein Research

Due to the invention of electrospray ionization (ESI) and Matrix-assisted laser desorption/ionization (MALDI) in 1960s to 1980s, as well as the rapid development of algorithms and instrument after 1990, mass spectrometry (MS) has become a powerful tool in medical and biomolecular research, especially for proteins.[25] Thanks to the high resolution

measurement of m/z ratio, multiple studies, particularly those about post-translational modifications, have highlighted the crucial role of tandem mass spectrometry (MS/MS) in detecting amino acid sequence modifications[26] as well as side-chain changes of amino acids.[27] Among those MS-based methods of peptide research, crosslinking mass spectrometry (XL-MS) brings considerable benefits in structural research. For those larger molecules that are not suitable for Fourier transform ion cyclotron-resonance (FTICR)-MS, measuring distances between positions indirectly with crosslinkers under the help of XL-MS seems a better solution[14]: by measuring the lengths of the crosslinkers that connect the two protein sites, it is possible to infer the overall structure of the protein even though the real 3D-structural data is not obtained. After the publishment of some earliest works that conducted back in early 2000s[28], this field is flourishing and the applications of new technologies bring more possibilities for future research: for example, prediction of protein structure could be much easier with the help of deep learning[29]. However, due to the lack of unified crosslinking method, analyzing tools, and reporting formats, peer researchers are facing more challenges when judging the accuracy and reliability of the data, and new learners are also easily confused by these complex terms and methods[14]. These issues are prompting scientists to focus on related fields and to make more progress.

Chapter 2. Experimental Section

Experiment Design and Research Methods

In 2008, Song et al. reported an approach of using [3+2] cycloaddition to attach dipolarophiles (C=C double bonds) to a 2,5-diaryltetrazole group that conjugated to the side

chain amino group of lysine under UV light. As a methodology for studying and modifying peptide structures, and investigating protein conformational and spatial information, nitrile imine photochemical crosslinking is being developed with more diversified and standardized crosslinking methods. Recently, more dipolarophiles were discovered to react with diaryltetrazole based on this cycloaddition reaction, such as amide group and carboxyl group.[21,22] Moreover, a rather unexpected reaction was reported that diaryltetrazole and primary amine are able to form triazole with a considerable yield, under basic condition and UV light.[23]The main goal of the research done by the Turecek research group is to discover more possible reaction mechanisms of nitrile imines to enrich the crosslinking toolbox and bring more possibilities to the crosslinking method.

In my research, 5-amino-acid-long peptide chains encoded as “XAAAK” (X could be Y, F, or W) from N to C terminus were synthesized on solid-support (Wang resin) with standardized coupling procedures. With the conjugation of a 2,5-diaryltetrazolecarbonyl tag at the side chain amine of lysine, the peptide chains were called “XAAA-*tet*-K”, for example, YAAA-*tet*-K represents a peptide chain with tyrosine on N terminus and lysine with diaryltetrazole attachment on C terminus. Crosslinking reactions were conducted in CID-MS or UVPD-MS with UV laser of 213 nm on a commercially available Thermo Fisher Orbitrap Ascend instrument. The spectra were analyzed for initial guesses of the structures, and these structures will be probed by computational analysis (Figure 2.1).

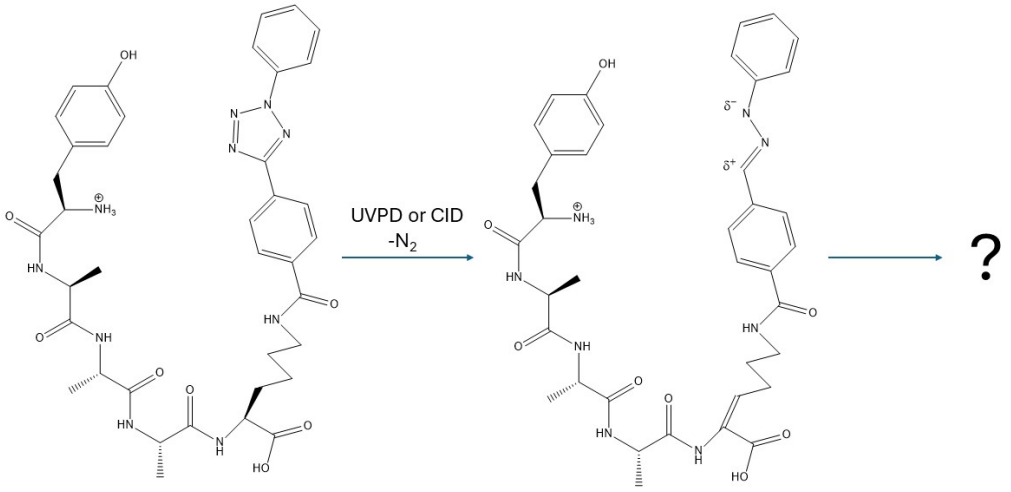

Figure 2.1. Nitrile Imine Formation and Crosslinking Possibility.

For calculations, a protocol previously developed by our group was followed (Figure 2.2).[13] First, Born-Oppenheimer molecular dynamics (BOMD) calculations of 20 ps trajectories with 1 fs steps were conducted under PM6-D3H4. These calculations used Berendsen thermostat[31] at 510, 610, and 810 K. Semiempirical Hamiltonian, dispersion and hydrogen-bonding interactions of the structures were used in these calculations. Two hundred snapshots were selected randomly from each BOMD trajectory, and these structures were optimized by PM6-D3H4.[32] Several structures with lowest energies were selected for further optimization.

Next, further structural optimization and calculations were conducted based on density functional theory (DFT) methods. The structures seected from BOMD then further optimized with B3LYP and 6-31+G(d,p) basis set with empirical dispersion corrections (GD3-BJ).[33,34,35] Structures with lower energies were selected and reoptimized with M06-2X with 6-31+G(d,p) basis set. These structures then were used to calculate single-point energies with M06-2X and def2qzvpp basis set.[36] All energy calculations were conducted under a 310 K

condition. In addition, calculations based on M06-2X with 6-31++G(d,p) were conducted for charge densities. Charge density data were used to calculate ion collision cross sections (CCS) to be compared with experimental data obtained by our collaborators. Through these calculations, the structures that had the lowest energy and match the measured CCS data were regarded as the ones that best represent the real structures of peptide chains, before or after the crosslinking.

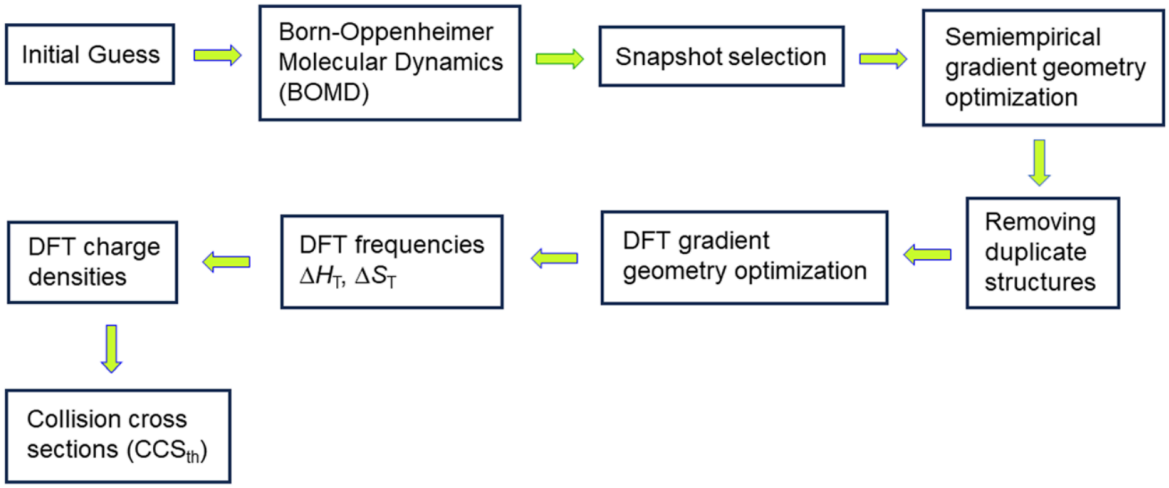

Figure 2.2. Workflow of computational analysis.[13]

