**Distribution Agreement**

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Sophia Shahin                                                                                  March 22nd, 2024

Bioinspired Chemical Approach for Synthesis of Allysine

by

Sophia Shahin

Dr. Monika Raj
Adviser

Chemistry

Dr. Monika Raj
Adviser

Dr. Richard Himes
Committee Member

Dr. Christina Gavegnano
Committee Member

Julio Medina
Committee Member

2024

Bioinspired Chemical Approach for Synthesis of Allysine

By

Sophia Shahin

Dr. Monika Raj

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2024

Abstract

Bioinspired Chemical Approach for Synthesis of Allysine
By Sophia Shahin

In this study, I present a novel chemical method for the selective conversion of dimethyl lysine residues to allysine. Inspired by a posttranslational modification catalyzed by lysyl oxidase, this chemistry offers the ability to selectively introduce aldehydes onto dimethyl lysine-containing molecules. Through numerous linear and cyclic peptide examples, I demonstrate the high chemoselectivity of this chemical method. I also expanded this chemical method for use on tertiary amine-containing small molecules. To demonstrate the utility of this chemical method, I diversified bioactive peptides with various affinity tags and fluorescent dyes. Finally, the potential of this method for developing cellular models for studying allysine-associated diseases is showcased.

Bioinspired Chemical Approach for Synthesis of Allysine

By

Sophia Shahin

Dr. Monika Raj

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2024

Acknowledgements

I would like to thank Dr. Monika Raj for her endless support and guidance throughout my time at Emory. She has been an incredible mentor, demonstrating what it means to be a passionate, driven, and successful woman in science while encouraging me to be the same.

I would like to thank Benjamin Emenike for his patience with answering my countless questions over the past three years and pouring equal effort into this project. Without his support, I would not have the privilege of conducting an honors thesis.

I would like to thank Raj lab graduate Dr. Rachel Wills and current Raj lab student Angele Bruce for teaching me confidence both inside and outside the lab. In addition, I would like to thank Angele Bruce for teaching me peptide synthesis and Dr. Wills for teaching me organic synthesis. I know it must not have been easy at times and to that I am eternally grateful for their patience.

I would like to thank current Raj lab students John Talbott and Riley Hughes for help revising my thesis.

I would like to thank all past and present Raj lab graduate students for their continuous mentorship throughout my time at Emory.

I would like to thank all past and present Raj lab undergraduate students for their friendship and support.

I would like to thank Dr. Christina Gavegnano for her kindness and for creating endless opportunities for me.

I would like to thank Dr. Richard Himes for his support both inside the classroom and during this thesis over the past year.

I would like to thank Julio Medina for his unwavering mentorship and to whom I credit much of my development as both a dancer and human.

I would like to thank all of my friends throughout my time at Emory who have become my home away from home. I thank them for being my biggest inspiration over the past four years. I am truly in awe at their passion, creativity, and drive.

I would like to thank Jack Swanson for encouraging me to not quit my chemistry major, even when it felt as though I was drowning, and for teaching me curiosity, not just about chemistry, but about the world.

And finally, I would like to thank my parents for their tireless support, endless encouragement, and infinite love. I would not be the woman I am today without them.

Table of Contents

# Introduction

Aldehydes, a class of organic chemical compounds characterized by a carbonyl group bonded to at least one hydrogen atom, play a significant role in biological systems, displaying diverse functionalities in naturally and non-naturally occurring processes. In addition to their considerable biological relevance, scientists across various disciplines have exploited aldehydes' electrophilic nature for many years. The utility of aldehydes lies in their ability to be functionalized using established chemistries. More clearly, by site-specifically functionalizing aldehydes with tags, including dyes, drugs, and affinity handles, the applications of aldehydes become abundant.[1] When applied to aldehyde-containing biological and small molecules, these applications include but aren't limited to synthesizing reversible covalent inhibitors and antibody-drug conjugates, two methodologies of increasing significance in current drug discovery efforts (Fig. 1). [2-4] The strategic functionalization of aldehydes also plays a critical role in certain cell imaging techniques. Chemically modified aldehydes can serve as probes for facilitating the visualization and elucidation of intricate cellular processes (Fig. 1).[3] Given the abundant applications of aldehydes in both biological and small molecules, it is imperative that there be robust methodologies for the chemical generation of aldehydes, advancing the study of diverse biological systems and development of novel therapeutics.
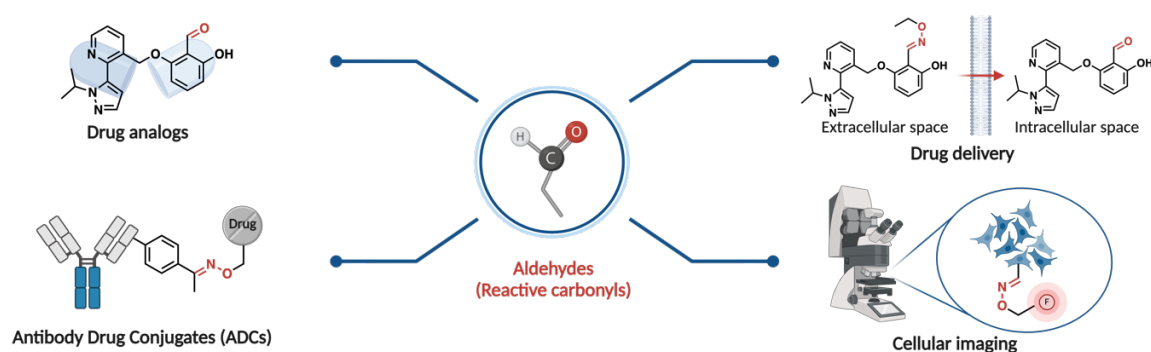


**Figure 1.** Known applications of aldehydes include drug analogs, drug delivery, antibody-drug conjugates, and cellular imaging. Figure re-printed from Angewandte Chemie International Edition publication.[5]

Despite the broad applications of aldehyde-containing molecules, current literature identifies a significant gap in the available chemical methods for the selective introduction of aldehydes onto peptides and proteins, two classes of biomolecules. In the human body, 20 canonical amino acids comprise the foundational building blocks for peptides, which have the capacity to form much larger proteins. When combined through condensation reactions, these naturally occurring amino acids give rise to peptide bonds, forming peptides and proteins of diverse lengths and sequences, each exhibiting various functions.[6] As previously mentioned, aldehyde-containing peptides and proteins demonstrate potential for numerous applications. However, the current known methods for introducing aldehydes onto these biological molecules face numerous limitations, impeding their applicability and efficacy. In a review article written by Spears et al., the currently available chemical methods for introducing aldehydes into biological molecules are discussed. The first relevant method capitalizes on the inherent reactivity of the N-terminus of these molecules. This aldehyde-introducing reaction is analogous to a vicinal diol cleavage, also known as the Malaprade reaction. Requiring an N-terminal serine or threonine residue, this reaction uses sodium periodate ($NaIO_4$) to promote the oxidative cleavage of either serine or threonine to produce an $\alpha$-oxo aldehyde (Fig. 2).[7] While this method is efficient and effective, an inherent limitation lies in its ability to introduce aldehydes exclusively at the N-terminus when serine or threonine is present. More clearly, it does not allow aldehydes to be introduced within the amino acid chain.
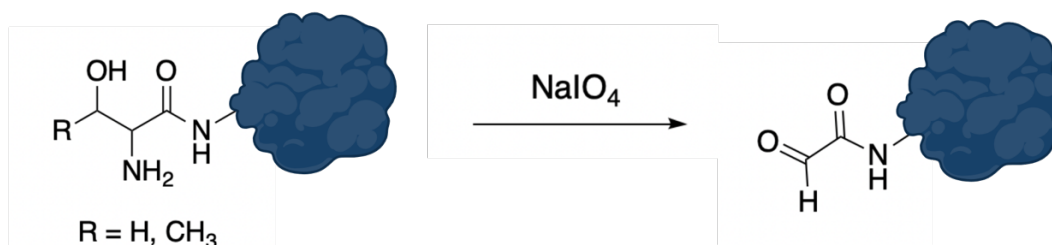


**Figure 2.** Reaction scheme of N-terminal serine or threonine cleavage to form N-terminal aldehyde using sodium periodate.

Another method for aldehyde generation on peptides and proteins reviewed by Spears et al. utilizes pyridoxal-5'-phosphate (PLP) as a reagent to promote a transamination reaction (Fig. 3).[5] This N-terminal-based methodology works to introduce aldehydes through the oxidation of amino acids at the N-terminus in a biomimetic manner. While this method is not limited by the presence of serine or threonine at the N-terminus, it does have several other limitations. First, when lysine, histidine, tryptophan, or proline are present at the N-terminus, several unwanted PLP-adduct byproducts are produced. Furthermore, this chemistry has been shown to display poor reactivity when isoleucine or valine occupies the N-terminus.[8] Due to these properties, this method is unviable for introducing aldehydes onto peptides and proteins in an efficient manner. Other methods have been reported to produce similar results at either the N or C-terminus,[9] exhibiting analogous capabilities but also limitations, thus perpetuating the overarching limitation of these methods characterized by their terminal-specific reactivity.
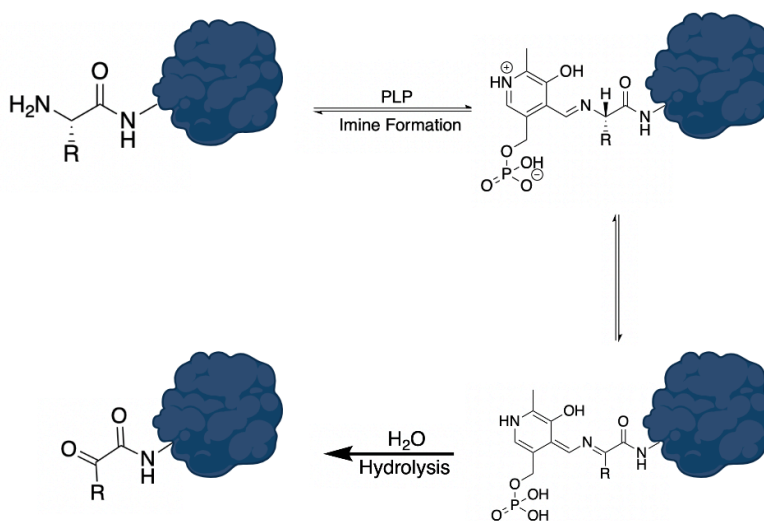


**Figure 3.** Reaction scheme of N-terminal aldehyde formation through transamination reaction with pyridoxal-5'-phosphate.

However, there are many naturally occurring enzymes that can selectively convert the side chains of amino acids to give rise to the formation of an aldehyde through posttranslational modifications (Fig. 4). For example, arginase and ornithine aminotransferase work together to selectively convert the side chain of arginine to glutamate semi-aldehyde.[10] Similarly, formylglycine-generating enzyme assists with the conversion of cysteine side chains to formyl glycine.[11] In addition, lysyl oxidase (LOX) can selectively convert lysine side chains to form allysine, an aldehyde that replaces lysine's primary amine.[12] However, after extensive research and to the best of my knowledge, no chemical methods are currently available to offer the selective introduction of aldehydes onto peptides and proteins through the conversion of an amino acid side chain. This gap in knowledge underscores the critical significance for the development of chemical methods to selectively introduce aldehydes onto peptides.
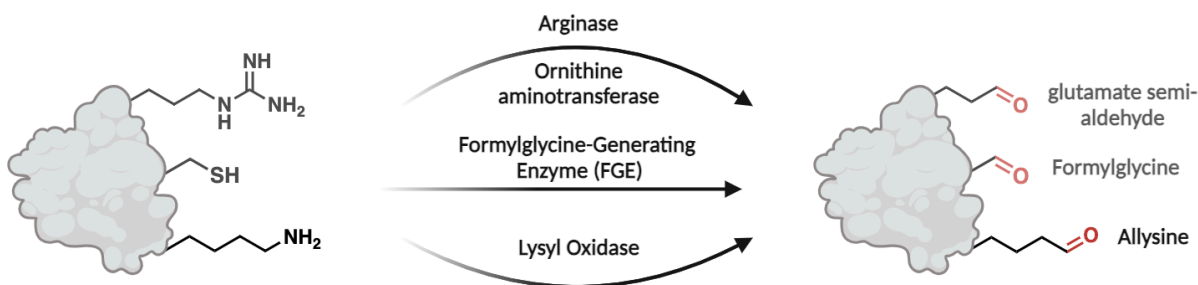


**Figure 4.** Known enzymes for introducing aldehydes on biological molecules include arginase, ornithine aminotransferase, formylglycine-generating enzyme, and lysyl oxidase. Figure re-printed from Angewandte Chemie International Edition publication.[5]

Lysyl oxidase (LOX), also known as protein-lysine 6 oxidase, is an enzyme involved in several cellular functions, specifically in the context of the extracellular matrix. Among its diverse enzymatic activities, LOX catalyzes the oxidation of lysine in elastin and collagen proteins. Elastin and collagen are two structural proteins responsible for tissue integrity and resilience. This oxidative transformation catalyzed by LOX results in the formation of an aldehyde on the lysine residue, allowing for the cross-linking of these fibrous proteins and contributing to their structural

stability (Fig. 5).[12a] Inspired by LOX and its ability to selectively oxidize lysine residues to aldehydes, we set out to mimic similar oxidative conditions. The generation of allysine, to the best of my knowledge, cannot be produced with other chemical methods and remains exclusive to LOX-catalyzed reactions.
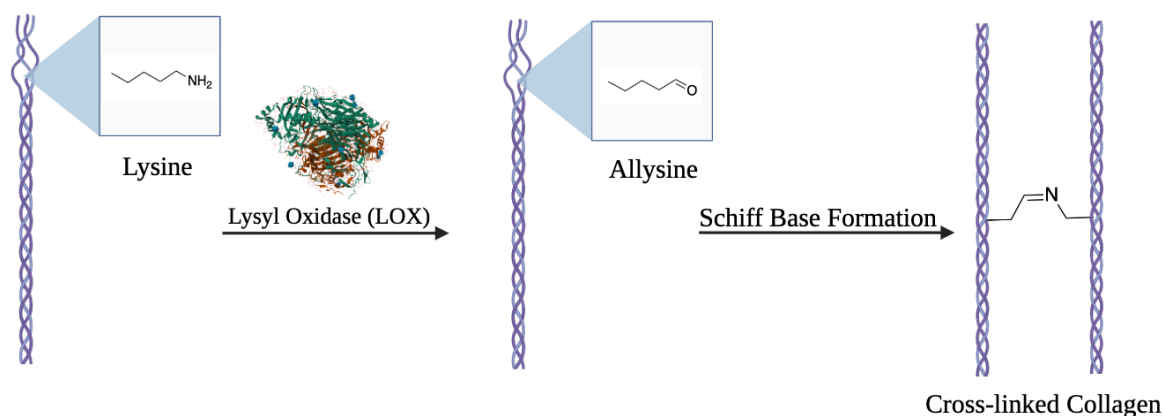


**Figure 5.** LOX oxidizes lysine to allysine, allowing for cross-links in proteins such as collagen.

This thesis proposes an approach aimed at the selective late-stage conversion of the side chain of dimethyl lysine into allysine while avoiding epimerization through what we have named Tertiary Amine Coupling Oxidation (TACO) (Fig. 6). In addition to introducing a novel synthetic method for the generation of allysine, I have demonstrated the applicability of this novel chemistry by introducing allysine on linear and cyclic peptides, some of which contain bioactive properties, highlighting TACO's pharmaceutical relevance. Moreover, I worked to diversify the allysine substrates with various tags, including dyes and affinity tags, to demonstrate the utility of this chemistry. These diversification reactions included hydroxylamine,[13] hydrazine,[14] thiolamine,[15] and reductive amination chemistries,[16] underscoring the versatility of this proposed oxidative reaction. This oxidative chemistry can be applied beyond biological molecules. In addition to peptides, I applied this chemistry to synthesize aldehydes from tertiary amine-containing small molecules. This application of our oxidative chemistry highlights the importance of this technique

in drug discovery. Aberrant overexpression of allysine has also been linked to various pathological conditions and diseases in literature. These diseases encompass fibrosis, cancer, and diabetes.[12b-d] Thus, we demonstrated the potential of this method to assist in the study of these allysine-mediated pathological molecular processes. This was achieved through mimicking the allysine posttranslational modification in cells.
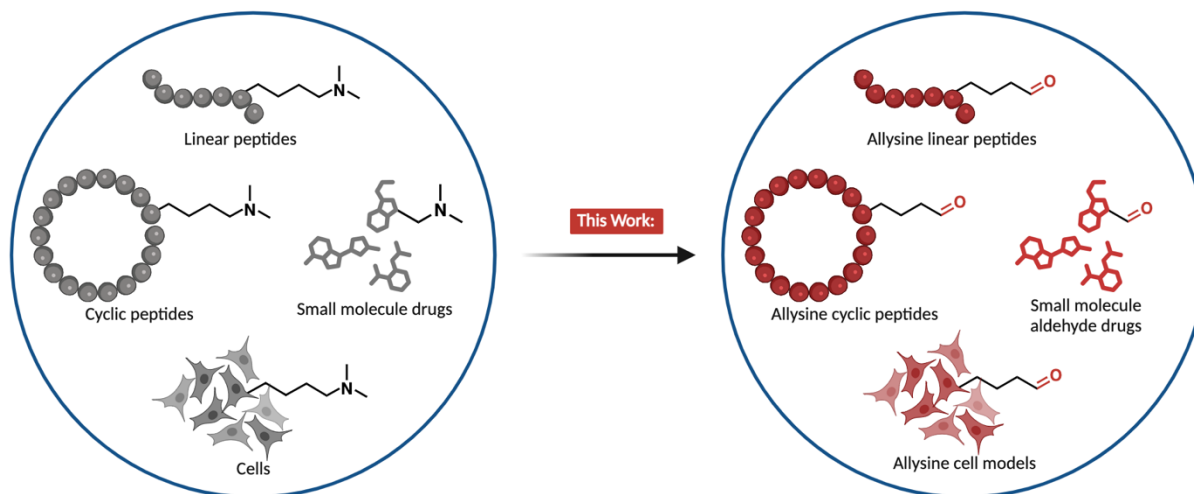


**Figure 6.** The scope of this work includes linear peptides, cyclic peptides, tertiary amine-containing small molecules, and tertiary amine cell models. Figure re-printed from Angewandte Chemie International Edition publication.[5]

# Results and discussion

### *Chapter 1: Background, inspiration, and approach*

This work is inspired by the posttranslational modification carried out by the enzyme lysyl oxidase. Lysyl oxidase is a known catalytic enzyme that oxidizes amines. Under normal physiological conditions, lysyl oxidase works to convert lysine residues into aldehydes in collagen and elastin proteins, contributing to tissue integrity. This aldehyde formation that occurs on the primary amine of lysine creates what is known as allysine.[12a] However, when the levels of allysine production are dysregulated, lysyl oxidase and the increased production of allysine have been linked to the pathogenesis of numerous disease states.[12b-d] Specifically, it has been observed that

the upregulation of allysine is associated with breast, head, and neck tumors. [10c] The inhibition of LOX has also been noted to cease tumor metastasis.[12e] By developing a chemical method to model the work done by LOX, scientists can better investigate its role in diseases.

The enzymatic pathway catalyzed by LOX uses an oxidative deamination reaction to form allysine. Central to this catalytic process is the involvement of the cofactor lysyl tyrosylquinone (LTQ). LTQ is strategically located in the active site of lysyl oxidase.[12a] While the exact mechanism of the oxidative deamination reaction is unknown, I propose that this reaction utilizes the primary amine on lysine to form a Schiff base (Fig. 7). The Schiff base formation is required for the subsequent steps in the mechanism, undergoing hydrolytic cleavage, and yielding the allysine product. When beginning to design our chemistry, we looked towards this reaction for inspiration to develop an effective method for selectively introducing allysine onto biological substrates containing lysine residues.
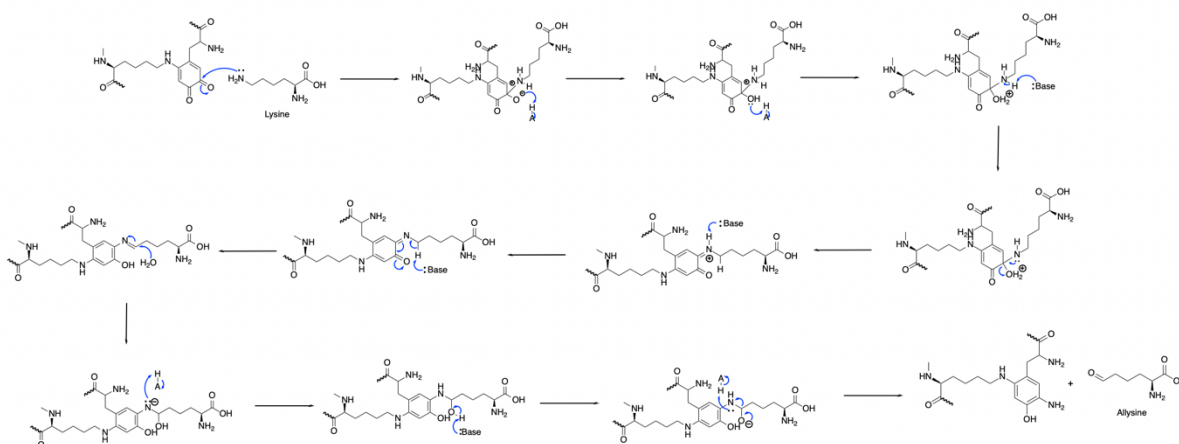


**Figure 7.** Proposed mechanism of allysine product using lysyl tyrosylquinone (LTQ), a cofactor located in the active site of lysyl oxidase.

In the initial stages of this research, my efforts were focused on optimizing the reaction conditions necessary to mimic the chemistry done by lysyl oxidase.[12a] Through Fmoc solid-phase peptide synthesis,[17] I synthesized a model tripeptide on rink amide resin with the sequence $H_2N$-

Phe-Lys-Val-CONH$_2$. The following oxidizing agents were screened due to their known, strong oxidizing abilities: NADP+, hydrogen peroxide, rose bengal, and selectfluor (Fig. 8). However, upon screening these oxidizing agents with the tripeptide, no reactivity was observed.
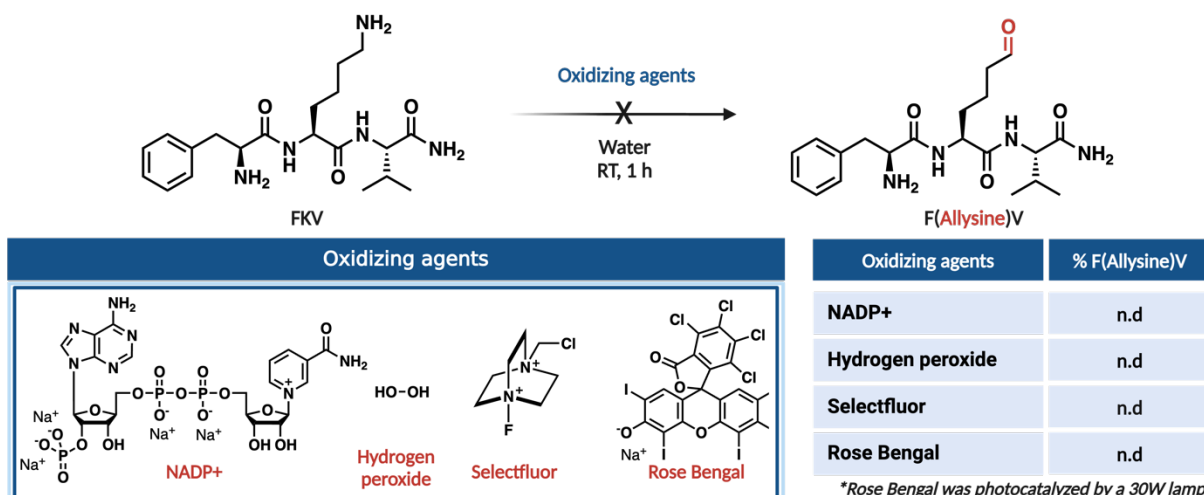


**Figure 8.** Optimization of the reaction with the following oxidizing agents: NADP+, Hydrogen Peroxide, Selectfluor, and Rose Bengal. No product was observed under any reaction conditions. Figure re-printed from Angewandte Chemie International Edition publication.[5]

In response to the observed experimental outcomes, I was prompted to look further into the reactivity of lysine for an oxidative reaction. In collaboration with Benjamin Emenike, computational studies were performed using Density Functional Theory (DFT) (Fig. 9). DFT is a quantum-mechanical method used to calculate molecules' electronic structure and reactivity.[18] Before performing the DFT calculations, we hypothesized that a methylated lysine may be more likely to react and, thus, form an aldehyde. Through these computational studies, we compared the reactivity of lysine in various methylation states: non-methylated (K), monomethylated (KMe$_1$), and dimethylated lysine (KMe$_2$). We began the computational studies by looking at the bond dissociation energies (BDE) for the initial reaction step that consists of a hydrogen ion removal. These calculations informed us that the BDE for the two methylated lysine states is increasingly more favorable than nonmethylated lysine. Specifically, the BDE was 96.4

Kcal/mol for monomethyl lysine and 94.9 Kcal/mol for dimethyl lysine, whereas canonical lysine has a bond dissociation energy of 96.7 Kcal/mol. This highlighted an initial energy barrier faced by canonical lysine that the other two methylation states avoided.

Next, we shifted our focus towards calculating the free energy of the iminium ion formation, the intermediate of the reaction. This revealed that the formation of the intermediate is significantly more favorable for the dimethyl lysine state (158 Kcal/mol) and the monomethyl lysine state (160 Kcal/mol) than with the canonical lysine state (165 Kcal/mol). Furthermore, when analyzing the electrostatic potential of the C=N bonds in this intermediate, we found that both canonical lysine and monomethyl lysine have a negative charge, -0.998 and -0.355, respectively. However, the electrostatic potential of the C=N bond observed for dimethyl lysine had a positive value of 0.209. This positive electrostatic potential demonstrates that the carbon in the C=N bond for the dimethylated lysine state is significantly more electrophilic than in the other methylation states. Thus, the subsequent hydrolysis reaction is more favored in the dimethyl lysine substrate. This computational insight laid the foundation for rationalizing the observed experimental reactivity and prompted us to focus on dimethyl lysine as our substrate rather than canonical lysine.
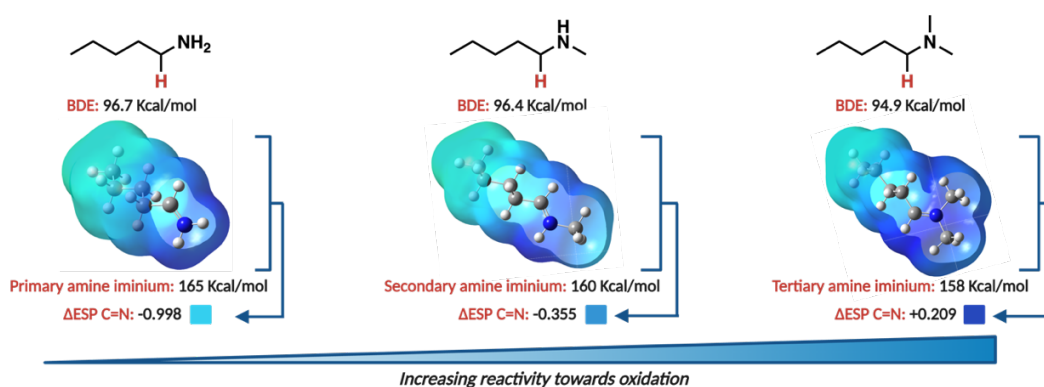


**Figure 9.** Density Functional Theory (DFT) calculations that demonstrate the reactivity of lysine, monomethyl lysine, and dimethyl lysine. Figure re-printed from Angewandte Chemie International Edition publication.[5]

Following these computational studies, experimental efforts were attempted to validate the hypothesized heightened reactivity of dimethyl lysine over canonical lysine. I attempted this on H$_2$N-Leu-Leu-Asp-Val-Leu-Leu-KMe$_2$-CONH$_2$, synthesized using Fmoc solid-phase peptide synthesis on rink amide resin with commercially available dimethyl lysine.[17] The four previously mentioned oxidizing agents were then screened again to test the reactivity of this peptide toward an oxidation reaction. Under these initial conditions, minimal conversion to the aldehyde product was observed with most of the oxidizing agents. Selectfluor emerged as the only exception, demonstrating a 5% conversion to aldehyde product. The observed reactivity trends from this initial screening of dimethyl lysine corroborate the computational predictions and reaffirmed our current hypothesis.

Upon reflection on why I wasn't observing higher conversion with selectfluor, I hypothesized that the low conversion was due to steric hindrance imposed by the selectfluor molecule. While selectfluor is the strongest of the screened oxidizing agents, it also has the bulkiest structure. [19] We hypothesized that to react with the tertiary amine on dimethyl lysine, selectfluor would have to act as a base and remove hydrogen from the sterically hindered carbon (Fig. 11). However, due to the bulkiness of selectfluor's structure,[19] I hypothesized that there may be a higher energy barrier for selectfluor to deprotonate the sterically hindered carbon, contributing to the low conversions we observed. Despite this barrier, the reaction with selectfluor allowed for slight conversion to aldehyde due to its ability to generate a 1,4-diazabicyclo[2.2.2]octane (DABCO) analogous base in situ.[17]

To further increase the reactivity of the iminium ion towards hydrolysis, we proposed that adding a known, strong base to the reaction would further promote the hydrolysis reaction step and allow for increased conversion to allysine. I proposed that pyridine, due to its ability to act as

a strong base and its planar structure, has the potential to further promote allysine formation while not imposing steric barriers.[20] Since selectfluor was the strongest oxidizing agent screened and the only one to portray reactivity, I decided to test pyridine (5 equivalents) with selectfluor (2 equivalents). Upon the addition of pyridine into the reaction vile with selectfluor and H$_2$N-Leu-Leu-Asp-Val-Leu-Leu-KMe$_2$-CONH$_2$ **1a**, I observed the formation of allysine at 85% conversion (Fig. 10). Additionally, I observed trace amounts of monomethyl lysine formation at 15% conversion.



| Oxidizing agents | Equivalence | % 1a : 2a : 3a |
|---|---|---|
| NADP+ | 5 | 100 : 0 : 0 |
| Hydrogen peroxide | 5 | 100 : 0 : 0 |
| Rose Bengal | 5 | 100 : 0 : 0 |
| Selectfluor | 5 | 95 : 5 : 0 |
| Selectfluor + Pyridine | 2 equiv. + 5 equiv. | 0 : 85 : 15 |

*Rose Bengal was photocatalyzed by a 30W lamp
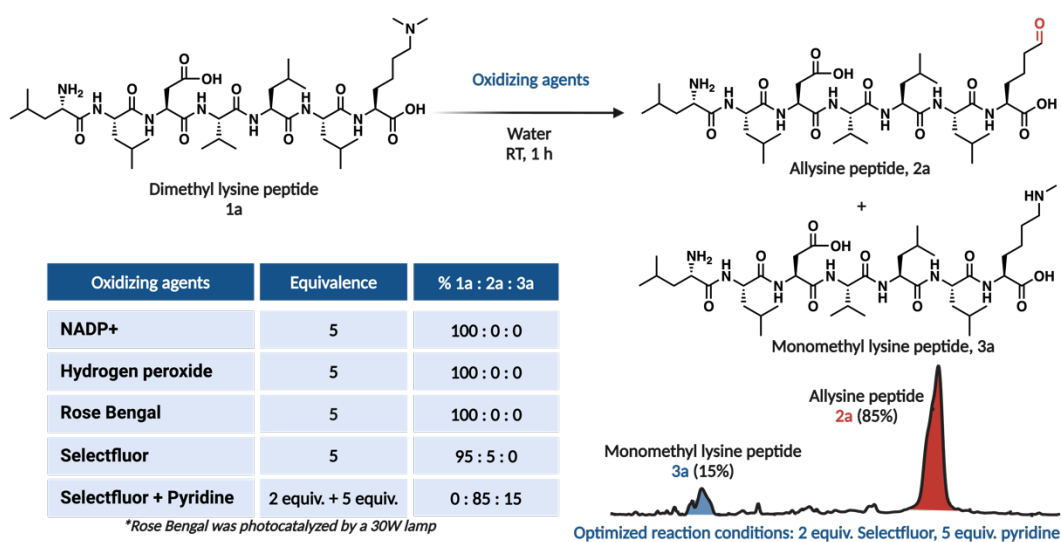
**Figure 10.** Optimization of oxidation chemistry with dimethyl lysine peptide using the following oxidizing agents: NADP+, Hydrogen Peroxide, Rose bengal, and selectfluor. Figure re-printed from Angewandte Chemie International Edition publication.[5]

To further support this observation and attempt to understand why monomethyl lysine was produced as a byproduct, I began to investigate what could be occurring mechanistically in the reaction (Fig. 11). The first step in the reaction involves the lone pair on nitrogen serving as a nucleophile that attacks the fluorine on selectfluor. In this scenario, fluorine serves as the electrophile due to the partial positive charge on fluorine caused by the neighboring positive nitrogen, pulling electron density away from fluorine.[19] The remaining lone pair on selectfluor

then removes a proton from one of the two methyl groups. While the selectfluor could react with the sterically hindered carbon, it is unfavored due to the bulky structure of selectfluor.[19] Due to this steric hindrance, the selectfluor is kinetically favored to react at either of the methyl groups on the tertiary amine. When a hydride is removed from one of the two methyl groups, the electrons shift down to form an iminium ion intermediate. Since water is both the solvent system and a reagent for this reaction, hydroxide can react with the imine, followed by a loss of formaldehyde, which leaves the monomethyl product. However, the choice to incorporate a strong base such as pyridine now allows for the hydride removal to occur at the sterically hindered carbon. This is because pyridine is an aromatic, planar molecule that can access the more sterically hindered hydrogens. The electrons on the OH functional group are then able to shift down and form the aldehyde product. To summarize, we believe the tertiary amine oxidation reaction is possible at both carbon sites and are competing reactions, confirming our experimental findings of both aldehyde and monomethyl product. However, the aldehyde product is usually observed in much higher conversion when using higher equivalents of pyridine.
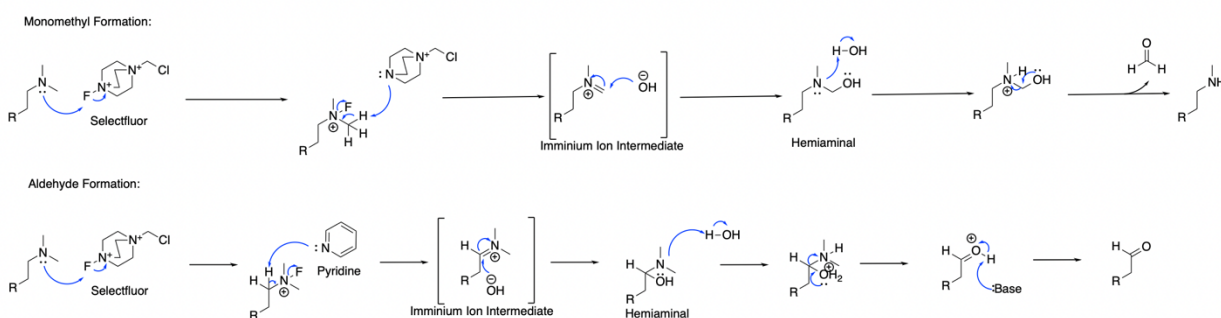


**Figure 11.** Hypothesized mechanism for observed experimental results including the formation of allysine and monomethyl lysine.