Chapter 3. Mass Spectrometry-Based Analysis of Photochemical Crosslinked peptides

3.1 Spectra Analysis and Calculation of YAAA-tet-K

CID-MS^2 and UVPD-MS^2 at 213 nm brought different spectra for YAAAK ions ($(YAAAK+H)^+$, *m/z* 771) (Figure 3.1). Both methods caused a loss of N_2 from the diaryltetrazole (*m/z* 743), but in spectrum of CID, a series of peaks with *m/z* 537, 466, and 395 were observed with greater intensities. These peaks represented the standard backbone cleavage of the peptide chain, and according to the peptide ion nomenclature system and,[37,38]

these fragments were referred to as $[y_3 + 2H]^+$ to $[y_1 + 2H]^+$. Similarly, the peak with *m/z* 608 is referred to as $[x_4 + 2H]^+$ due to the elimination of a C_8H_9NO fragment from the tyrosine residue after photodissociation of the tetrazole group. From CID-MS^3 spectrum of the *m/z* 608 fragment, a peak with *m/z* 537 was found and assigned as $[x_3 + 2H]^+$ based on the same nomenclature system. However, the *m/z* 456 and 385 peaks indicated another fragmentation pattern that shared some similarities with the $[y_n + 2H]^+$ system, and this system was referred to as "υ_n" ions, in our previous work.[22] An oxygen transfer occurred during the process of isolation and fragmentation of *m/z* 608 ion, and a new C–O double bond was formed between the electron-deficient carbon and a peptide oxygen atom. At the same time, two N-H single bonds were formed on nitrile-imine nitrogen atoms. Even though the

sites of backbone cleavage were the same, the structural changes distinguished the "υ_n" fragments from $[y_n + 2H]^+$ fragments by different m/z values.

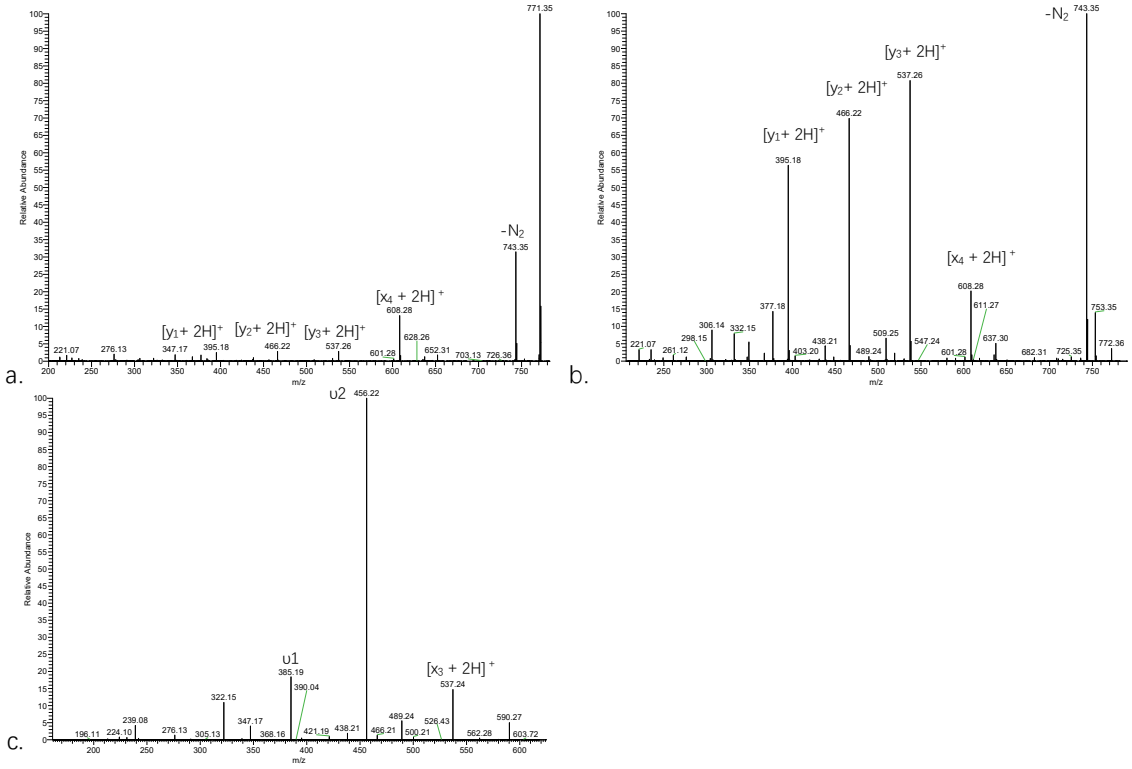

Figure 3.1. a. UVPD-MS^2 and b. CID-MS^2 of *m/z* 771, c. UVPD-MS^3 of *m/z* 608, of YAAA-*tet*-K sequence

An uncommon fragmentation pattern was found for both UVPD and CID-MS^3 spectra of the *m/z* 743 peak, which consisted of *m/z* 672, 601, and 530 ions, which is a fragment ion series representative of crosslinking (Figure 3.2). The consecutive elimination of alanine residues (indicated as -[Ala], -[AlaAla], -[AlaAlaAla] in figures) indicated that the tyrosine on the N terminus crosslinked with the nitrile imine in certain way that left the alanine residues to be chopped off from the chain, like we discussed in previous works.[21,22] However, the fragments with *m/z* 580, 509, 438, and 367 formed another fragmentation sequence of standard backbone cleavage of linear peptide, which showed that some ions were not able to

crosslink. The peak with *m/z* 508 indicated the loss of an aniline molecule from the fragment with *m/z* 601, and the peak with *m/z* 437 gave evidence of one alanine residue cleavage from *m/z* 508 fragment. CID-MS^4 spectra of the *m/z* 635 peak formed by UVPD-MS^3 indicated that the peptide ion still maintained a cyclic shape even after the cleavage of phenylhydrazine, because of the presence of internal loss of alanine residues (*m/z* 564, 493 peaks). (Figure 3.2).