*Chapter 2: Substrate scope of synthesizing allysine in linear and cyclic peptides*

After obtaining the optimized reaction conditions for the formation of allysine on peptides, I began to explore the scope of this chemistry. Produced in a multitude of cells and glands, linear

peptides exhibit diverse functions and serve many applications in the human body. The classification of each amino acid within a peptide as hydrophobic, hydrophilic, acidic, or basic influences the function of the peptide. The function of a peptide can also be significantly governed by the exact sequence and the order of amino acids incorporated into the peptide chain. Due to these peptides' widely varying and sometimes unpredictable properties,[6] my first step in investigating the scope of this new chemical method was to test numerous linear peptides with varying lengths and amino acid compositions. These peptides were synthesized using Fmoc solid-phase peptide synthesis with wang resin,[17] varying the location of dimethyl lysine in each example. In addition to exploring various sequence compositions, I introduced a non-canonical amino acid, propargyl-glycine, into one example. Choosing to incorporate a non-canonical amino acid into one of the sequences is significant due to their frequent usage in peptide-based drugs, capacity to increase the enzymatic stability of biomolecules, and ability to increase cell permeability.[21] The peptide sequences chosen for this study include various bioactive properties, including anticancer, anti-inflammatory, antimicrobial, and antihypertensive.

The sequences tested included $NH_2$-Leu-Ile-KMe$_2$-Pro-Phe-COOH **1b**, $NH_2$-Arg-Leu-Pro-Tyr-Met-Pro-Tyr-Gly-Gly-KMe$_2$-Gly-COOH (anticancer) **1c**,[22] $NH_2$-Leu-Asp-KMe$_2$-Val-Asn-Arg-COOH (anti-inflammatory) **1d**,[23] $NH_2$-Phe-Ser-Asp-Lys-Lys-Ile-KMe$_2$-Lys-COOH (antimicrobial) **1e**,[24] $NH_2$-Ala-KMe$_2$-Gly-Ser-Lys-Ala-Phe-(Pra)-Ala-COOH **1f**, and $NH_2$-Val-Phe-KMe$_2$-Asn-Arg-COOH (antihypertensive) **1g**.[25] Following exposure to our optimized conditions, each of these linear peptides exhibited high conversion rates, ranging from 58-89%, of dimethyl lysine to allysine (Fig. 12). Additionally, minimal amounts of monomethyl lysine formation were observed, with no reactive byproducts of amino acids detected. The only additional

modification observed was in the peptide, NH$_2$-Arg-Leu-Pro-Tyr-Met-Pro-Tyr-Gly-Gly-KMe$_2$-Gly-COOH **1c**, where the methionine residue was oxidized into the sulfoxide product.

Remarkably, the peptide NH$_2$-Leu-KMe$_2$-Val-Pro-Phe-Leu-KMe$_2$-Val-Pro-Phe-COOH **1h**, containing two dimethyl lysine residues, underwent a notable conversion upon reaction with selectfluor and pyridine, producing a double-modified allysine product. Furthermore, I noted the presence of a single-modified allysine product with the formation of monomethyl lysine on the other dimethyl lysine residue. Moreover, a double monomethyl-lysine-modified product was identified. These findings underscore the remarkable chemoselectivity and specificity of this chemistry, herein establishing the first chemical method for converting dimethyl lysine to allysine.
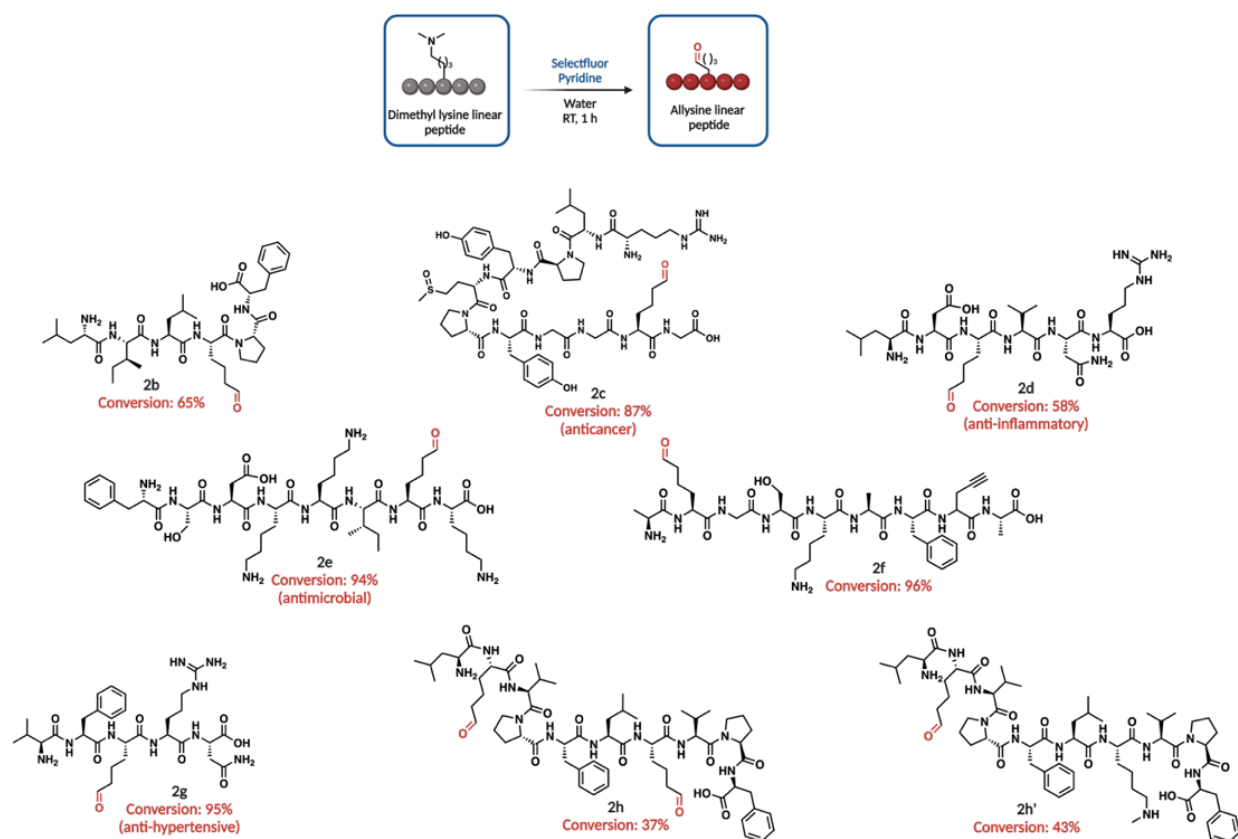


**Figure 12.** Scope for modification of linear peptides containing dimethyl lysine to allysine. Percent conversion determined by HPLC. Figure re-printed from Angewandte Chemie International Edition publication.[5]

To further validate this approach, I applied this chemistry towards various bioactive cyclic peptides. The extended application of this chemistry to cyclic peptides offers much relevance due to the benefits that cyclic peptides offer. The cyclic structure eliminates termini on peptides, making them more resistant to enzyme-based hydrolysis, which typically initiates cleavage at the termini. Moreover, the lack of amino and carboxy termini on cyclic peptides enhances their cellular permeability. Additionally, the innate rigidity of peptide macrocycles leads to lower entropy, facilitating more effective binding to target molecules.[26] Given these advantages, many peptide-based drugs are macrocycles, deeming it necessary that our chemistry be assessed for use on cyclic peptides.

The cyclic peptides chosen for this substrate scope study had varying amino acid compositions and ring sizes that spanned between 18-30 atom rings. The chosen cyclic peptides include Phepropeptin cyc-(LILKMe$_2$PF) **1i**, a proteasome inhibitor,[27] Vasopressin cyc-(CYFNQCPRKMe$_2$)-disulfide **1j**, an antidiuretic hormone (ADH),[28] Baceridin cyc-(IKMe$_2$WLLV) **1k**, a proteasome inhibitor inducing apoptosis in tumor cells,[29] cyc-(KSKRKMe$_2$PFES) **1l,** and Gramicidin cyc-(LKMe$_2$VPFLKMe$_2$VPF) **1m**, an antibiotic.[30] Upon exposure to the optimized reaction conditions, each example yielded allysine-cyclic peptides with high conversion rates ranging from 60-99% (Fig. 13). Similar to the linear peptides, trace amounts of cyclic monomethyl lysine peptide were also detected.

The Vasopressin peptide is formed as a result of disulfide bridge formation at the cyclization site. Remarkably, this peptide successfully generated the allysine cyclic peptide with an exceptionally high conversion rate, while avoiding ring opening due to disulfide cleavage. However, it should be noted that in the case of H$_2$N-Phe-Cys-Val-CONH$_2$, we observed that the presence of the free thiol on cysteine allowed for the oxidization to a disulfide bond, resulting in

dimer formation. The free thiol was able to be regenerated upon the addition of tris(2-carboxyethyl)phosphine (TCEP). Additionally, the formation of sulfenic acid was observed for free thiol-containing peptides. Another intriguing observation was in the case of Baceridin, cyc-(IKMe₂WLLV) **1k**, which incorporates Trp; fluorination of the Trp side chain was noted in the resulting allysine and monomethyl peptide. In the case of cyclic Gramicidin, upon reaction with pyridine and selectfluor, similar to its linear counterpart, a mixture of di-allysine-cyclic peptide, mono-allysine cyclic peptide, and monomethyl lysine cyclic peptide was formed.
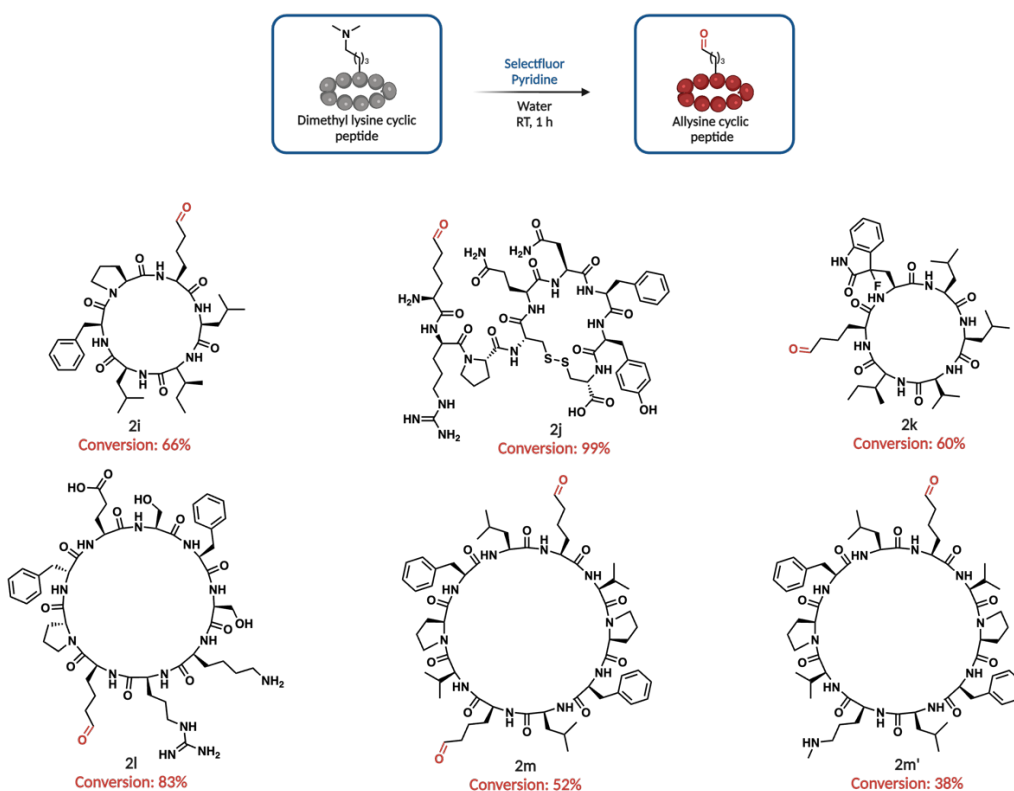


**Figure 13.** Scope for modification of cyclic peptides containing dimethyl lysine to allysine. Percent conversion determined by HPLC. Figure re-printed from Angewandte Chemie International Edition publication.[5]

## *Chapter 3: Substrate scope for modification of small molecules*

The presence of an aldehyde moiety on small molecules can serve many applications, specifically in the realm of drug discovery regarding synthesizing reversible covalent inhibitors.

In recent years, the United States Food and Drug Administration (FDA) has been approving an increasing number of covalent small molecule drugs. This can be attributed to the many benefits that covalent inhibitor drugs offer. Since covalent drugs can bind reversibly to target proteins, they are able to create a stronger bond in comparison to non-covalent inhibitors. Thus, reversible covalent inhibitors have been shown to increase the potency of drugs as well as increase the half-life. [31] Covalent inhibitors utilize a specific functional group, referred to as the warhead, to serve as an electrophilic binding site for target proteins. As shown in literature, aldehydes can serve as this electrophilic warhead through a reversible reaction, forming a Schiff base.[32] OXBRYTA (voxelotor) is a reversible covalent inhibitor that uses an aldehyde warhead which was approved by the FDA in 2019 for the treatment of sickle cell disease in both adults and pediatric patients.[33] Similarly, Roblitinib (FGF401) is currently undergoing clinical trials as a reversible covalent inhibitor using an aldehyde warhead that can reversibly bind to kinase FGFR4, a receptor linked to the metastasis of certain cancers.[34] Furthermore, Qiao et al. designed a series of aldehyde warhead-based reversible-covalent inhibitors for the main protease ($M^{pro}$) of SARS-CoV-2.[35] As seen by these examples, aldehyde warhead-based reversible-covalent inhibitors currently occupy a large portion of drug discovery research (Fig. 14).
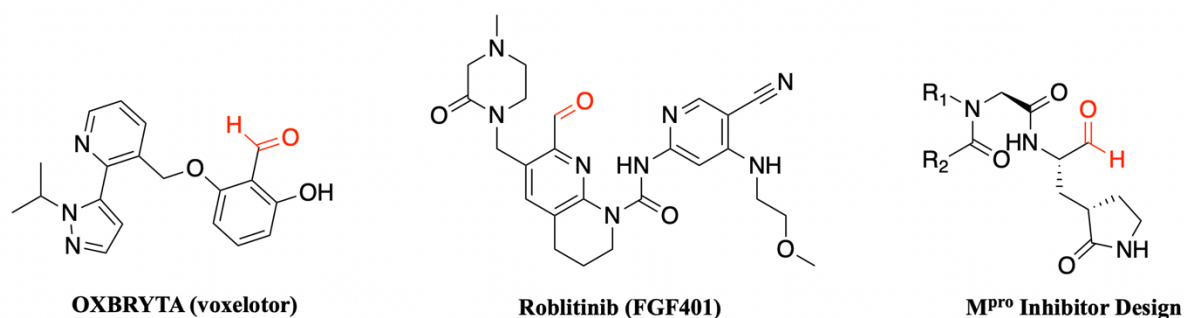


**Figure 14.** Examples of reversible covalent inhibitors using an aldehyde warhead including Oxbryta, Roblitinib, and $M^{pro}$ inhibitor.

Due to its high selectivity for the tertiary amine on dimethyl lysine residues, TACO offers similar success when applied to tertiary amine-containing small molecules. By introducing an aldehyde handle onto small molecules, we aimed to enable the conversion of non-covalent drug inhibitors into reversible covalent inhibitors.[35] I first explored the reactivity of N,N-dimethyl-4-phenylbutylamine **1n** with our optimized chemistry. I selected this molecule for its relatively simple chemical structure to gauge the chemistry's potential without interference from other functional groups. Upon exposure to our optimized macromolecular conditions, N,N-dimethyl-4-phenylbutylamine underwent conversion into 4-phenylbutanal **2n** with an 80% yield (Fig. 15). This preliminary reaction demonstrated promising potential for the chemistry of other, more complex small molecules. The next small molecule I investigated was N,N-dimethylbenzylamine **1o**, which similarly yielded benzaldehyde **2o** in 88% yield (Fig. 15). Furthermore, when applied to 4-(dimethylamino)-2-methoxyphenol **1p**, vanillin, the corresponding bioactive aromatic aldehyde **2p** commonly found in sweets [36] was obtained with an 84% yield (Fig. 15). Finally, I wanted to assess the applicability of this chemical method with a structurally complex small molecule, mimicking the complexity of small molecule drugs. Citalopram hydrogen bromide **1q**, a serotonin reuptake inhibitor and known antidepressant medication, contains a tertiary amine in addition to having a relatively complex structure, specifically containing an aromatic nitrile and a fused tetrahydrofuran ring.[37] Notably, the molecule's structure remained unaffected by the reaction conditions, and the corresponding aldehyde **2q** was obtained in a remarkable 93% yield, underscoring the potential for introducing covalent-binding small molecule drugs (Fig. 15). All small molecule aldehyde products were characterized using proton and carbon NMR spectroscopy, providing robust confirmation of their structures.
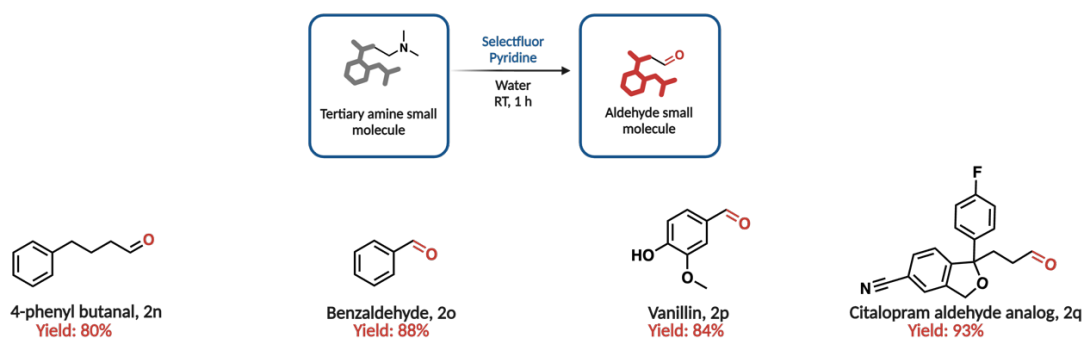
**Figure 15.** Scope of modification for tertiary amine-containing small molecules. Figure re-printed from Angewandte Chemie International Edition publication.[5]