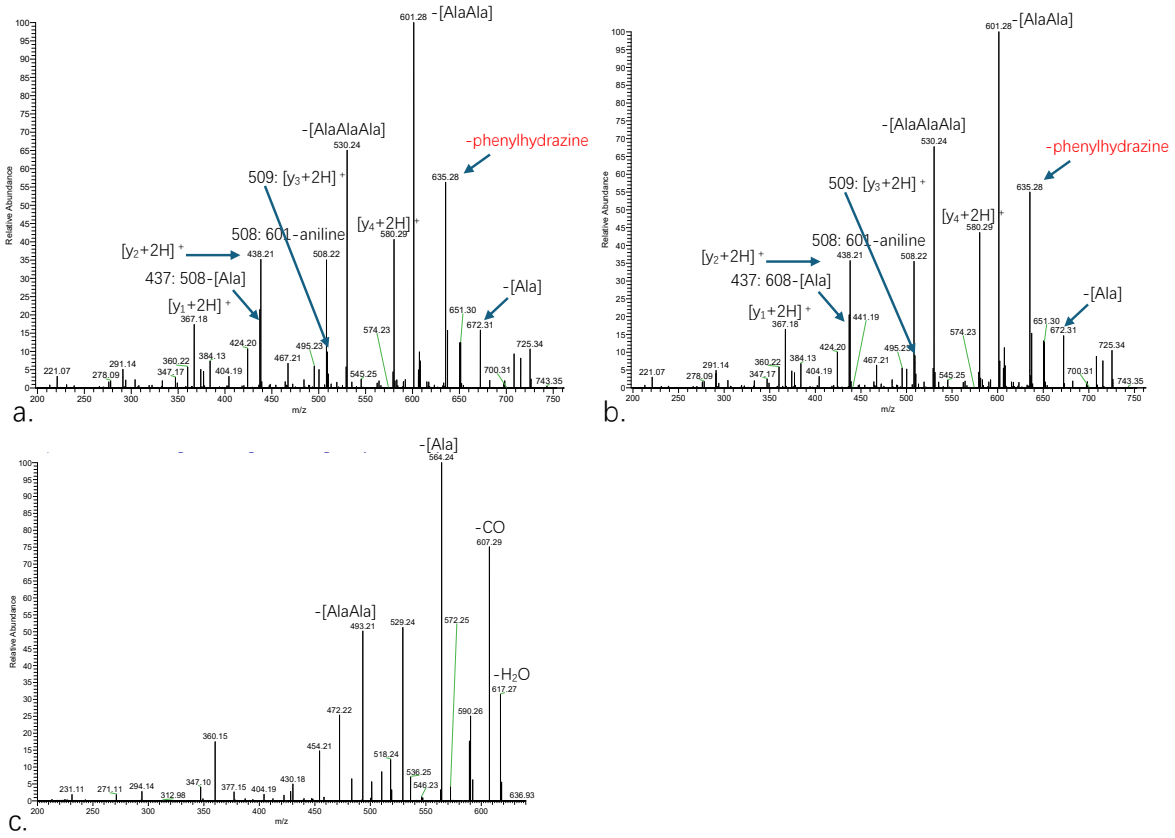

Figure 3.2. a. UVPD-MS^3 of *m/z* 743. b. CID-MS^3 of *m/z* 743 and c. UVPD-MS^4 of *m/z* 635 spectra of YAAA-*tet*-K sequence.

Primary guesses about the crosslinked peptides were proposed after analyzing the spectra, and stable structures with low-Gibbs energy were obtained by calculations (Figure 3.3). **Y1**-**Y3** showed the structures before the cleavage of nitrogen, and **Y2** obtained the

lowest Gibbs energy of 0.0 kJ mol^{-1}, at 310 K. Structure **Y4** represented an intermediate just after the cleavage of N_2, whereby the nitrile imine group was formed but the crosslinking reaction had yet to happen. This structure had a relative Gibbs energy of 110 kJ mol^{-1}. In our guesses, according to the spectra, two types of crosslinking products with different structures were formed. **Y7** and **Y8** represented the mechanism in which a new C–O double bond was formed by oxygen transfer and a C-N single bond was formed between the nitrile-imine nitrogen and the carbon of tyrosine. This was similar to those reported previously for GAAA-*tet*-K.[21] According to the M06-2X/def2qzvpp calculated energies**, Y7** had a lower Gibbs energy of 58 kJ mol^{-1} in 310 K. On the other hand, based on the UVPD-MS^4 of *m/z* 635, we proposed a new structure for the crosslinking product showed as **Y5** and **Y6**: a C–O single bond was formed between the tyrosine side chain and the carbon atom of tetrazole ring, while this carbon atom linked with the nitrile-imine nitrogen with a C–N double bond. At the same time the phenol hydrogen of the tyrosine side chain was transferred to a nitrile-imine

nitrogen. In this case, structure **Y5** was more stable and had a lower Gibbs energy that we assigned the reference value of 0.0 kJ mol^{-1}.

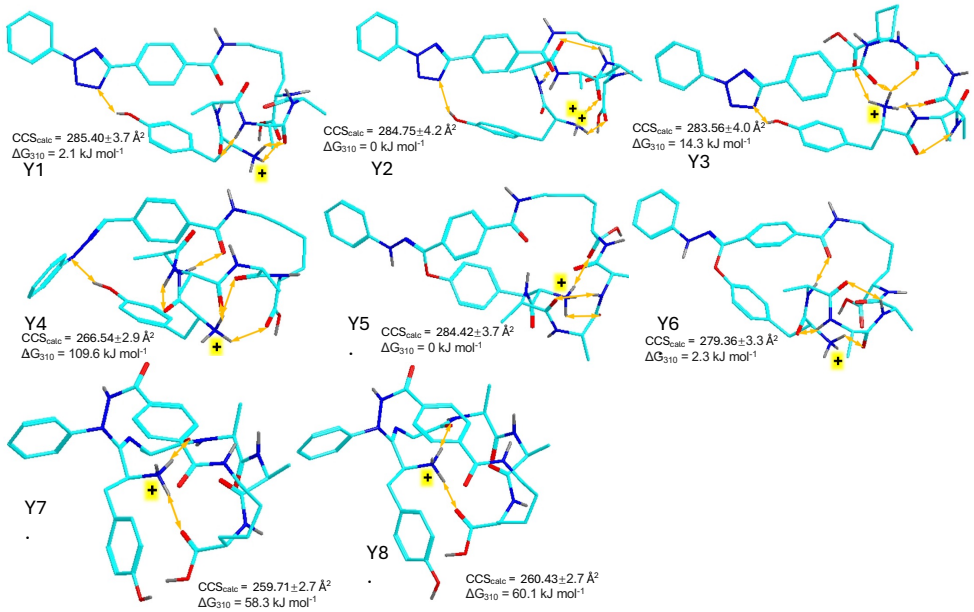

Figure 3.3. M06-2X/6-31+G(d,p) optimized structures of representative low-energy ions associated with YAAA-*tet*-K. Atom color coding is as follows: cyan = C, blue = N, red = O, gray = H. Only exchangeable hydrogens are shown to avoid clutter. Hydrogen bonds are visualized by ochre double-headed arrows. Relative Gibbs energies are from M062X/def2qzvpp single-point energy calculations including B3LYP/6-31+G(d,p) zero-point corrections, and 310 K enthalpies and entropies.

3.2 Spectra Analysis and Calculation of FAAA-tet-K

The FAAA-*tet*-K sequence exhibited a similar fragmentation behavior and spectra as YAAA-*tet*-K in MS^2. Compared to UVPD-MS^2, CID-MS^2 exhibited a greater loss of nitrogen (*m/z* 727) and greater $[x_4 + 2H]^+$ (*m/z* 608) and $[y_3 + 2H]^+$ to $[y_1 + 2H]^+$ (*m/z* 537, 466, 395) peaks. The UVPD-MS^3 of the *m/z* 608 peak exhibited a similar fragmentation pattern as

shown in the YAAA-*tet*-K spectrum, $[x_3 + 2H]^+$, "**υ_2**" and "**υ_1**" fragments were found in this spectrum. This indicated that oxygen transfer and the formation of "**υ_n**" fragments in YAAA-*tet*-K and FAAA-*tet*-K may not closely associated with tyrosine and phenylalanine (Figure 3.4).