## *Chapter 4: Late-stage diversification of peptides via allysine with varying tags*

To demonstrate the applicability of allysine-modified peptides, I further diversified the aldehydes through various known chemistries, including oxime, hydrazone, reductive amination, and thiazolidine chemistries (Fig. 16).[13-16] We initially opted to diversify the linear peptide containing two allysine groups with o-2-propynylhydroxylamine and O-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine in 2% acetic acid in water. Upon analysis, I observed nearly full conversion to double oxime-conjugation products at both allysine residues **4a-4b**. The mono-modified oxime product was not observed under either of the reaction conditions. Subsequently, a linear allysine peptide was modified with o-2-propynylhydroxyl amine, yielding a single oxime-conjugation product **4c** in high conversion, specifically 98%. Using biotin hydrazide and dansyl hydrazide, I was able to observe mono-modified hydrazone products with good conversions, 57% **4d** and 73% **4e**, respectively. Employing established reductive amination chemistry with propargyl amine led to the formation of an alkyne-ligated peptide with 67% conversion **4f**. Lastly, I observed a thiazolidine-modified product **4g-4h** in full conversion by using cysteine methyl ester in 2% acetic acid in water.
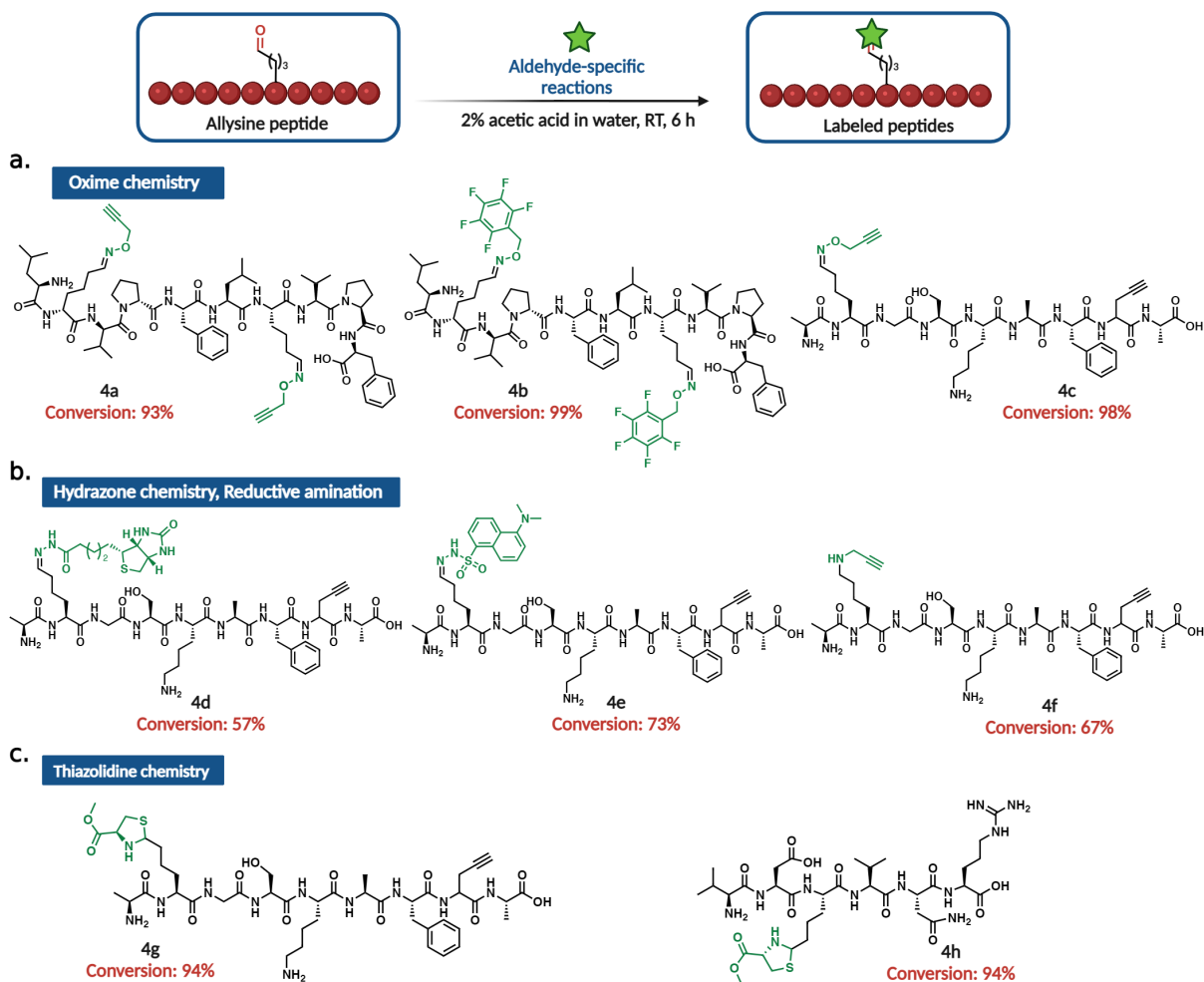
**Figure 16.** Late-stage diversification of linear peptides containing allysine. The reactions performed include various oxime, hydrazone, reductive amination, and thiazolidine chemistries. Figure re-printed from Angewandte Chemie International Edition publication.[5]

## *Chapter 5: Synthesis of allysine on solid support*

The synthesis of allysine in the solution phase presents a promising, novel chemistry. However, a potential limitation of this chemistry became evident during optimization and substrate scope assessment: the highly oxidative conditions generated by selectfluor lead to undesired modifications of certain amino acids, including tryptophan, histidine, cysteine, and methionine. As commercially available, these amino acids contain protecting groups on their side chains. Specifically, tryptophan features a Boc protecting group, while histidine and cysteine are protected by trityl groups; methionine lacks a side chain protecting group altogether.[38] To overcome these

unwanted modifications, we hypothesized that conducting a reaction in the solid phase while the peptide remained bound to resin, with the side chain protecting groups intact, could prevent the undesired side chain modifications. [17] However, experimental observations revealed that the resin-bound peptides failed to undergo conversion to form allysine. We hypothesized that the choice of resin may be the underlying reason for these results.

Polystyrene-based resins, such as rink amide and wang resins, are the most commonly used resins during Fmoc solid-phase peptide synthesis (Fig. 17). [17] However, due to the requirement of water as both a solvent and reagent in TACO chemistry, we suspect that the hydrophobic nature of polystyrene hindered the resin's ability to swell in water, thereby preventing the peptide from exposing its side chains to the reaction conditions required for allysine. Following extensive research into various hydrophilic resins, we opted to investigate rink amide PEGA resin, known for its compatibility with water (Fig. 17). The rink amide PEGA resin incorporates a polyethylene glycol moiety, imparting hydrophilic properties that promote the swelling of the resin in water.[39] As a result, the peptide can better interact with the reaction conditions and overcome the limitations encountered with the traditional polystyrene resins.



Polyethylene Glycol          Polystyrene
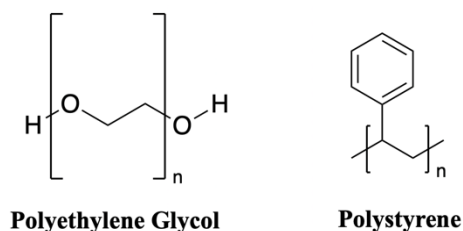
**Figure 17.** Structure of polystyrene and polyethylene glycol repeating unit.

To evaluate the efficacy of this approach, three tetrapeptides were synthesized using Fmoc solid-phase peptide synthesis on rink amide PEGA resin with one oxidizable residue on each (Fig. 18a).[17] These peptides are FKMe$_2$HV **2r**, FKMe$_2$WV **2s**, and FKme$_2$CV **2t**, with the oxidizable

residues being histidine, tryptophan, and cysteine. Remarkably, I observed the allysine peptides in high conversion without any modification to the oxidizable byproducts. Observing the initial success of this chemical method on these tetrapeptides, I synthesized a longer peptide, FKMe$_2$RWGHVCSMV **2u**, encompassing all the oxidizable residues: tryptophan, cysteine, histidine, and methionine. Analysis revealed the formation of the allysine product without the occurrence of fluorinated and oxidized byproducts, apart from methionine, whose commercial structure does not contain a side chain protecting group.

I further diversified this longer peptide with o-2-propynylhydroxylamine hydrochloride, 2-azasprio[3.5]nonane hydrochloride, and methylhydrazine to demonstrate the versatility of this chemistry (Fig. 18b). [40] Similarly, I observed high conversions without any modifications of side chains to the derivatized peptides **3u-5u**. These products were analyzed by LC/MS, confirming their structures and purity. Overall, using rink amide PEGA resin coupled with TACO offers a promising approach for synthesizing allysine peptides, overcoming the challenges associated with oxidation-sensitive residues.
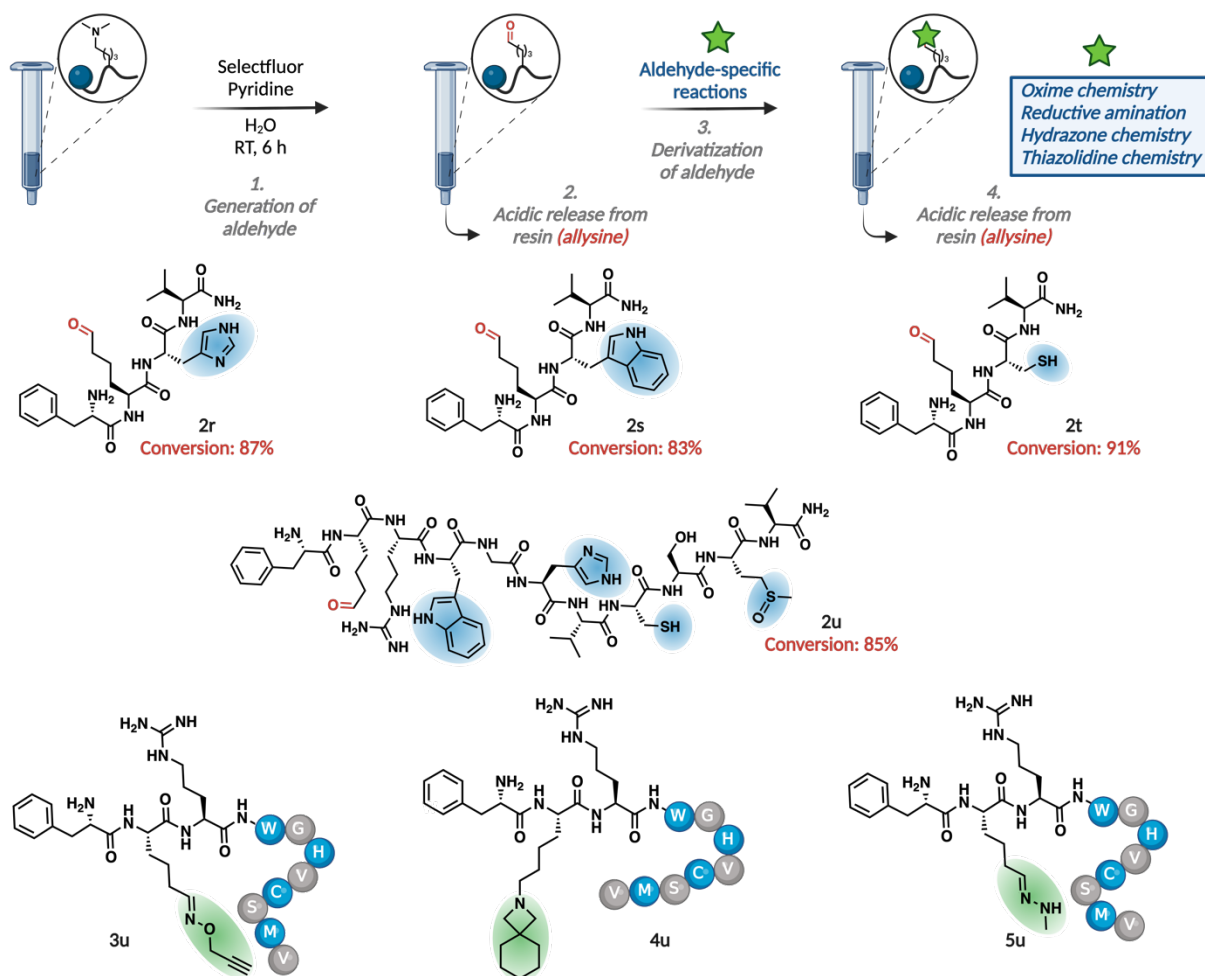
**Figure 18.** (a) On-resin synthesis of allysine linear peptides containing oxidation-sensitive amino acids including histidine, tryptophan, cysteine, and methionine. (b) Late-stage diversification of the on-resin allysine peptides using oxime, reductive amination, and hydrazone chemistries. Figure re-printed from Angewandte Chemie International Edition publication.[5]

## *Chapter 6: Cellular imaging applications of TACO reaction*

### Cellular imaging of peptide-fluorophore conjugate

A potential application of this chemistry is the ability to conjugate fluorescent molecules onto the biomolecules. Fluorescent probes are commonly used to assist in the detection of peptides, proteins, or drugs inside cells to gain information about their location as well as to observe biological processes.[3] To further demonstrate the applications of allysine-modified peptides, in collaboration with Benjamin Emenike, we sought to demonstrate allysine as a handle for selective

labeling with a fluorophore, enabling the tracking of peptides within cells. This was achieved through the conjugation of AlexaFluor-647 hydroxylamine dye to an allysine-modified linear peptide **2i** (Fig. 19). Subsequently, the peptide-fluorophore conjugate was incubated for 12 hours with T-47D cells. Confocal imaging analysis revealed the successful delivery of the peptide-fluorophore conjugate into the cells, as evidenced by the fluorescence signals within the cell (Fig. 20). In contrast, the negative control lacking the peptide-fluorophore conjugate did not exhibit any fluorescence, confirming the specificity of the labeling approach.
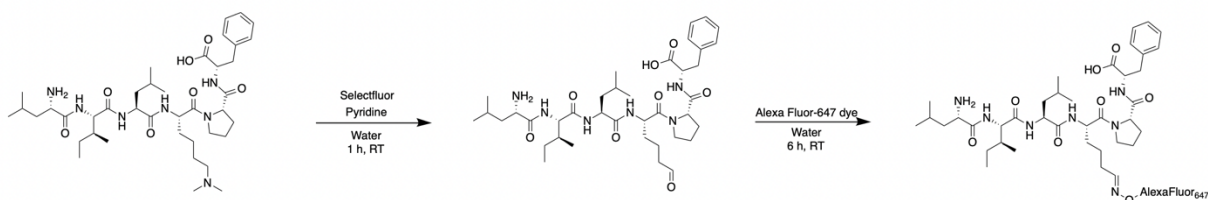


**Figure 19.** Reaction scheme for allysine linear peptide formation from dimethyl lysine starting material in addition to scheme for allysine and Alexa Fluor 647 conjugation.
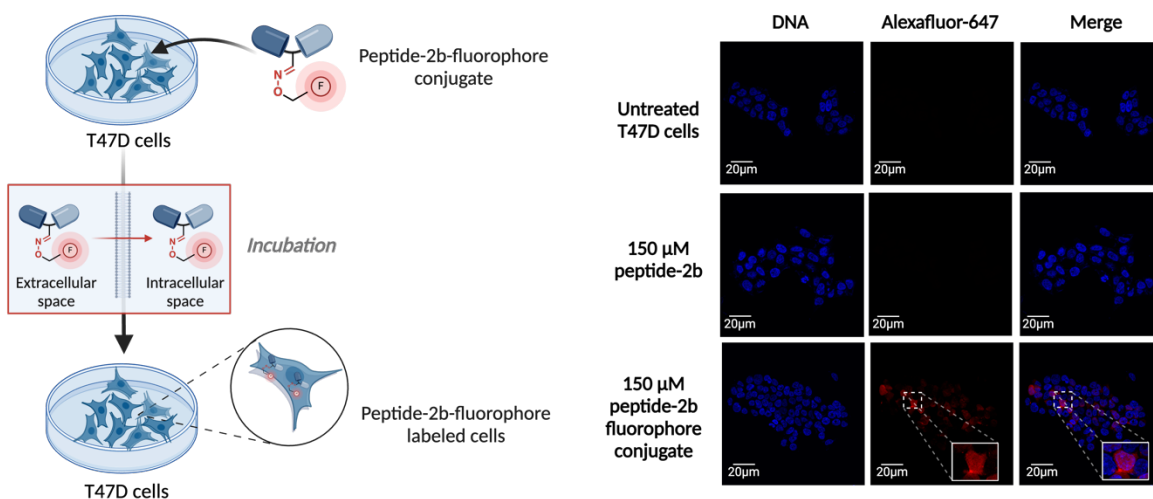


**Figure 20.** Tracking of peptide-fluorophore conjugate inside breast cancer cell line, T-47D cells. Figure re-printed from Angewandte Chemie International Edition publication.[5]