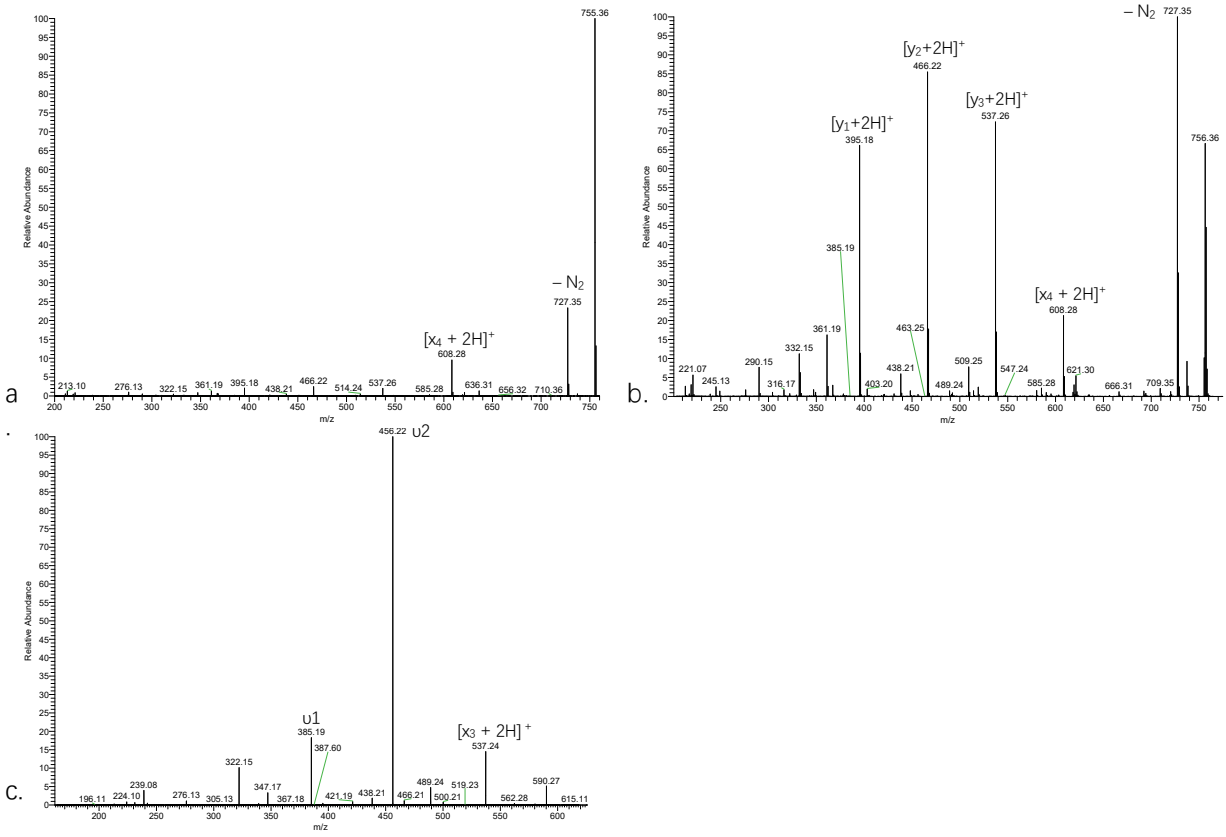

Figure 3.4. a. UVPD-MS^2 and b. CID-MS^2 of m/z 755, c. UVPD-MS^3 of m/z 608, of FAAA-*tet*-K sequence

The unique fragment ion series caused by crosslinks was found in the UVPD-MS^3 spectrum of the *m/z* 727 ion, as well. This series consisted of *m/z* 656, 585, and 514 peaks. At the same time, $[y_n + 2H]^+$ ions were found with minor intensities (*m/z* 580, 438, 367). Peaks at *m/z* 634, 492, and 421 were associated with the loss of aniline, and the loss of phenylhydrazine was indicated by the *m/z* 619 peak. For the UVPD-MS^4 spectrum of *m/z*

585, most peaks were associated with alanine residue loss, including *m/z* 468 and *m/z* 451 peaks that represented the loss of CO, H_2O, and NH_3 molecules besides three alanine residues. The peak of loss of alanine-lysine residues (*m/z* 368) were unique for FAAA-*tet*-K (Figure 3.5).

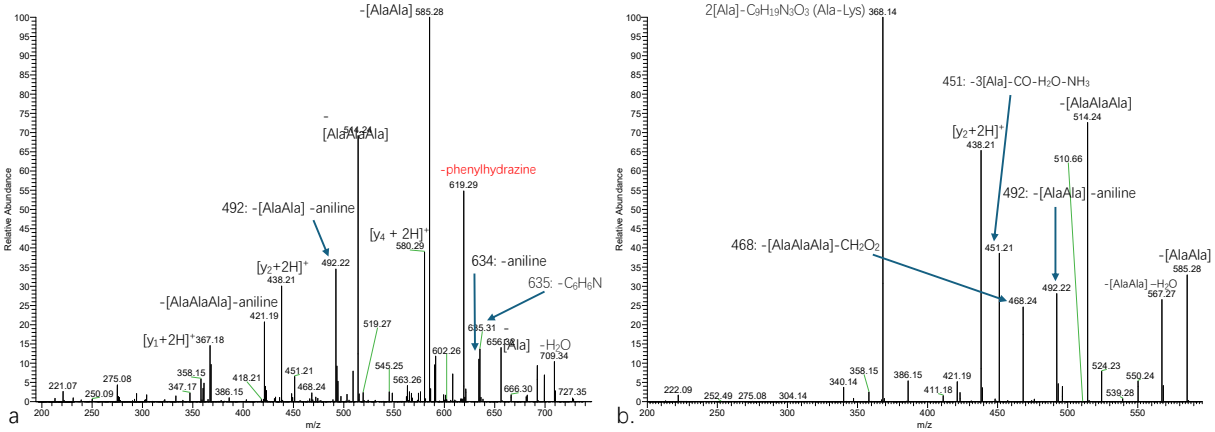

Figure 3.5. a. UVPD-MS^3 of m/z = 727. b. UVPD-MS^4 of m/z = 585 of FAAA-*tet*-K sequence.

Based on our primary guesses and calculation methods, we identified stable structures with low Gibbs energies. **F1**-**F3** indicated the molecules before the photodissociation of the tetrazole group, and **F2** was a product with th lowest Gibbs energy of -6.8 kJ mol^{-1}. **F4** and

F5 represented the structures after photodissociation, and structure **F4** exhibited the lowest Gibbs energy of 0 kJ mol^{-1} (Figure 3.6).