**Cellular imaging of allysine using aldehyde-reactive fluorophore**

A significant application of our chemistry lies in its capability to generate cellular models for studying allysine-associated diseases. Given the correlation between allysine overexpression

and its implication in pathological conditions such as cancer, fibrosis, and diabetes, it is imperative to develop molecular models to allow for the study of these diseases. However, model systems are relatively unavailable for overexpressing allysine, rendering this research area understudied.[41-43] Here, with contributions from Benjamin Emenike, I present a demonstration for the chemical generation of allysine in T-47D cells, a breast cancer cell line, portraying one of the many applications of TACO. First, the cells were exposed to reductive amination conditions in a dose-dependent manner to create dimethyl lysine in cells. Subsequently, dimethyl lysine proteins in cells were modified using selectfluor and pyridine to create allysine proteins. These cells were treated with AlexaFluor-647 hydroxylamine dye, creating a dose-dependent fluorescent signal response, indicating the varying concentrations of allysine proteins in different cell populations (Fig. 21). The presence of allysine was confirmed as the cause of fluorescence, as seen by our negative control, which showed no fluorescent signal in unmodified cells and dimethyl lysine-containing cells. Further visualization of the allysine adducts would enhance our comprehension of the cellular processes associated with allysine. This approach shows promise for elucidating the role of allysine in disease development and progression and assisting researchers in the exploration of the molecular mechanisms associated with allysine-linked pathologies.
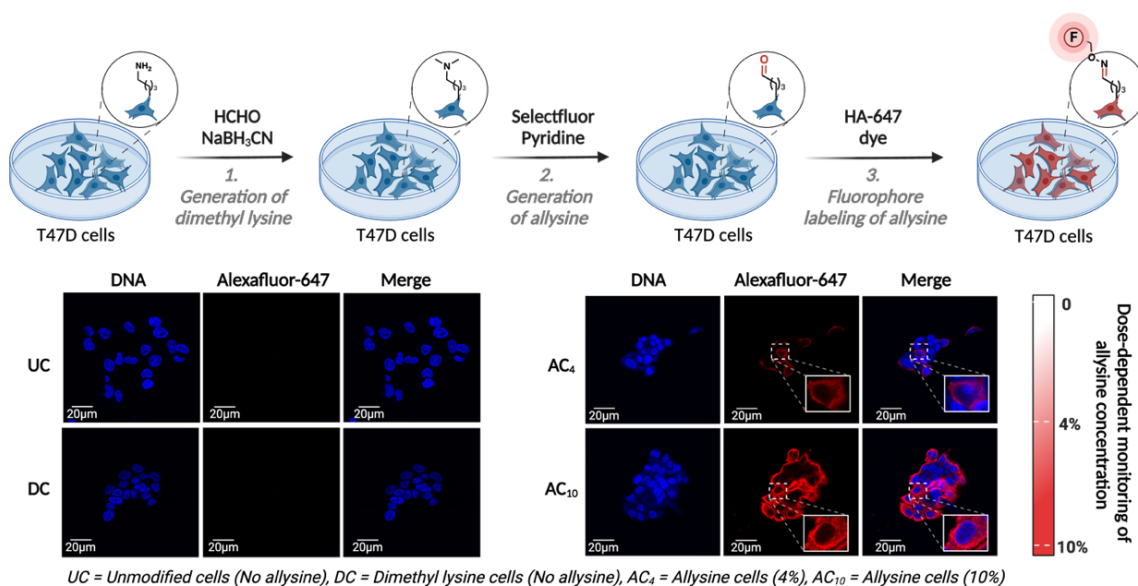
**Figure 21.** Cellular imaging of concentration-dependent allysine proteins. Figure re-printed from Angewandte Chemie International Edition publication.[5]

## Conclusion and Future Directions

In this work, we have introduced a biomimetic approach for the generation of allysine, an aldehyde functional group naturally formed through posttranslational modifications with the assistance of the enzyme lysyl oxidase. Inspired by lysyl oxidase, we have optimized oxidative reaction conditions to develop the first known chemical method for selectively introducing allysine onto biomolecules. This is accomplished by the selective oxidation of dimethyl lysine, replacing the tertiary amine with a formyl group. This highly chemoselective reaction exclusively modifies dimethyl lysine residues in the presence of all other amino acids. This is the first chemical approach capable of the selective conversion of an amino acid side chain to an aldehyde.

Throughout this thesis, the utility of this novel chemical method has been demonstrated through the synthesis of allysine-modified linear and cyclic peptides, followed by the late-stage diversification of therapeutically relevant linear peptides with various tags including hydroxylamine, hydrazine, thiazolidine, and reductive amination chemistries. Moreover, we have

extended this chemical method for use on tertiary amine-containing small molecules. Furthermore, our solid-support chemistry has enabled the introduction of allysine without modification of oxidation-sensitive residues such as histidine, tryptophan, and cysteine. Lastly, we portrayed the utility of this approach for the generation of cellular models of overexpressed allysine, demonstrating its potential applications in various fields. The robust chemoselectivity and mild reaction conditions make this chemistry highly useful for applications in medicinal chemistry, proteomics, biochemistry, material chemistry, and the synthesis of antibody-drug conjugates.

An additional application of this research is the investigation of cytotoxicity studies of allysine-baceridin, a cyclic bioactive peptide that we have successfully modified using our chemistry. Baceridin, a known tumor-suppressing cyclic peptide, inhibits the active site of the proteasome in tumor cells, thereby inducing apoptosis (Fig. 22). In the active site of this proteasome, a lysine residue exists that could potentially react with the aldehyde to form a covalent inhibitor.[29] By introducing an aldehyde onto this cyclic peptide, we hypothesize that allysine-baceridin may exhibit enhanced tumor-suppressing properties. We aim to test this hypothesis in future cytotoxicity studies.
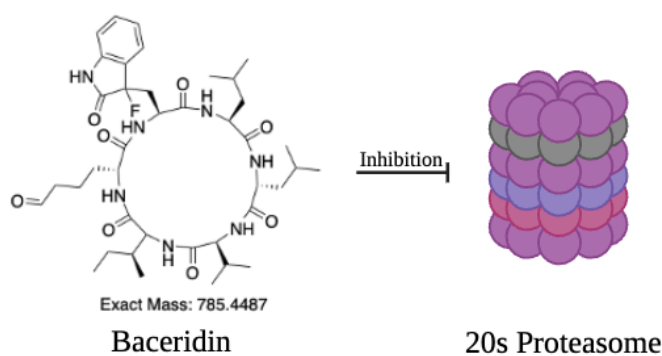


**Figure 22.** Proposed inhibition of 20s proteasome with cyclic allysine-baceridin.

I am a co-first author on this project with Benjamin Emenike. This research has been accepted by Angewandte Chemie International Edition for publication.

**Acknowledgement of contributions**

S.S., B.E., and M.R. designed the project. S.S. synthesized all linear and cyclic peptides in this thesis and analyzed results using HRMS and HPLC. S.S. performed synthetic steps on small molecule examples and analyzed by NMR. S.S. performed all peptide diversification reactions and analyzed using HRMS and HPLC. B.E. performed all computational calculations. B.E. conducted all cellular studies seen in this work. B.E. produced figures for Angewandte Chemie publication[5] and gave permission for use in this thesis. S.S. produced all other figures in this thesis. Mechanistic figures provided in this thesis were created using ChemDraw and BioRender. M.R., B.E., and S.S. wrote the manuscript for publication, which S.S. adapted and expanded upon for this thesis. S.S wrote the Supporting Information document for this thesis.

**References**

(1)    E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed. Engl*. **2009**, 48, 6974-6998.

(2)    O. El-Mahdi, O. Melnyk, *Bioconjugate Chem*. **2013**, 24, 735-765.

(3)    a) J. L. Lau, M. K. Dunn, *Bioorg. Med. Chem.* **2018**, 26, 2700-2707; b) A. Henninot, J. C. Collins, J. M. Nuss, *J. Med. Chem.* **2018**, 61, 1382-1414.

(4)    E. Sarubbi, P. F. Seneci, M. R. Angelastro, N. P. Peet, M. Denaro, K. Islam, *FEBS Lett.* **1993**, 319, 253-256.

(5)    B. Emenike, S. Shahin, M. Raj, *Angew. Chem. Int. Ed.* (In Press)

(6)    J. Forbes, K. Krishnamurthy, *StatPearls*, **2024**.

(7)    R. Spears, M. Fascione, *Organic & Biomolecular Chemistry* **2016**, 14 (32), 7622–7638.

(8)    a) J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi, M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, 45, 5307-5311; b) R. A. Scheck, M. T. Dedeo, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2008**, 130, 11762-11770;

(9)    a) A. Moulin, J. Martinez, J.-A. Fehrentz, *J. Pept. Sci.* **2007**, 13, 1-15; b) J. A. Fehrentz, M. Paris, A. Heitz, J. Velek, F. Winternitz, J. Martinez, *J. Org. Chem.* **1997**, 62, 6792-6796. c) J. Chen, X. Gong, J. Li, Y. Li, J. Ma, C. Hou, G. Zhao, W. Yuan, B. Zhao, *Science* **2018**, 360, 1438-1442; d) M. Zhang, X. Zhang, J. Li, Q. Guo, Q. Xiao, Chin. *J. Chem.* **2011**, 29, 1715-1720; e) Y. E. Liu, Z. Lu, B. Li, J. Tian, F. Liu, J. Zhao, C. Hou, Y. Li, L. Niu, B. Zhao, *J. Am. Chem. Soc.* **2016**, 138, 10730-10733; f) L. S. Witus, C. Netirojjanakul, K. S. Palla, E. M. Muehl, C.-H. Weng, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2013**, 135, 17223-17229.

(10)   a) A. Ginguay, L, Cynober, E. Curis, I Nicolis, *Biology (Basel)* **2017**, 6 (1), 18; b) Y. J. Suzuki, *Life* **2022**, 12, 967; c) G. Romagnoli, M. D. Verhoeven, R. Mans, Y. Fleury Rey, R. Bel-Rhlid, M. van den Broek, R. Maleki Seifar, A. Ten Pierick, M. Thompson, V. Müller, S. A. Wahl, J. T. Pronk, J. M. Daran, *Mol. Microbiol.* **2014**, 93, 369-389.

(11)   M. J. Appel, C. R. Bertozzi, *ACS Chem. Biol.* **2015**, 10, 72-84.

(12)   a) H.-J. Moon, J. Finney, T. Ronnebaum, M. Mure, *Bioorg. Chem.* **2014**, 57, 231-241; b) B. Piersma, R. A. Bank, *Essays Biochem.* **2019**, 63, 377-387; c) Q. Xiao, G. Ge, *Cancer Microenviron.* **2012**, 5, 261-273; d) C. Lunda, A. Arjona, C. Duenas, M. Estevez, *Antioxidants (Basel)* **2021**, 10, 474. e) L. Yang, S. Hsia, T. Shieh, *Int. J. Mol. Sci.* **2017**, 18, 62.

(13)   a) A. Dirksen, T. M. Hackeng, P. E. Dawson, Angew. Chem. Int. Ed. **2006**, 45, 7581-7584; b) S. Wang, G. N. Nawale, S. Kadekar, O. P. Oommen, N. K. Jena, S. Chakraborty, J. Hilborn, O. P. Varghese, *Sci. Rep.* **2018**, 8, 2193.

(14)   a) A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, J. Am. Chem. Soc. **2006**, 128, 15602-15603; b) A. Dirksen, P. E. Dawson, *Bioconjugate Chem.* **2008**, 19, 2543-2548.

(15)   a) J. Shao, J. P. Tam, J. Am. Chem. Soc. **1995**, 117, 3893-3899; b) K. Nakatsu, A. Okamoto, G. Hayashi, H. Murakami, *Angew. Chem. Int. Ed. Engl.* **2022**, 61, e202206240.

(16)   D. Chen, M. M. Disotuar, X. Xiong, Y. Wang, D. H.-C. Chou, *Chemical Science* **2017**, 8, 2717-2722.

(17)   W. Chan, P. White, *Oxford University Press*, **1999**.

(18)   T. V. Mourik, M. Buhl, M. Gaigeot, *Philos. Trans. A. Math Phys. Eng. Sci.* **2014**, 372.

(19)   P. T. Nyffeler, S. G. Duron, M. D. Burkart, S. P. Vincent, C. Wong, *Angew. Chem. Int. Ed. Engl.* **2004**, 44, 192-212.

(20)   a) S. S. Rawalay, H. Shechter, *J. Org. Chem.* **1967**, 32, 3129-3131; b) T. Seki, T. Fujiwara, Y. Takeuchi, J. Fluor. Chem. **2011**, 132, 181-185.

(21)   H. Zou, L. Li, T. Zhang, M. Shi, N. Zhang, J. Huang, M. Zian, *Biotechnology Advances* **2013**, 36, 1917-1927.

(22)   S. Sanjukta, A. K. Rai, *Trends Food Sci.* **2016**, 50, 1-10.

(23)   X. Fan, L. Bai, L. Zhu, L. Yang, X. Zhang, *J. Agric. Food Chem.* **2014**, 62, 9211-9222.

(24)   A. L. Capriotti, G. Caruso, C. Cavaliere, R. Samperi, S. Ventura, R. Z. Chiozzi, A. Laganà, *J. Food Compos. Anal.* **2015**, 44, 205-213.

(25)   X. Zhang, W. Shi, H. He, R. Cao, T. Hou, J. Funct. Foods **2020**, 73, 104100.

(26)   S. Joo, *Biomol. Ther. (Seoul)* **2012**, 20, 19-26.

(27)   R. Sekizawa, I. Momose, N. Kinoshita, H. Naganawa, M. Hamada, Y. Muraoka, H. Iinuma, T. Takeuchi, *J. Antibiot. (Tokyo)* **2001**, 54, 874-881.

(28)   J. Demiselle, N. Fage, P. Radermacher, P. Asfar, *Ann. Intensive Care* **2020**, 10, 9.

(29)   J. Niggemann, P. Bozko, N. Bruns, A. Wodtke, M. T. Gieseler, K. Thomas, C. Jahns, M. Nimtz, I. Reupke, T. Brüser, G. Auling, N. Malek, M. Kalesse, *Chembiochem* **2014**, 15, 1021-1029.

(30)   Q. Guan, S. Huang, Y. Jin, R. Campagne, V. Alezra, Y. Wan, *J. Med. Chem.* **2019**, 62, 7603-7617.

(31)   J. Lee, S. B. Park, *Pharmaceuticals (Basel)* **2022**, 15, 1478.

(32)   Faridoon, R. Ng, G. Zhang, J. J. Li, *Med. Chem. Res.* **2023**, 32, 1039-1062.

(33)   S. T. Rehan, H. U. Hussain, F. Malik, R. M. Usama, M. J. Tahir, M. S. Asghar, *Health Sci. Rep.* **2022**, 5, 713.

(34)   R. Fairhurst, R. T. Knoepfel, N. Buschmann, C. Leblanc, R. Mah, M. Todorov, P. Nimsgern, S. Ripoche, M. Niklaus, N. Warin, V. Luu, M. Madoerin, J.Wirth, D. Graus-Porta, A. Weiss,  M. Kiffe, M. Wartmann, J. Kinyamu-Akunda, D. Sterker, C. Stamm, F. Adler, A. Buhles, H. Schadt, P. Couttet, J. Blank, I. Galuba, J. Trappe, J. Voshol, N. Ostermann, C. Zou, J. Berghausen, A. Del Rio Espinola, W. Jahnke, P. Furet, *J. Med. Chem.* **2020**, 63 (21), 12542–12573.

(35)   J. Qiao, Y. Li, R. Zeng, F. Liu, R. Luo, C. Huang, Y. Wang, J. Zhang, B. Quan, C. Shen,  X. Mao, X. Liu, W. Sun, W. Yang, X. Ni, K. Wang, L. Xu, Z. Duan, Q. Zou, H. Zhang, W. Qu,  Y. Long, M. Li, R. Yang, X. Liu, J. You, Y. Zhou, R. Yao, W. Li, J. Liu, P. Chen, Y. Liu, G. Lin, X. Yang, J. Zou, L. Li, Y. Hu, G. Lu, W. Li, Y. Wei, Y. Zheng, J. Lei, S. Yang, *Science* **2021**, 371 (6536), 1374–1378.

(36)   S. S. Arya, J. E. Rookes, D. M. Cahill, S. K. Lenka, *Adv. Tradit. Med. (ADTM)* **2021**, 21, 1-17.

(37)   K. Bezchlibnyk-Butler, I. Aleksic, S. H. Kennedy, *J. Psychiatry Neurosci.* **2000**, 25, 241-254.

(38)   R. Behrendt, P. Whtie, J. Offer, *J. Pept. Sci.* **2016**, 22(1), 4-27.

(39)   Rink Amide PEGA resin Novabiochem® | Sigma-Aldrich.

(40)   B. Emenike, J. Donovan, M. Raj, *J. Am. Chem. Soc.* **2023**, 145, 16417-16428.

(41)   T. Liburkin-Dan, S. Toledano, G. Neufeld, *Int. J. Mol. Sci.* **2022**, 23.

(42)   S. D. Vallet, C. Berthollier, R. Salza, L. Muller, S. Ricard-Blum, *Cancers* **2021**, 13, 71.

(43)   Y. Shan, N. Sayed, K. C. Chong, H. J. Ting, X. Liu, B. Li, J. Liu, M. Wu, J.-W. Wang, B. Liu, *ACS mater. lett.* **2023**, 3171-3176.

# Supporting Information

*This SI has been re-printed and adapted from Angew Chem Publication.*[5]

Sophia Shahin[#], Benjamin Emenike[#], Monika Raj[*]

**I. General.** All commercial materials (Sigma-Aldrich, Fluka and Novabiochem) were used without further purification. All solvents were reagent or HPLC (Fisher) grade. All reactions were performed under air in glass vials. Yields refer to chromatographically pure compounds; % conversions were determined by integrating the area under the HPLC peaks of reaction mixture. HPLC and MS were used to monitor reaction progress, and products were characterized using MS and NMR.

**II. Materials.** Fmoc-amino acids, Rink amide resin, PEGA-amide resin, 3-[bis(dimethylamino)methyliumyl]-3H-benzotriazol-1-oxide hexafluorophosphate (HBTU), 1-hydroxy-7-azabenzotriazole (HOAt), N,N′-iisopropylcarbodiimide (DIC), and N,N-diisopropylethylamine (DIEA) were obtained from CreoSalus (Louisville, Kentucky). Piperidine and trifluoroacetic acid (TFA), were obtained from Alfa Aesar (Ward Hill, Massachusetts). N,N-dimethylformamide (DMF), dichloromethane ($CH_2Cl_2$), methanol (MeOH) and acetonitrile (ACN) were obtained from VWR (100 Matsonford Road Radnor, Pennsylvania). Selectfluor (1-Chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane bis(tetrafluoroborate) and pyridine were obtained from Sigma.

**III. Purification. HPLC:** Purification of peptides was performed using high performance liquid chromatography (HPLC) on an Agilent 1100 series HPLC equipped with a C-18 reverse phase column with a particle size of 5 μm. All separations involved a mobile phase of 0.1% formic acid in water (solvent A) and 0.1 % formic acid in acetonitrile (solvent B). The HPLC method used a linear gradient of 0-80% solvent B over 30 minutes at ambient temperature with a flow rate of 1 mL min$^{-1}$. The eluent was monitored by absorbance at 220 nm.

**IV. Instrumentation and sample analysis. NMR.** $^1H$ and $^{13}C$ spectra were acquired at 25 °C in DMSO-$d_6$, CDCl$_3$ using an Agilent DD2 (600 MHz) spectrometer with a 3-mm He triple resonance (HCN) cryoprobe. All $^1H$ NMR chemical shifts (δ) were referenced relative to the residual DMSO-$d_6$ peak at 2.50 ppm, CDCl$_3$ peak at 7.26 ppm or internal tetramethylsilane (TMS) at 0.00 ppm. $^{13}C$ NMR chemical shifts were referenced to DMSO-$d_6$ at 39.52 ppm and CDCl$_3$ at 77.2 ppm. $^{13}C$ NMR spectra were proton decoupled. NMR spectral data are reported as chemical shift (multiplicity, coupling constants (J), integration). Multiplicity is reported as follows: singlet (s), broad singlet (br s), doublet (d), doublet of doublets (dd), doublet of triplets (td), triplet (t) and multiplet (m). Coupling constant (J) in hertz (Hz).

**Analytical HPLC.** Analytical HPLC chromatography (HPLC) was performed on an Agilent 1200 series equipped with a C-18 reversed-phase column with a particle size of 5 μm. The reaction was monitored by analytical reverse phase HPLC using a gradient of water versus acetonitrile. All separations involved mobile phase with 0.1 % formic acid in water (solvent A) and 0.1 % formic acid in acetonitrile (solvent B). Analytical HPLC method used for analysis of reaction: Gradient: 0 to 80 % B in 30 min; 80-100 % B in 31-35 min at a flow rate of 1 mL/min.

**LC/MS.** High resolution LC-MS conditions for all purified peptides: Analyses were performed on an ultraperformance LC system (ACQUITY, Waters Corp., USA) coupled with a quadrupole time-of-flight mass spectrometer (Q-ToF Premier, Waters) with electrospray ionization (ESI) in positive mode using Mass lynx software (V4.1) or high-performance LC system (Agilent, 1100 series) coupled with triple quadrupole. LC-MS (Agilent technologies 6460) with electrospray ionization

(ESI) in positive mode using Agilent mass hunter (10.0). Unless otherwise mentioned a sample was injected either onto a C4 column (Phenomenex Aeris™ 3.6 µm WIDEPORE C4 200 Å, LC Column 50 x 2.1 mm) with a 400 µL/min flow rate of mobile phase of solution A (90 % $H_2O$, 10 % acetonitrile and 0.1 % formic acid (FA)) and solution B (95 % acetonitrile, 5 % $H_2O$, and 0.1 % formic acid) beginning gradient- Time- 0 min 10 % B; 5 min 28 % B; 20 min 38 % B; 22 min 90 % B; C18 column (ACQUITY UPLC BEH 1.7 µm 1x 50 mm) with a 200 µL/min flow rate of mobile phase of solution A (90 % $H_2O$, 10 % acetonitrile and 0.1 % formic acid) and solution B (90 % acetonitrile, 10 % $H_2O$, and 0.1 % formic acid) beginning gradient- Time- 1 min 0% B; 1-10 min 100% B for chromatography analysis (or) directly injected with mobile phase 90 % $H_2O$: 10 % ACN, 0.1% formic acid at 400 µL/min flow rate in ESI positive mode.

**HRMS.** High resolution MS data were acquired on Thermo Exactive Plus using a heated electrospray source. The solution was infused at a rate of 10-25 µL/min/electrospray using 3.3 KV. The typical settings were Capillary temp 320 °C. S-lens RF level was between 30-80 with an AGC setting of 1 E6. The maximum injection time was set to 50 ms. Spectra were taken at 140,000 resolutions at m/z 200 using Tune software and analyze with Thermo's Freestyle software.

**V. Fmoc Solid-Phase Peptide Synthesis (Fmoc-SPPS).[1]** Peptides were synthesized manually on a 0.25 mm scale using Rink amide resin. Resin was swollen with $CH_2Cl_2$ for 1 h at room temperature. Fmoc was deprotected using 20% piperidine in DMF for 5 min to obtain a deprotected peptide-resin. First, Fmoc protected amino acid (1.25 mm/5 equiv.) was coupled using HOAt (1.25 mm/5 equiv.) and DIC (1.25 mm/5 equiv.) in DMF for 15 min at room temperature. Fmoc-protected amino acids (0.75 mm/3 equiv.) were sequentially coupled on the resin using HBTU (0.75 mm/3 equiv.) and DIEA (1.5 mm/6 equiv.) in DMF for 5 min at room temperature. Peptides were cleaved from the resin using a cocktail of 95:2.5:2.5, trifluoroacetic acid: water:triethylsilane (TES) for 2 h. The resin was removed by filtration and the resulting solution was concentrated. The residue was diluted with ACN/water mixture. The resulting solution was purified by HPLC.

**VI. General procedure 1: Oxidative transformation of dimethyllysine peptides.**

To 1.0 mg of dimethyllysine peptide dissolved in 400 µL of water, was added pyridine (5 eq) and selectfluor (2 eq). The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by integrating the area under the HPLC peaks of reaction mixture.

**VII. General procedure 2: Oxidative transformation of tertiary amine small molecules.**

To 500 mg of tertiary amine compounds dissolved in 3 mL of water:ACN (1:1), was added pyridine (5 eq) and selectfluor (2 eq). The reaction mixture was stirred for 1 h. Upon completion as analyzed by TLC, the reaction mixture was concentrated and purified by silica gel column chromatography to generate the aldehyde product. ^{1}H and ^{13}C NMR spectra were acquired for characterization.

## VIII. General procedure 3: Oxidative transformation of dimethyllysine peptide on solid-support.

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing dimethyllysine peptide dissolved in 18 mL of water, was added pyridine (100 µL) and selectfluor (200 mg). The reaction mixture was stirred for 6 h. Upon completion, resin was washed with DMF, MeOH, and DCM before cleavage in (95% TFA, 2.5% Triethyl silane, 2.5% water). Following cleavage, peptide was dried under air and subsequently injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.

## IX. Supplementary Figure 1. Evaluation of allysine synthesis from lysine oxidation

**NADP+**
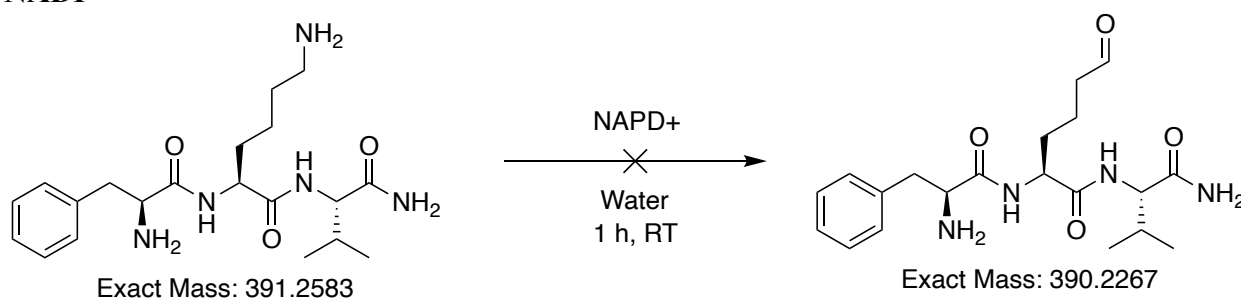


Exact Mass: 391.2583          Exact Mass: 390.2267

To 1.0 mg of FKV peptide dissolved in 400 µL of water, was added 5 eq of NADP$^+$. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.

**FKV peptide.** LCMS: m/z 392.2809(calcd [M+H]$^+$ = 392.2656), m/z 196.6440 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5545 (calcd [2M+H]$^+$ = 783.5554). (HPLC analysis at 220 nm). Retention time in HPLC: 4.690.

**FKV (reaction).** LCMS: m/z 392.2812(calcd [M+H]$^+$ = 392.2656), m/z 196.6441 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5554 (calcd [2M+H]$^+$ = 783.5554). (HPLC analysis at 220 nm). Retention time in HPLC: 4.607.

## HPLC Trace of FKV Peptide



## MS-Trace of peak 4.690



Exact Mass: 391.2583

## HPLC Trace of Reaction

**MS-Trace of peak 4.607**



Exact Mass: 391.2583

## Hydrogen Peroxide



Exact Mass: 391.2583

Exact Mass: 390.2267

To 1.0 mg of FKV peptide dissolved in 400 µL of water, was added 5 eq of hydrogen peroxide. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.
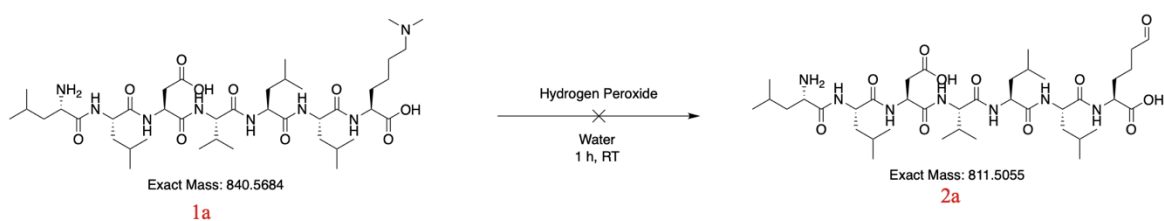
**FKV peptide.** LCMS: m/z 392.2809(calcd $[M+H]^+$ = 392.2656), m/z 196.6440 (calcd $[M+2/2]^+$ =196.6328), m/z 783.5545 (calcd $[2M+H]^+$ = 783.5554). (HPLC analysis at 220 nm). Retention time in HPLC: 4.690.

**FKV (reaction)**. LCMS: m/z 392.2811 (calcd $[M+H]^+$ = 392.2656), m/z 196.6441 (calcd $[M+2/2]^+$ =196.6328), m/z 783.5553 (calcd $[2M+H]^+$ = 783.5312). (HPLC analysis at 220 nm). Retention time in HPLC: 5.246.

**HPLC Trace of FKV Peptide**



**MS-Trace of peak 4.690**



Exact Mass: 391.2583
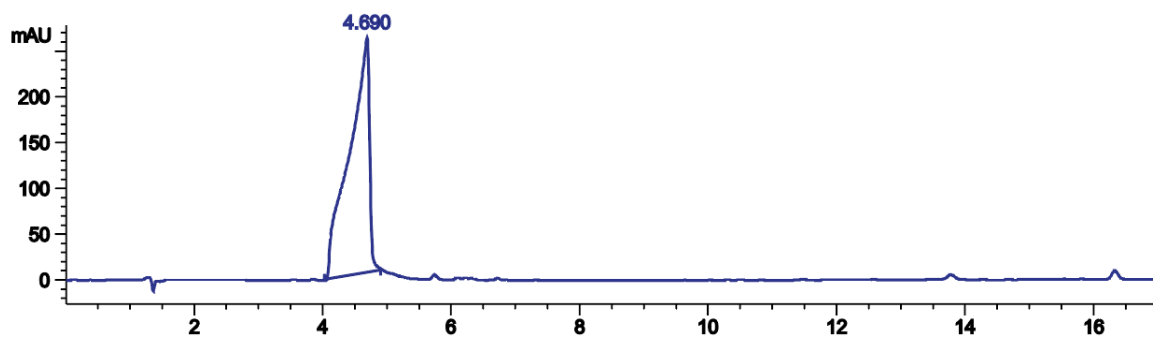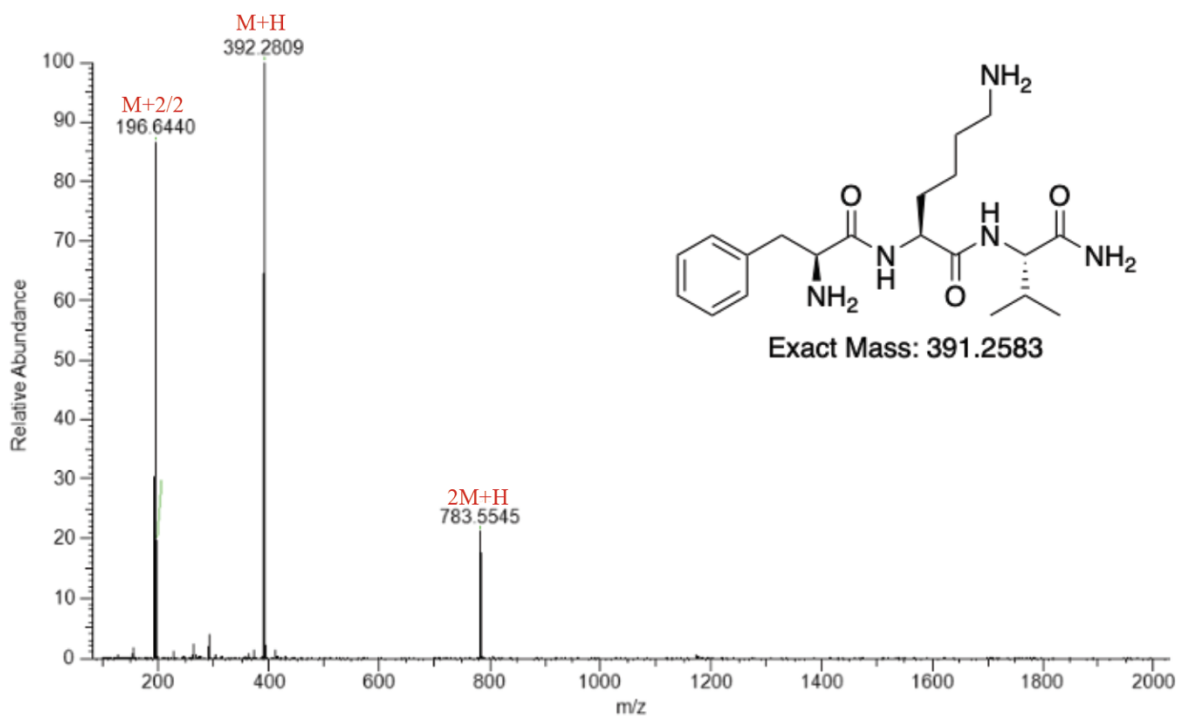
**HPLC-Trace of Reaction**
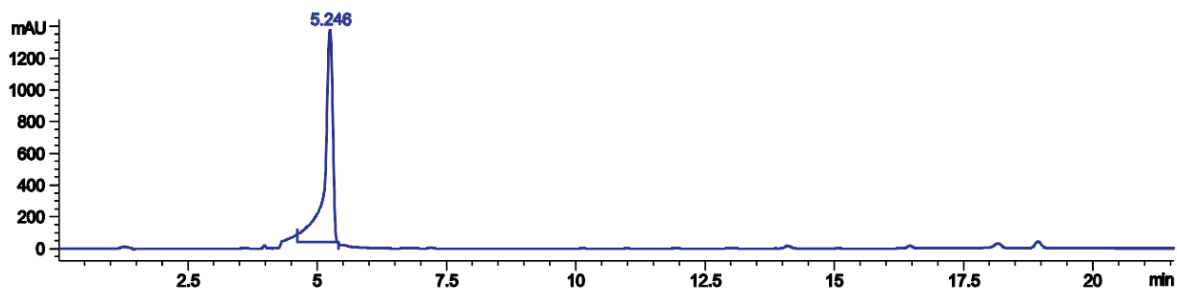


**MS-Trace of peak 5.246**
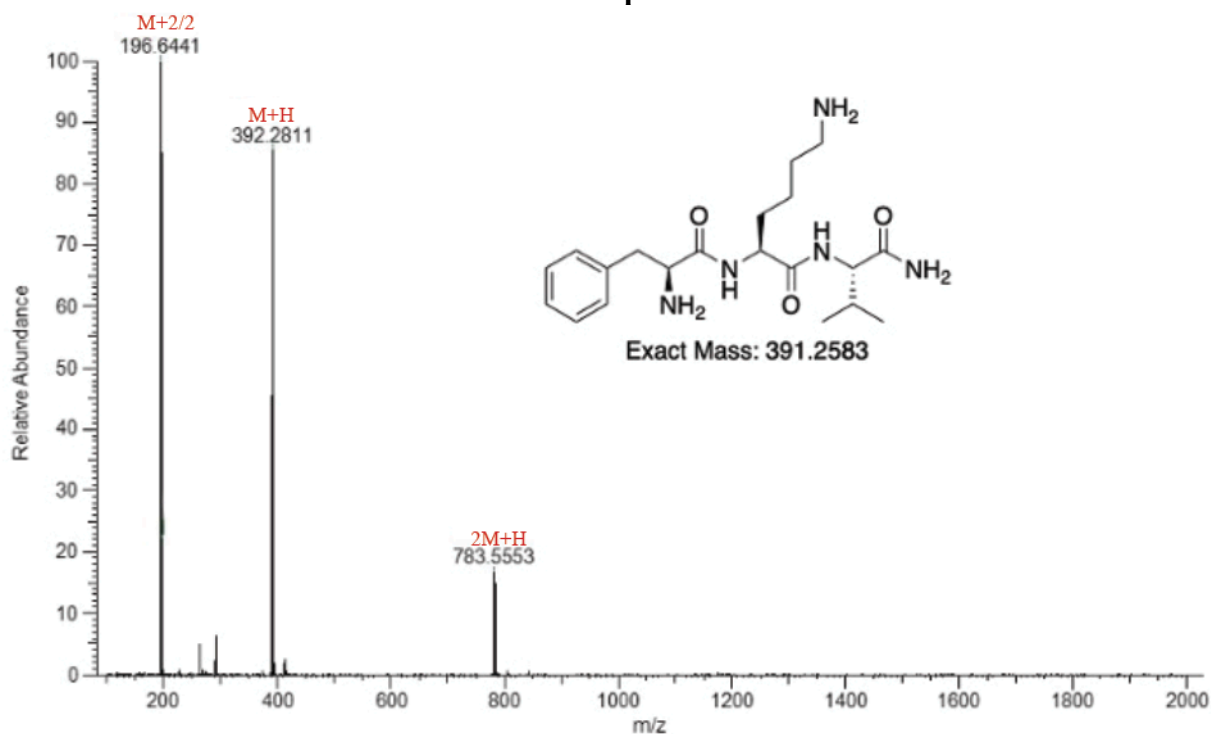


**Selectfluor**



To 1.0 mg of FKV peptide dissolved in 400 µL of water, was added 5 eq of selectfluor. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-

MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.