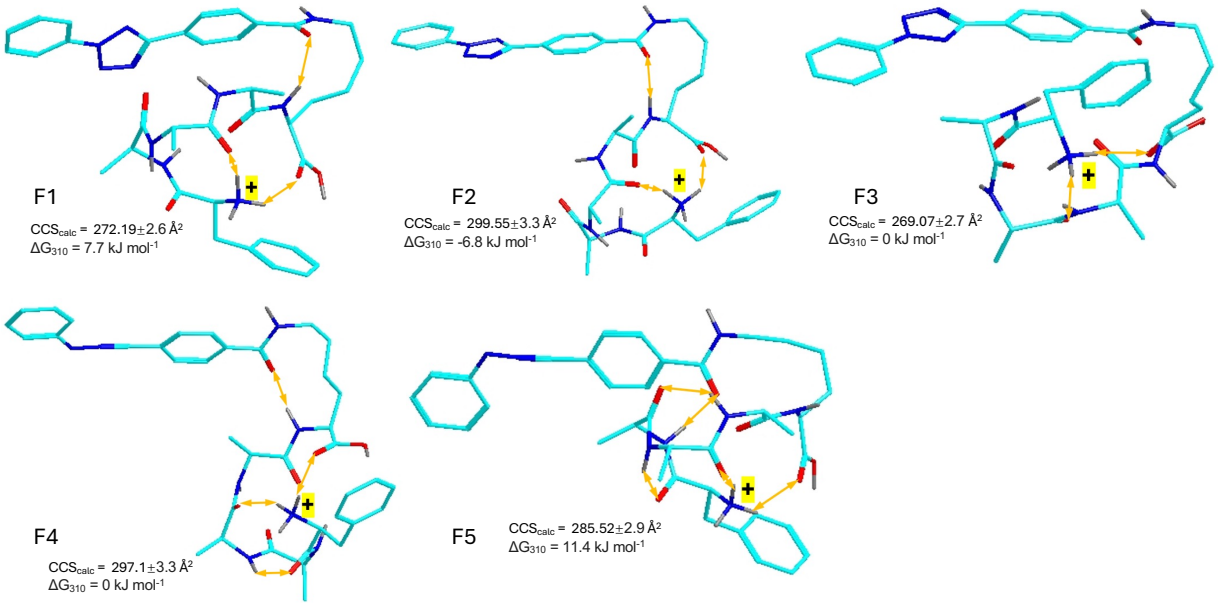

Figure 3.6. M06-2X/6-31+G(d,p) optimized structures of representative low-energy ions associate with FAAA-*tet*-K. Relative Gibbs energies are from M062X/def2qzvpp single-point energy calculations including B3LYP/6-31+G(d,p) zero-point corrections, and 310 K enthalpies and entropies. Atom color coding same as **Figure 3.3.**

3.3 Spectra Analysis and Calculations for WAAA-tet-K

Like spectra of YAAA-*tet*-K and FAAA-*tet*-K spectra, similar fragment ion series of backbone cleavage was found in MS^2 spectra of WAAA-*tet*-K as well. Similarly, $[y_3 + 2H]^+$ to $[y_1 + 2H]^+$ (*m/z* 537, 466, 395), $[x_4 + 2H]^+$ (*m/z* 608), and $-N_2$ (*m/z* 766) peaks had greater relative intensities in CID-MS^2 spectrum. In the UVPD-MS^3 spectrum of *m/z* 608, similar $[x_3 + 2H]^+$ (*m/z* 537) and υ_2, υ_1 fragments (*m/z* 456, 385) were also observed with considerable intensities (Figure 3.7).

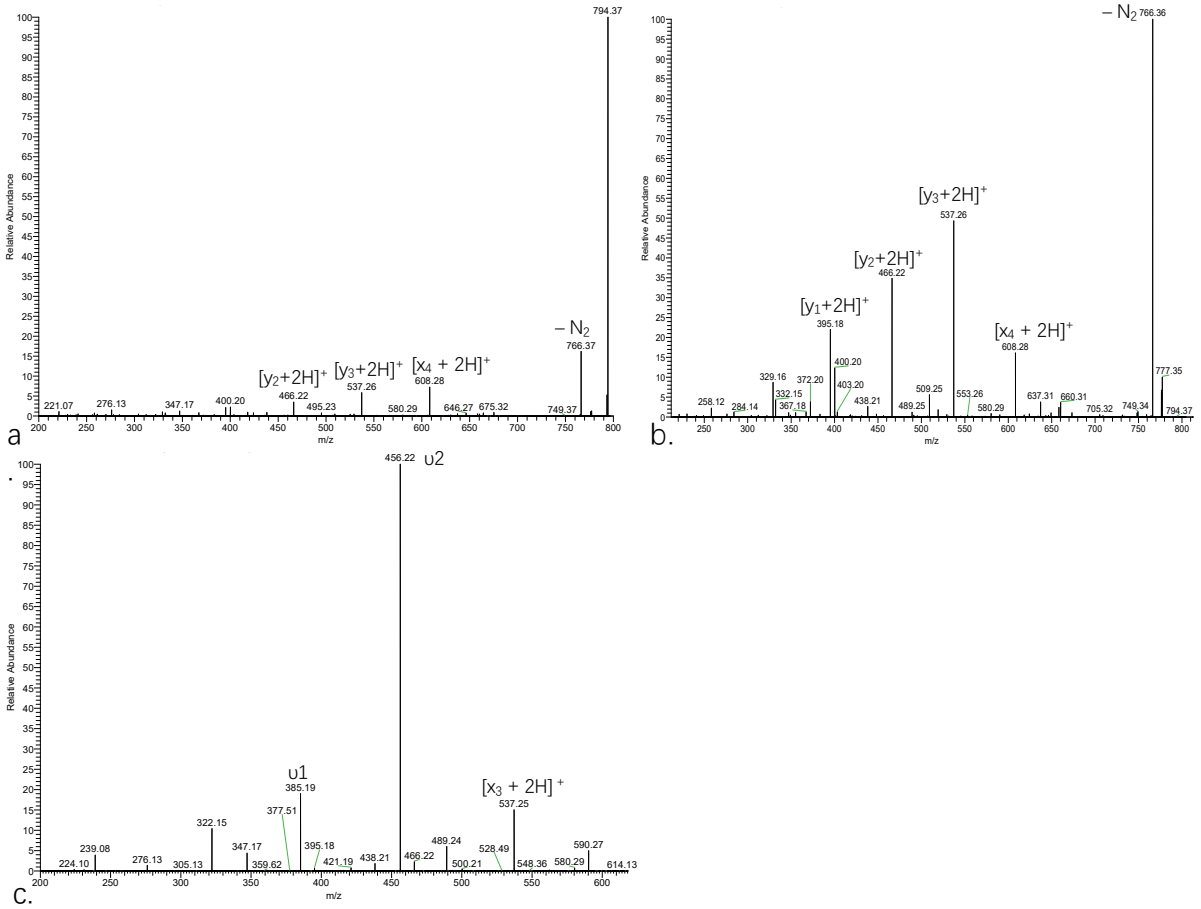

Figure 3.7. a. UVPD-MS^2 and b. CID-MS^2 of m/z = 794, c. UVPD-MS^3 of m/z = 608 of WAAA-*tet*-K sequence

The UVPD-MS^3 spectrum of the *m/z* 766 ion was more complex than the other MS^3 spectra. Due to the structural complexity of the tryptophan side chain, multiple unique fragments were found in this spectrum. The *m/z* 649 and 637 peaks indicated different fragmentations of the Trp side chain; based on accurate mass measurement of the fragments, these peaks represented the loss of C_8H_7N, and C_9H_7N. These peaks were unique for WAAA-*tet*-K spectra, compared to the spectra of other sequences. Crosslinking was indicated by the peaks with *m/z* 695, 624, and 553, that represented fragment ions formed by cleavage of internal alanine residues. The fragment ion at m/z 495 underwent a further loss of C_9H_7N and two alanine residues, according to accurate mass measurements. Like with the previous

sequences, loss of phenylhydrazine was found for WAAA-*tet*-K with *m/z* 658, and the peak at *m/z* 673 represented the loss of aniline. Fragment ions due to backbone cleavage of typical linear peptide were found with less intensity at *m/z* 580 ($[y_4 + 2H]^+$) and *m/z* 438 ($[y_1 + 2H]^+$) (Figure 3.8).

The UVPD-MS^4 spectrum of the *m/z* 624 ion, the peaks associated with internal alanine residue cleavage were found with high relative intensities. A neutral fragment with three alanine residues lost was indicated as the *m/z* 553 peak. The fragments of Trp side chain cleavage were shown as *m/z* 531 and *m/z* 495. The peak at *m/z* 438 was found with a lower intensity and was labeled as $[y_2 + 2H]^+$, indicating regular linear peptide backbone cleavage. In UVPD-MS^4 of the *m/z* 658 ion, only one peak associated with internal alanine loss was found at *m/z* 587, -1[Ala], but the loss of tryptophan side chain was shown as a peak at *m/z* 529 (Figure 3.8).