**FKV peptide.** LCMS: m/z 392.2809(calcd [M+H]$^+$ = 392.2656), m/z 196.6440 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5545 (calcd [2M+H]$^+$ = 783.5554). (HPLC analysis at 220 nm). Retention time in HPLC: 4.690.

**FKV (reaction).** LCMS: m/z 392.2809 (calcd [M+H]$^+$ = 392.2656), m/z 196.6441 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5546 (calcd [2M+H]$^+$ = 783.5312). (HPLC analysis at 220 nm). Retention time in HPLC: 5.187.

**HPLC Trace of FKV Peptide**



**MS-Trace of peak 4.690**

## HPLC-Trace of Reaction



## MS-Trace of peak 5.187



Exact Mass: 391.2583

Rose Bengal



Exact Mass: 391.2583

Rose Bengal

Water
1 h, RT

Exact Mass: 390.2267
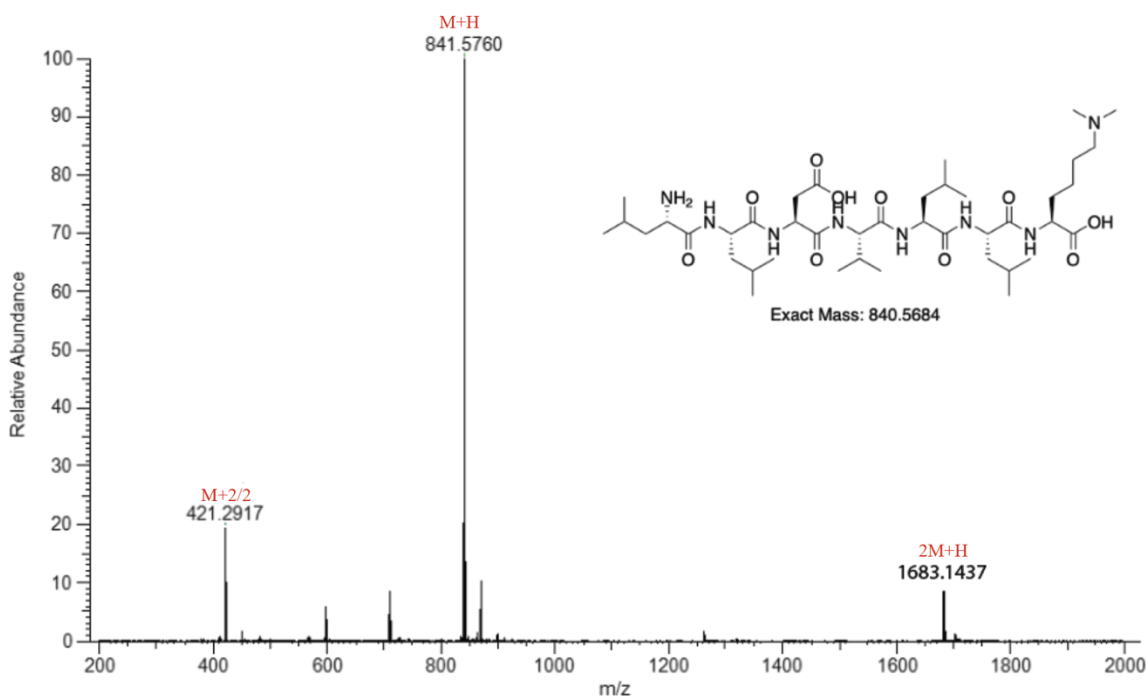
To 1.0 mg of FKV peptide dissolved in 400 μL of water, was added 2 mol% of rose bengal. The reaction mixture was photocatalyzed by a 30W lamp and stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.

**FKV peptide.** LCMS: m/z 392.2809(calcd [M+H]$^+$ = 392.2656), m/z 196.6440 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5545 (calcd [2M+H]$^+$ = 783.5554). (HPLC analysis at 220 nm). Retention time in HPLC: 4.690.

**FKV (reaction)**. LCMS: m/z 392.2809 (calcd [M+H]$^+$ = 392.2656), m/z 196.6441 (calcd [M+2/2]$^+$ =196.6328), m/z 783.5546 (calcd [2M+H]$^+$ = 783.5312). (HPLC analysis at 220 nm). Retention time in HPLC: 3.799.

**HPLC Trace of FKV Peptide**
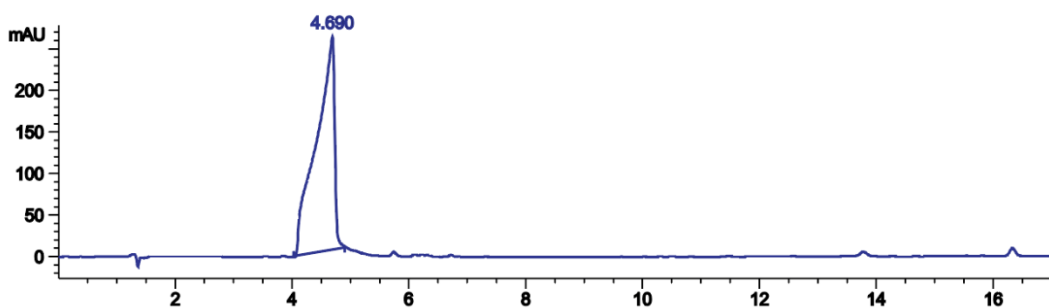


**MS-Trace of peak 4.690**
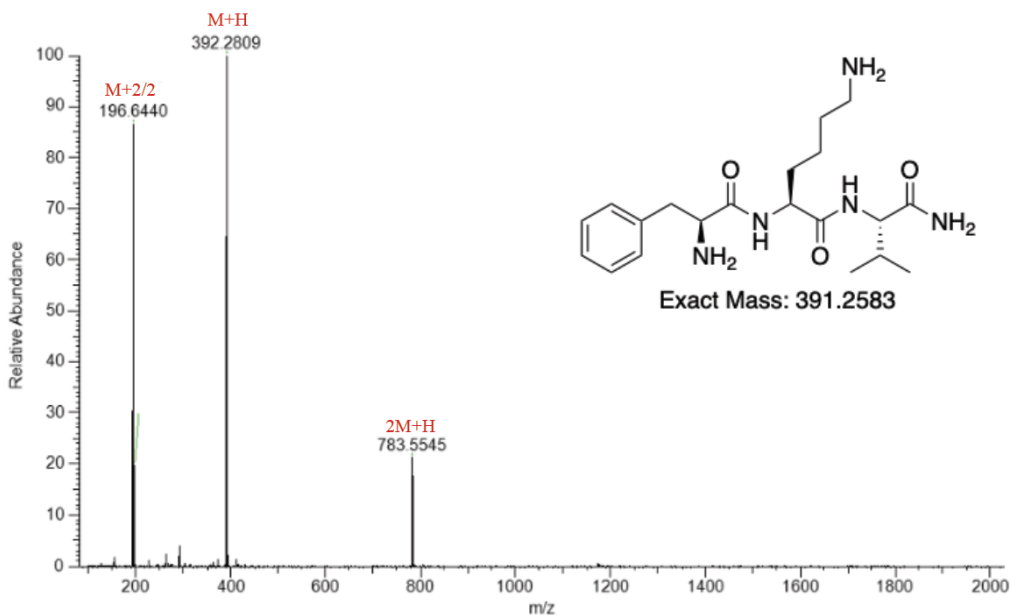
**HPLC-Trace of Reaction**
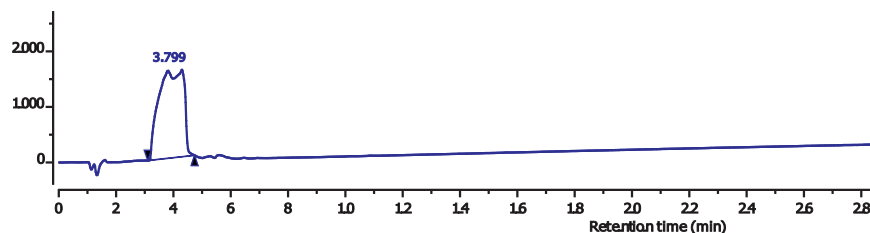


**MS-Trace of peak 3.799**
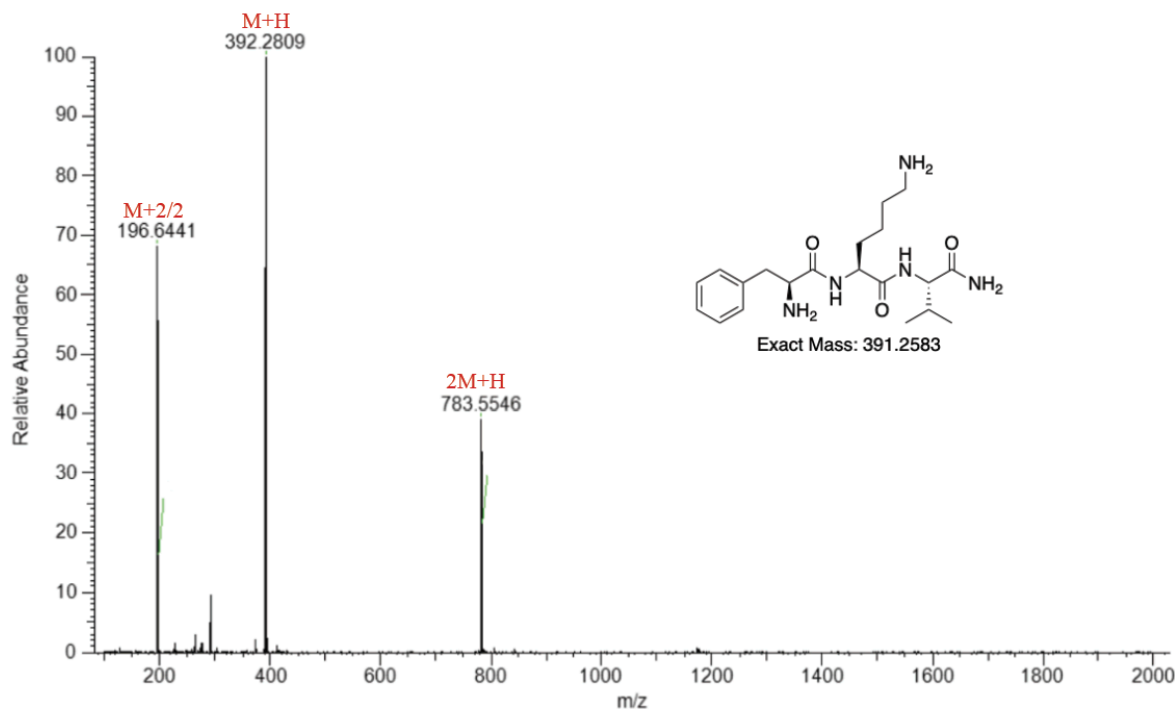


## X. Supplementary Figure 2. Computational evaluation of oxidative transformations of amines.

**Computational Details.**

All reported calculations were performed using the Gaussian-16 suite of programs[2] and Schrodinger computational suite. Geometry optimization and frequency calculations of all reported structures were done at the B3LYP-D3(BJ)/[6-31G(d,p)] level of theory. In other words, we used the B3LYP density functional[3-5] powered by Grimme's empirical dispersion-correction (D3)[6] and Becke-Johnson (BJ) damping-correction[7-9] in conjunction with split-valence 6-31G(d,p) basis sets for all atoms. Bulk solvent effects were incorporated for all geometry optimization and frequency calculations at the SMD level.[10] Here, we chose water as a solvent. The reported thermodynamic data were computed at a temperature of 298.15K and at 1atm of pressure.

**Evaluation of bond dissociation energy (BDE) of amines**

DFT-based Bond Dissociation Energy (BDE) calculation was performed using Schrodinger computational suite. Geometry optimization and frequency calculations were done with the B3LYP/6-31G* level of theory. Optimizations and BDE calculations led to a lower BDE for tertiary amine C-H bond over secondary and primary amine C-H bonds.

Primary amine (96.7 Kcal/mol):
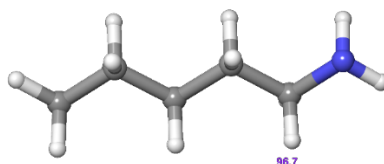


Secondary amine (96.4 Kcal/mol):
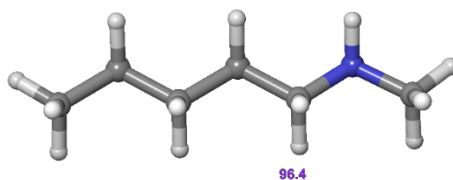


Tertiary amine (94.9 Kcal/mol):
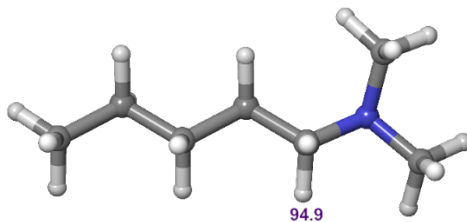


**Evaluation of thermodynamics, Electro-Static-Potential (ESP) and reactivity of amine generated iminium ions.**

Thermodynamic and ESP computations were performed using the Gaussian-16 suite of programs. Geometry optimization and frequency calculations calculations were done with the B3LYP/6-31G* level of theory. Thermodynamic calculations clearly highlight the favorable stability of iminium ion from tertiary amine (158 Kcal/mol) over secondary (160 Kcal/mol) and primary amines (165 Kcal/mol) iminium ions. ESP calculations further highlight the reactivity of tertiary amine generated iminium ion with a local C=N charge density of +0.209 (highly electrophilic). Significantly reduced charge density is observed for secondary (-0.355) and primary (-0.998) iminium ions.

Primary amine iminium: 165 Kcal/mol
ΔESP C=N: -0.998

Secondary amine iminium: 160 Kcal/mol
ΔESP C=N: -0.355

Tertiary amine iminium: 158 Kcal/mol
ΔESP C=N: +0.209
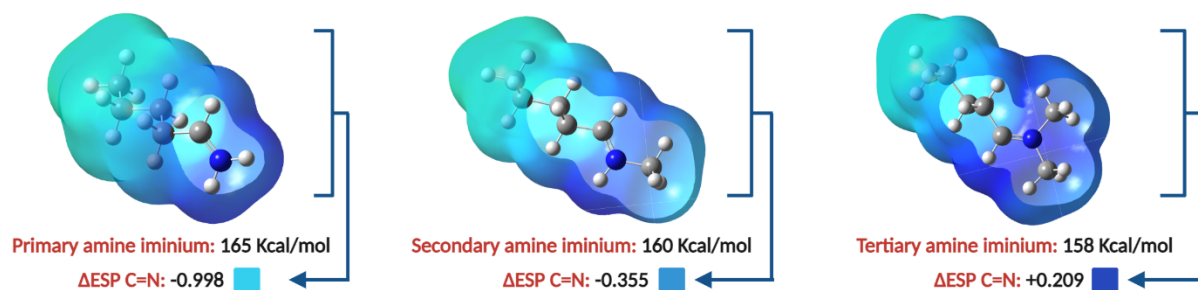
## XI. Supplementary Figure 3. Optimization reactions for aldehyde generation from dimethyl lysine.

**NADP+**



Exact Mass: 840.5684
1a

NAPD+
Water
1 h, RT

Exact Mass: 811.5055
2a

To 1.0 mg of dimethyllysine peptide LLDVLLKme$_2$ (1a) dissolved in 400 µL of water, was added 5 eq of NADP$^+$. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. No aldehyde product was observed.

**LLDVLLKme$_2$ peptide 1a**. LCMS: m/z 841.5739 (calcd [M+H]$^+$ = 841.5757), m/z 421.2908 (calcd [M+2/2]$^+$ = 421.2878), ), m/z 1683.1403 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 9.702.

**LLDVLLKme$_2$ (reaction).** LCMS: m/z 841.6080 (calcd [M+H]$^+$ = 841.5757), m/z 421.3079 (calcd [M+2/2]$^+$ = 421.2842), m/z 1683.1399 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 11.346.

# HPLC Trace of peptide 1a



# MS-Trace of peptide 1a



# HPLC-Trace of Reaction

**MS-Trace of peak 11.346**
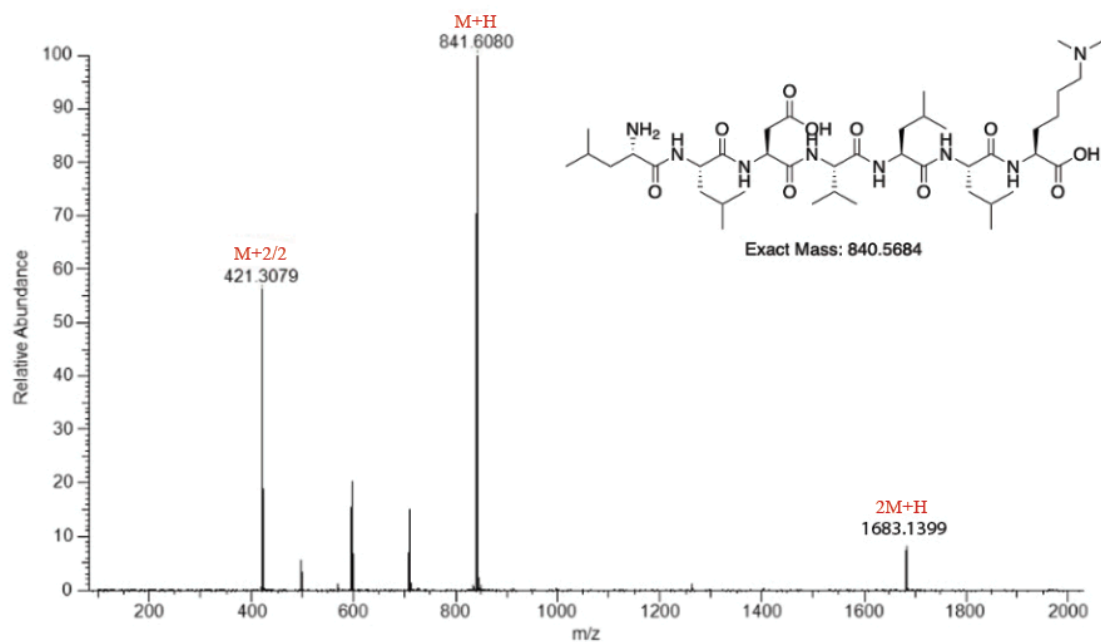


**Hydrogen Peroxide**



To 1.0 mg of dimethyllysine peptide LLDVLLKme$_2$ (1a) dissolved in 400 µL of water, was added 5 eq of hydrogen peroxide. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. No aldehyde product was observed. No aldehyde peptide was observed.