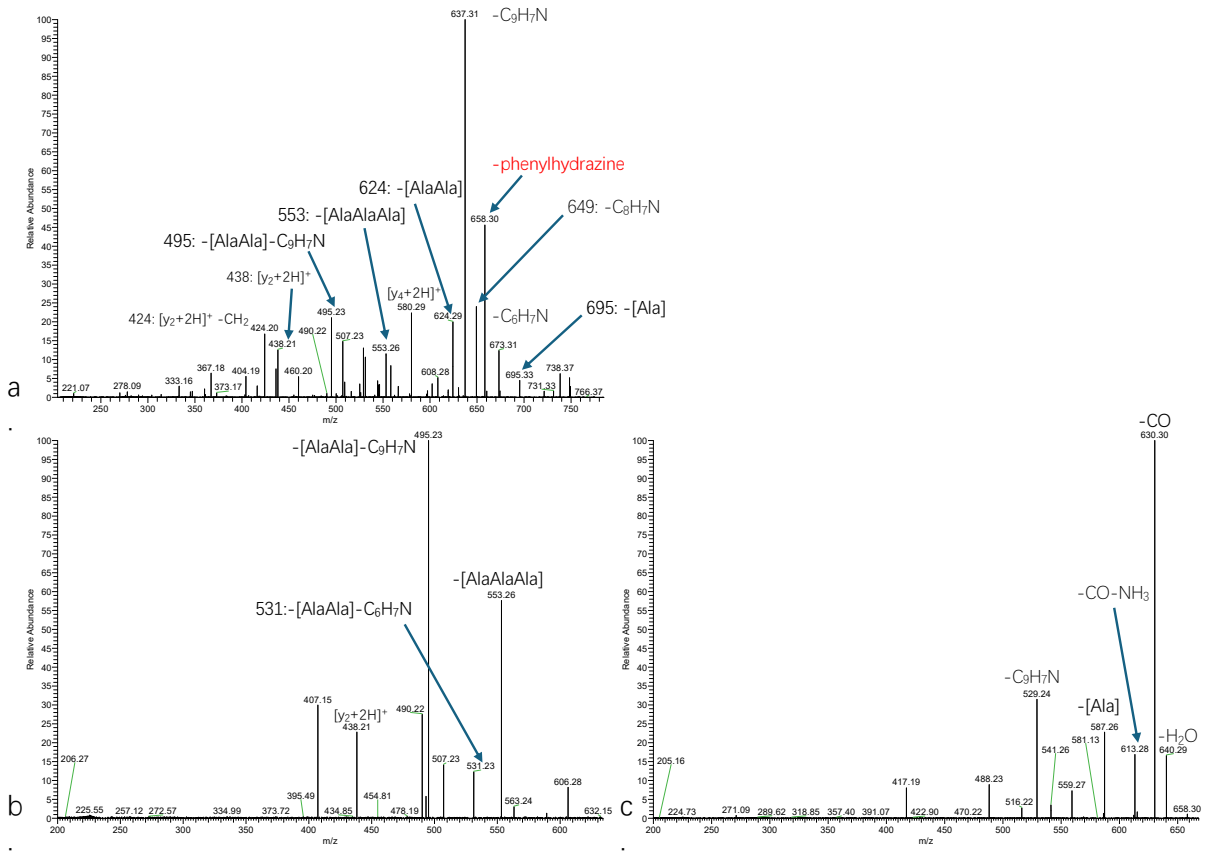

Figure 3.8. a. UVPD-MS^3 of m/z = 766. b. UVPD-MS^4 of m/z = 624, and c. UVPD-MS^4 of m/z = 658 of WAAA-*tet*-K

Based on our analysis about the spectra, initial guesses of ion structures were optimized calculations (Figure 3.9). **W1-W3** represented the structure before the photolysis reaction of tetrazole ring, of which **W1** was the one with the lowest relative Gibbs energy (0 kJ mol^{-1}). Intermediates after the loss of N_2 were represented by **W4** and **W5** of which **W4** was regarded as the more stable structure with a relative Gibbs energy of 0 kJ mol^{-1}.

For the products of the crosslinking reaction, proposed structures could be divided into two major groups of stereoisomers, and stable structures were found in both groups. As shown in Figure 3.10, the amide bond between tryptophan and alanine residue could be found either in a cis (**W7-W9**) or in a trans state (**W10-W12**) after the crosslinking reaction. Among the

structures with a cis amide, **W7** was labeled as the most stable structure with a Gibbs energy of -39 kJ mol^{-1}. Due to the protonation of the imine group that formed after the crosslink reaction, the positive charge was transferred to the alpha carbon of the tryptophan residue in the trans isomer. Among these structures, **W10** was the most stable one with a lowest relative Gibbs energy of -44 kJ mol^{-1}. Besides these structures, **W6** in Figure 3.9 represented another possible crosslinking product in which the nitrile imine group was protonated. The positive charge is located at the nitrile imine group and this structure had a relative Gibbs energy of 7.3 kJ mol^{-1} (Figure 3.9).

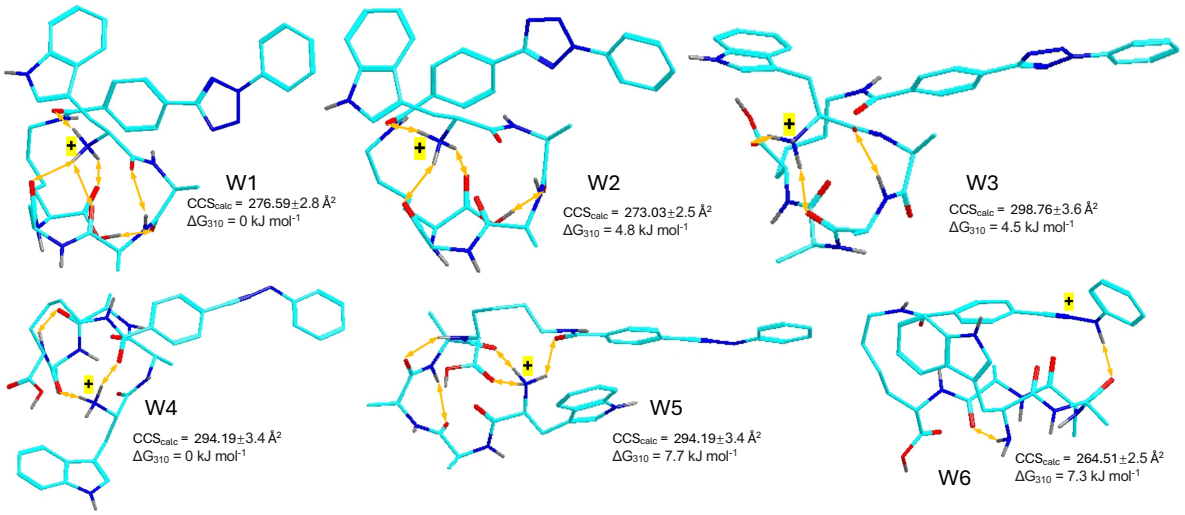

Figure 3.9. M06-2X/6-31+G(d,p) optimized structures **W1-W6** of representative low-energy ions associated with WAAA-*tet*-K. Relative Gibbs energies are from M062X/def2qzvpp single-point energy calculations including B3LYP/6-31+G(d,p) zero-point corrections, and 310 K enthalpies and entropies. Atom color coding same as **Figure 3.3.**

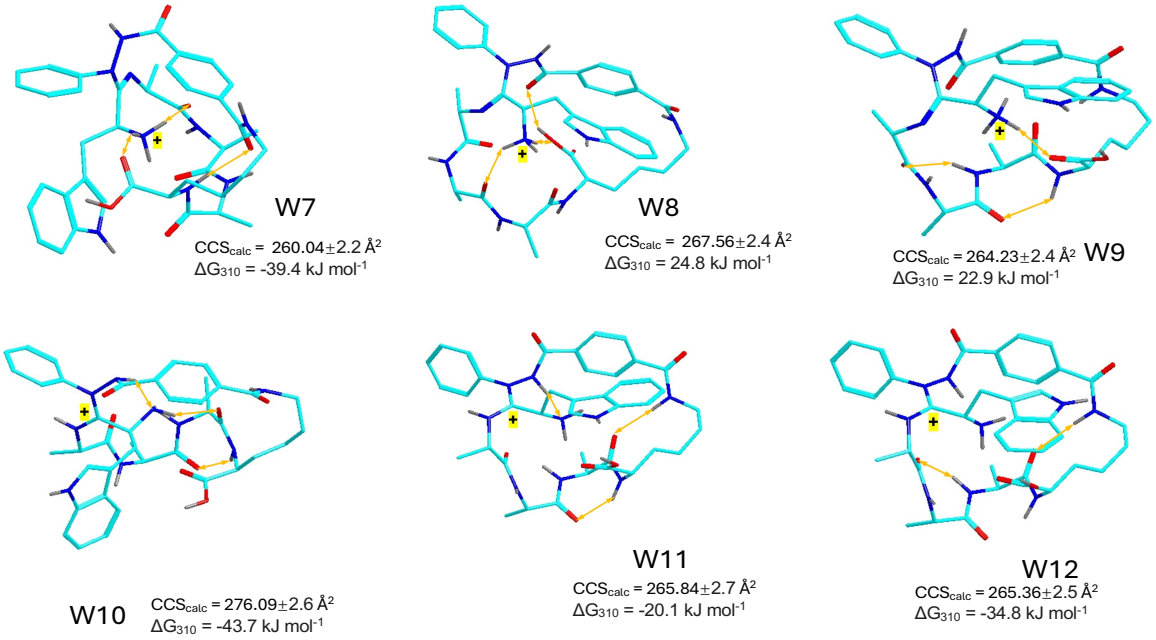

Figure 3.10. M06-2X/6-31+G(d,p) optimized structures **W7-W12** of representative low-energy ions associated with WAAA-*tet*-K . Relative Gibbs energies are from M062X/def2qzvpp single-point energy calculations including B3LYP/6-31+G(d,p) zero-point corrections, and 310 K enthalpies and entropies. Atom color coding same as **Figure 3.3.**

3.4 Crosslinking Yield Analysis

Based on the UVPD-MS^2 spectra, the crosslinking yields of the peptide-diaryltetrazole conjugates were obtained. Among these three sequences, WAAA-*tet*-K was found with the lowest crosslinking yield of 33% while FAAA-*tet*-K and YAAA-*tet*-K gave of 44% and 47%, respectively. However, in UVPD-CID-MS^3, WAAA-*tet*-K sequence was found with the highest crosslinking yield of 93%. The peaks that represented tryptophan side chain (*m/z* 637) and

phenylhydrazine (*m/z* 658) cleavage contributed greatly to the crosslinking yields.

	UVPD-MS^2 FTMS	UVPD-CID-MS^3 FTMS
FAAA-*tet*-K	44	74
YAAA-*tet*-K	47	86
WAAA-*tet*-K	33	93

Table 1. Crosslinking Yields

Chapter 4. Conclusion and Results

In UVPD-MS2 at 213 nm and CID, nitrile imine intermediates were generated through nitrogen-loss dissociation of the diaryltetrazole group. In peptide-2,5-diaryltetrazole conjugates, intramolecular cyclization reactions were observed between the nitrile imine intermediates and amino acid residues. With the analysis of MS spectra and following computational studies, structures of the cyclization crosslinking products were inferred, and reaction mechanisms were proposed. In this study, two types of crosslinking structures were proposed based on the analysis of spectra: the first one involved the C-N bond formation between the amide carbon and nitrile-imine nitrogen, while the other one was about the C-O bond formation between tyrosine side chain and carbon atom of tetrazole ring. Besides new crosslinking products, the unique sidechain cleavage of the tryptophan residue that was associated with crosslinking was reported. These results exhibited more possibilities of crosslinking reactions between amino acid and nitrile imine intermediate in the gas phase, but further research such as comparison between calculational and experimental collision cross section are still needed to confirm these structures and mechanisms in a more persuasive way.

Chapter 5. Acknowledgements

Research at the University of Washington was supported by the Chemistry Division of the U.S. National Science Foundation. Additional support from the Klaus and Mary Ann Saegebarth Endowment is gratefully acknowledged.

BIBLIOGRAPHY

1. Wong, S. S.; Wong, L. C. Chemical crosslinking and the stabilization of proteins and enzymes. *Enzyme and Microbial Technology* **1992**, 14 (11), 866-874. https://doi.org/10.1016/0141-0229(92)90049-T
2. Mattson, G.; Conklin, E.; Desai, S.; Nielander, G.; Savage, M. D.; & Morgensen, S. A practical approach to crosslinking. *Molecular Biology Reports* **1993**, 17 (3), 167–183. doi:10.1007/bf00986726
3. Decker, C. Photoinitiated crosslinking polymerisation. *Progress in Polymer Science* **1996**, *21*(4), 593–650. https://doi.org/10.1016/0079-6700(95)00027-5
4. Hennink, W., & Van Nostrum, C. Novel crosslinking methods to design hydrogels. *Advanced Drug Delivery Reviews* **2002**, *54*(1), 13–36. https://doi.org/10.1016/s0169-409x(01)00240-x
5. Tan, J., Luo, Y., Guo, Y., Zhou, Y., Liao, X., Li, D., Lai, X., & Liu, Y. Development of alginate-based hydrogels: Crosslinking strategies and biomedical applications. *International Journal of Biological Macromolecules* **2023**, *239*, 124275. https://doi.org/10.1016/j.ijbiomac.2023.124275
6. Luo, Y., Tan, J., Zhou, Y., Guo, Y., Liao, X., He, L., Li, D., Li, X., & Liu, Y. From crosslinking strategies to biomedical applications of hyaluronic acid-based hydrogels: A review. *International Journal of Biological Macromolecules* **2023**, *231*, 123308. https://doi.org/10.1016/j.ijbiomac.2023.123308
7. Alavarse, A. C., Frachini, E. C. G., Da Silva, R. L. C. G., Lima, V. H., Shavandi, A., & Petri, D. F. S. Crosslinkers for polysaccharides and proteins: Synthesis conditions,

mechanisms, and crosslinking efficiency, a review. *International Journal of Biological Macromolecules* **2022**, *202*, 558–596. https://doi.org/10.1016/j.ijbiomac.2022.01.029

8. Huang, R., Hua, J., Ru, M., Yu, M., Wang, L., Huang, Y., Yan, S., Zhang, Q., & Xu, W. Superb Silk Hydrogels with High Adaptability, Bioactivity, and Versatility Enabled by Photo-Cross-Linking. *ACS Nano* **2024**, *18*(23), 15312–15325. https://doi.org/10.1021/acsnano.4c05017
9. Jayachandran, B.; Parvin, T. N.; Alam, M. M.; Chanda, K.; & Mm, B. Insights on Chemical Crosslinking Strategies for Proteins. *Molecules* (Basel, Switzerland) **2022**, 27 (23), 8124. https://doi.org/10.3390/molecules27238124
10. Olszewska, A.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M. Chloroacetamide-Linked Nucleotides and DNA for Cross-Linking with Peptides and Proteins. *Bioconjugate Chemistry* **2016**, 27 (9), 2089–2094. https://doi.org/10.1021/acs.bioconjchem.6b00342
11. Herner, A.; Marjanovic, J.; Lewandowski, T. M.; Marin, V.; Patterson, M.; Miesbauer, L.; Ready, D.; Williams, J.; Vasudevan, A.; Lin, Q. 2-Aryl-5-carboxytetrazole as a New Photoaffinity Label for Drug Target Identification. *Journal of the American Chemical Society* **2016**, 138 (44), 14609–14615. https://doi.org/10.1021/jacs.6b06645
12. Brunner, J. NEW PHOTOLABELING AND CROSSLINKING METHODS. *Annual Review of Biochemistry* **1993**, *62*(1), 483–514. https://doi.org/10.1146/annurev.bi.62.070193.002411
13. Tureček, F. Covalent crosslinking in gas-phase biomolecular ions. An account and perspective. *Physical Chemistry Chemical Physics* **2023**, *25*(47), 32292–32304. https://doi.org/10.1039/d3cp04879a

14. Piersimoni, L., Kastritis, P. L., Arlt, C., & Sinz, A. Cross-Linking mass spectrometry for investigating protein conformations and Protein–Protein Interactions─A method for all seasons. *Chemical Reviews* **2021**, *122*(8), 7500–7531. https://doi.org/10.1021/acs.chemrev.1c00786

15. Uppadhayay, R. K., Kumar, A., Teotia, J., & Singh, A. Multifaceted chemistry of tetrazole. Synthesis, uses, and pharmaceutical applications. *Russian Journal of Organic Chemistry* **2022**, *58*(12), 1801–1811. https://doi.org/10.1134/s1070428022120090

16. Manzoor, S., Tariq, Q., Yin, X., & Zhang, J. Nitro-tetrazole based high performing explosives: Recent overview of synthesis and energetic properties. *Defence Technology* **2021**, *17*(6), 1995–2010. https://doi.org/10.1016/j.dt.2021.02.002

17. Wentrup, C., Damerius, A., & Reichen, W. Intramolecular cyclization of nitrile imines. Synthesis of indazoles, fluorenes, and aza analogs. *The Journal of Organic Chemistry* **1978**, *43*(10), 2037–2041. https://doi.org/10.1021/jo00404a043

18. Huisgen, R., Seidel, M., Sauer, J., McFarland, J., & Wallbillich, G. Communications: the formation of nitrile imines in the thermal breakdown of 2,5-Disubstituted tetrazoles. *The Journal of Organic Chemistry* **1959**, *24*(6), 892–893. https://doi.org/10.1021/jo01088a034

19. Su, Y., Zhao, Y., Chang, B., Zhao, X., Zhang, R., Liu, X., Huang, D., Wang, K., Huo, C., & Hu, Y. [3 + 2] Cycloaddition of para-Quinone Methides with Nitrile Imines: Approach to Spiro-pyrazoline-cyclohexadienones. *The Journal of Organic Chemistry* **2019**, *84*(11), 6719–6728. https://doi.org/10.1021/acs.joc.9b00434

20. Voronin, V. V., Ledovskaya, M. S., Gordeev, E. G., Rodygin, K. S., & Ananikov, V. P. [3 + 2]-Cycloaddition of in Situ Generated Nitrile Imines and Acetylene for Assembling of 1,3-Disubstituted Pyrazoles with Quantitative Deuterium Labeling. *The Journal of Organic Chemistry* **2018**, *83*(7), 3819–3828. https://doi.org/10.1021/acs.joc.8b00155

21. Wan, J., Nytka, M., Qian, H., Vu, K., Lemr, K., & Tureček, F. Nitrile imines as peptide and oligonucleotide Photo-Cross-Linkers in Gas-Phase ions. *Journal of the American Society for Mass Spectrometry* **2024**, *35*(2), 344–356. https://doi.org/10.1021/jasms.3c00379

22. Vlk, M., Wan, J., Nytka, M., Vu, T. N. K., Lemr, K., & Tureček, F. Photochemical and Collision-Induced Cross-Linking of ASP, GLU, ASN, and GLN residues in Peptide-Nitrile Imine conjugate ions in the gas phase. *Journal of the American Society for Mass Spectrometry* **2024**, 36(1), 175−186. https://doi.org/10.1021/jasms.4c00394

23. Zhang, J., Liu, J., Li, X., Ju, Y., Li, Y., Zhang, G., & Li, Y. Unexpected Cyclization Product Discovery from the Photoinduced Bioconjugation Chemistry between Tetrazole and Amine. *Journal of the American Chemical Society* **2024**, *146*(3), 2122–2131. https://doi.org/10.1021/jacs.3c11574

24. Marshall, D. L., Menzel, J. P., McKinnon, B. I., Blinco, J. P., Trevitt, A. J., Barner-Kowollik, C., & Blanksby, S. J. Laser photodissociation action spectroscopy for the Wavelength-Dependent evaluation of photoligation reactions. *Analytical Chemistry* **2021**, *93*(22), 8091–8098. https://doi.org/10.1021/acs.analchem.1c01584

25. Mann, M., Hendrickson, R. C., & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annual Review of Biochemistry* **2001**, *70*(1), 437–473. https://doi.org/10.1146/annurev.biochem.70.1.437

26. Larsen, M. R., Trelle, M. B., Thingholm, T. E., & Jensen, O. N. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *BioTechniques* **2006**, *40*(6), 790–798. https://doi.org/10.2144/000112201

27. Neagu, A., Jayathirtha, M., Baxter, E., Donnelly, M., Petre, B. A., & Darie, C. C. Applications of Tandem Mass Spectrometry (MS/MS) in protein analysis for biomedical research. *Molecules* **2022**, *27*(8), 2411. https://doi.org/10.3390/molecules27082411

28. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., & Dollinger, G. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proceedings of the National Academy of Sciences* **2000**, *97*(11), 5802–5806. https://doi.org/10.1073/pnas.090099097

29. Stahl, K., Graziadei, A., Dau, T., Brock, O., & Rappsilber, J. Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nature Biotechnology* **2023**, *41*(12), 1810–1819. https://doi.org/10.1038/s41587-023-01704-z

30. Song, W., Wang, Y., Qu, J., Madden, M. M., & Lin, Q. A photoinducible 1,3-Dipolar cycloaddition reaction for rapid, selective modification of Tetrazole-Containing proteins. *Angewandte Chemie International Edition* **2008**, *47*(15), 2832–2835. https://doi.org/10.1002/anie.200705805

31. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., DiNola, A., & Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

32. Řezáč, J., Fanfrlík, J., Salahub, D., & Hobza, P. Semiempirical quantum chemical PM6 method augmented by dispersion and H-Bonding correction terms reliably describes various types of noncovalent complexes. *Journal of Chemical Theory and Computation* **2009**, *5*(7), 1749–1760. https://doi.org/10.1021/ct9000922

33. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review. A, General Physics* **1988**, *38*(6), 3098–3100. https://doi.org/10.1103/physreva.38.3098

34. Grimme, S., Ehrlich, S., & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32*(7), 1456–1465. https://doi.org/10.1002/jcc.21759

35. Nickerson, C. J., Bryenton, K. R., Price, A. J. A., & Johnson, E. R. Comparison of Density-Functional Theory dispersion corrections for the DES15K Database. *The Journal of Physical Chemistry A* **2023**, *127*(41), 8712–8722. https://doi.org/10.1021/acs.jpca.3c04332

36. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry Chemical Physics* **2006**, *8*(9), 1057. https://doi.org/10.1039/b515623h

37. 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. *International Journal of Mass Spectrometry* **2015**, *390*, 24–27. https://doi.org/10.1016/j.ijms.2015.07.021.

38. Roepstorff, P.; Fohlman, J. Letter to the editors. *Biomedical Mass Spectrometry* **1984**, *11* (11), 601. https://doi.org/10.1002/bms.1200111109.