**LLDVLLKme$_2$ peptide 1a**. LCMS: m/z 841.5739 (calcd [M+H]$^+$ = 841.5757), m/z 421.2908 (calcd [M+2/2]$^+$ = 421.2878), ), m/z 1683.1403 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 9.702.

**LLDVLLKme$_2$ peptide 1a (reaction).** LCMS: m/z 841.5735 (calcd [M+H]$^+$ = 841.5757), m/z 421.2905 (calcd [M+2/2]$^+$ = 421.2842), m/z 1683.1357 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 10.983.

## HPLC Trace of peptide 1a



## MS-Trace of peptide 1a



## HPLC Trace of Reaction

## MS-Trace of peak 10.983
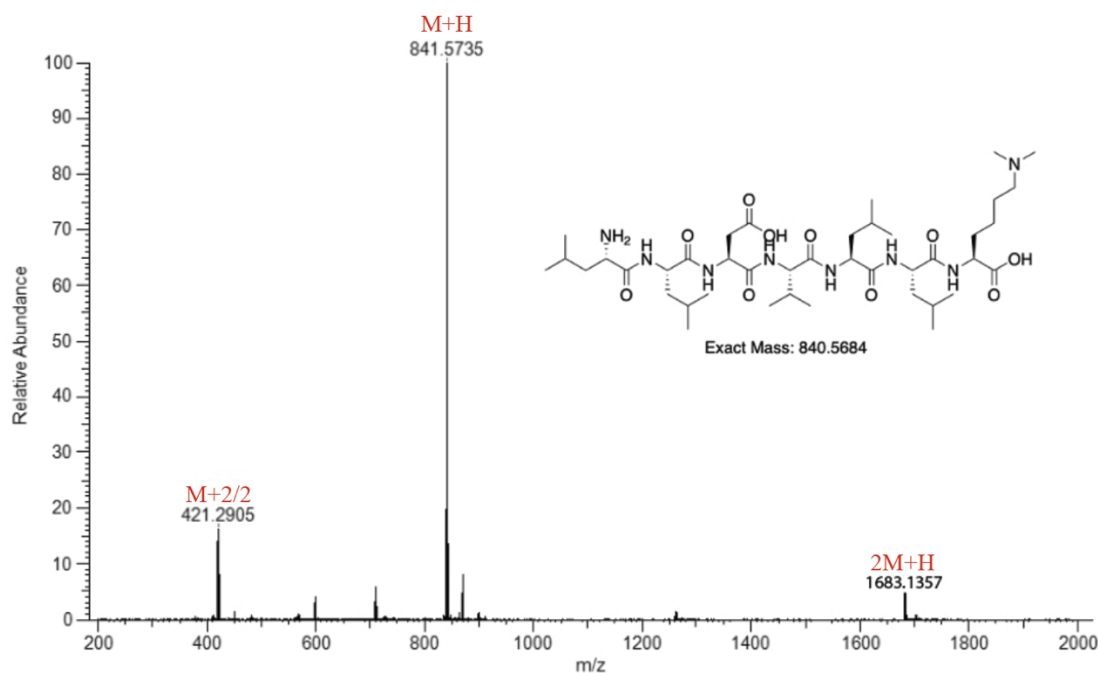


**Rose Bengal**



To 1.0 mg of dimethyllysine peptide LLDVLLKme$_2$ (1a) dissolved in 400 µL of water, was added 2 mol% of rose Bengal. The reaction mixture was photocatalyzed by a 30W lamp and stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. No aldehyde product was observed. No aldehyde peptide was observed.

**LLDVLLKme$_2$ peptide 1a**. LCMS: m/z 841.5739 (calcd [M+H]$^+$ = 841.5757), m/z 421.2908 (calcd [M+2/2]$^+$ = 421.2878), ), m/z 1683.1403 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 9.702.

**LLDVLLKme$_2$ peptide 1a (reaction).** LCMS: m/z 841.5760 (calcd [M+H]$^+$ = 841.5757), m/z 421.2917 (calcd [M+2/2]$^+$ = 421.2842), m/z 1683.1437 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 9.375.
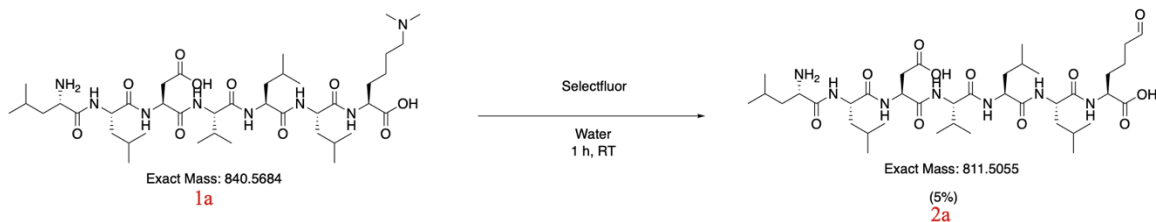
## HPLC Trace of peptide 1a



## MS-Trace of peptide 1a



Exact Mass: 840.5684

## HPLC-Trace of Reaction

## MS-Trace of peak 9.375



M+H 841.5760

M+2/2 421.2917

2M+H 1683.1437

Exact Mass: 840.5684

**Selectfluor**



Exact Mass: 840.5684
1a

Selectfluor
Water
1 h, RT

Exact Mass: 811.5055
(5%)
2a
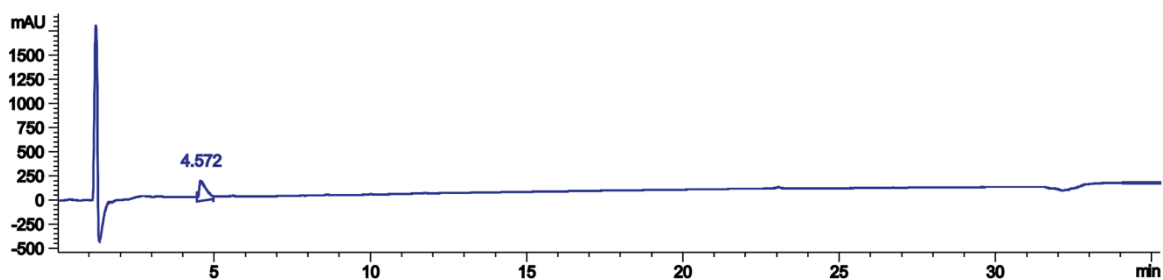
To 1.0 mg of dimethyllysine peptide LLDVLLKme$_2$ (1a) dissolved in 400 µL of water, was added 5 eq of selectfluor. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. 5% of aldehyde peptide was observed.

**LLDVLLKme$_2$ peptide 1a**. LCMS: m/z 841.5739 (calcd [M+H]$^+$ = 841.5757), m/z 421.2908 (calcd [M+2/2]$^+$ = 421.2878), ), m/z 1683.1403 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 9.702.

**LLDVLLKme$_2$ peptide 1a (reaction).** LCMS: m/z 841.5746 (calcd [M+H]$^+$ = 841.5757), m/z 421.2910 (calcd [M+2/2]$^+$ = 421.2842), m/z 1683.1387 (calcd [2M+H]$^+$ = 1683.2175). (HPLC analysis at 220 nm). Retention time in HPLC: 10.983.

**LLDVLLK(CHO) peptide 2a.** LCMS: m/z 812.4933 (calcd [M+H]$^+$ = 812.5128), m/z 889.5208 (calcd [M+2K]$^+$ = 889.4318). (HPLC analysis at 220 nm). Retention time in HPLC: 18.801.
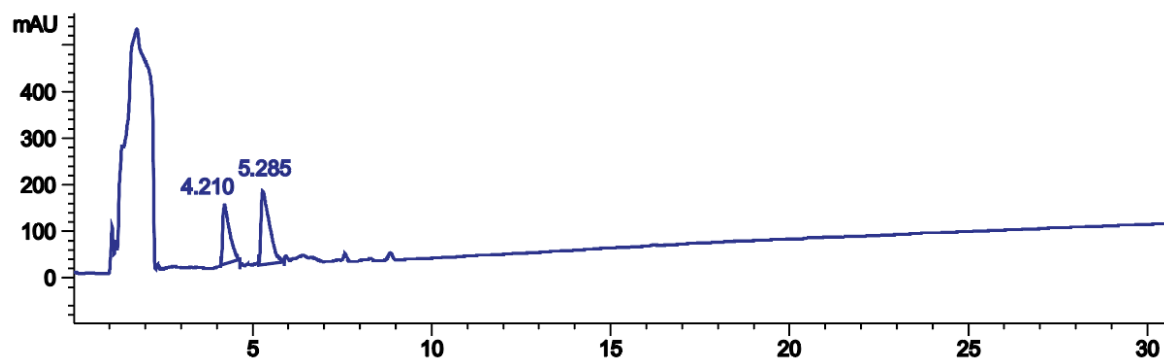
## HPLC Trace of peptide 1a



## MS-Trace of peptide 1a



## HPLC Trace of Reaction

**MS-Trace of peak 10.893**



M+H
841.5746

M+2/2
421.2910

2M+H
1683.1387

Exact Mass: 840.5684

**MS-Trace of peak 18.801**



M+2K
889.5208

M+H
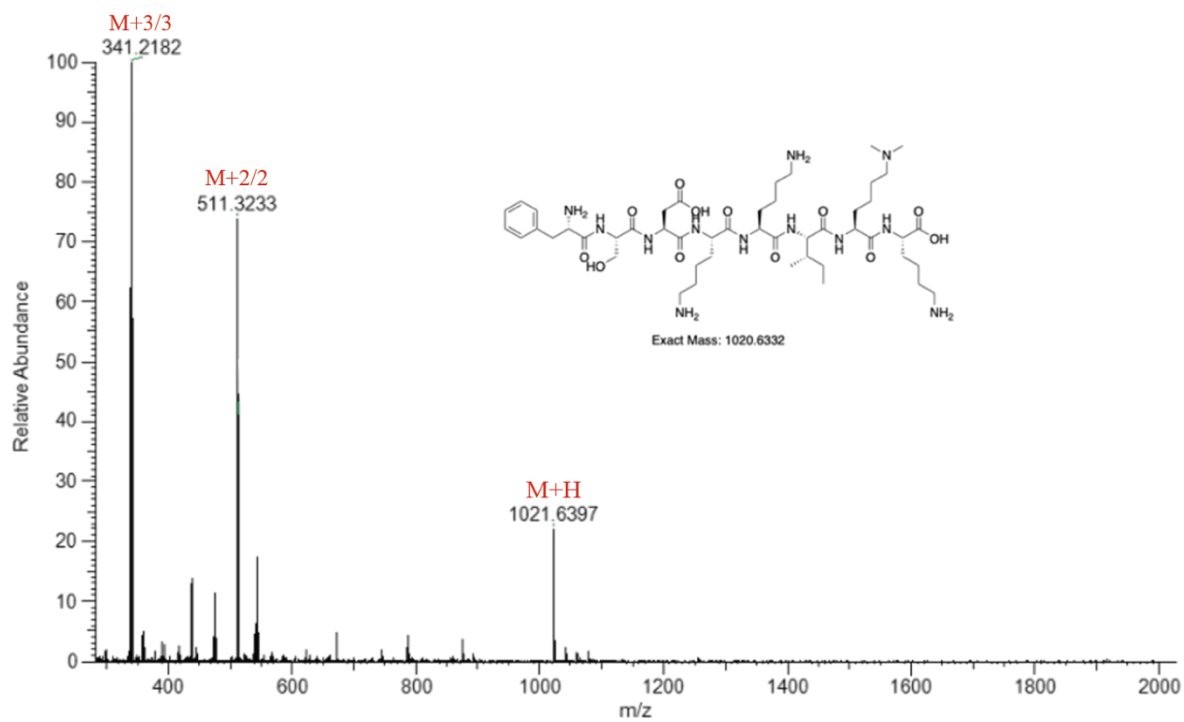812.4933

Exact Mass: 811.5055
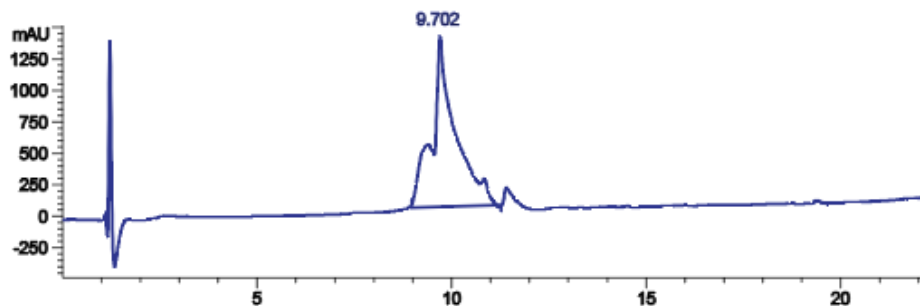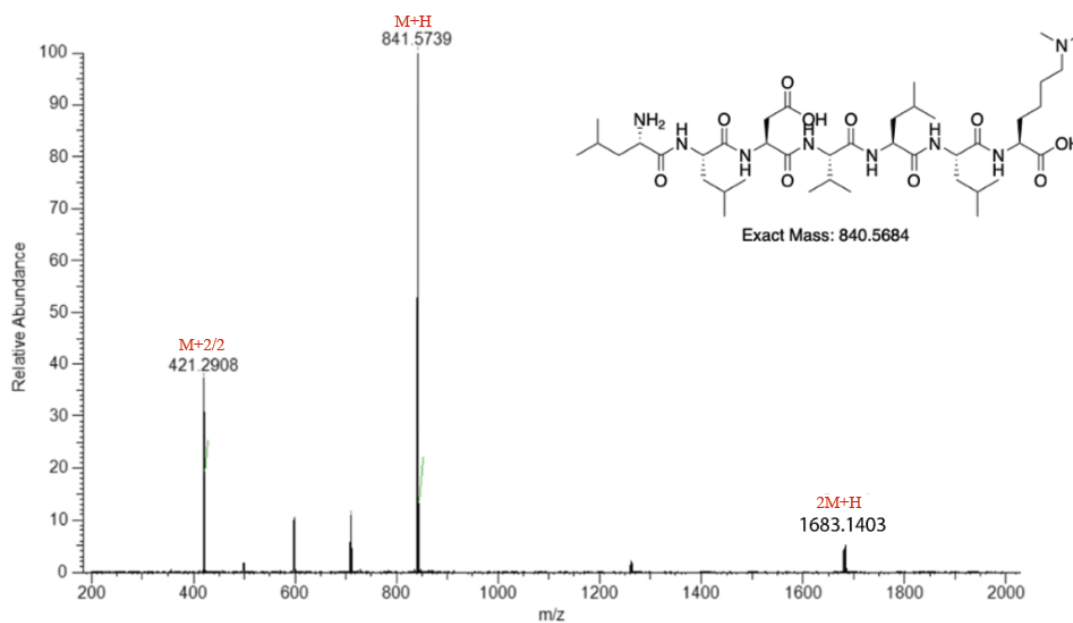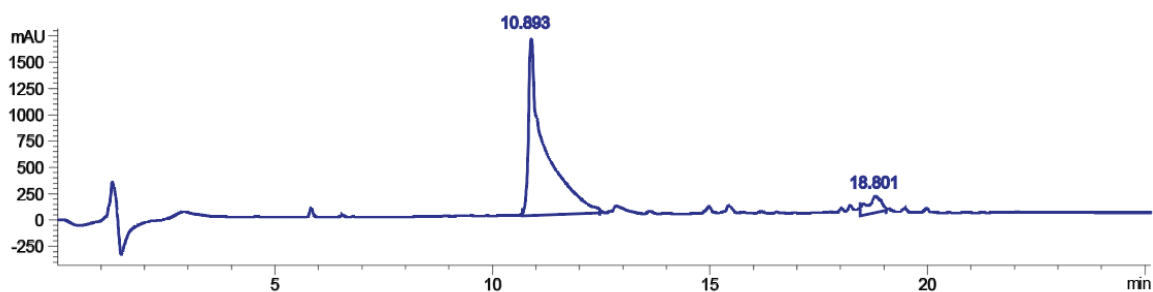
## Selectfluor with Pyridine
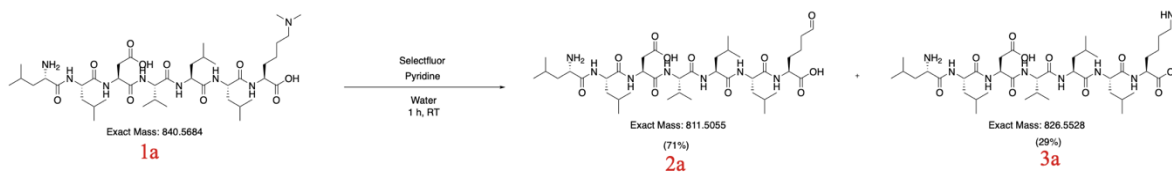


To 1.0 mg of dimethyllysine peptide LLDVLLKme₂ (1a) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 85%.

**LLDVLLKme₂ peptide 1a**. LCMS: m/z 841.5729 (calcd $[M+H]^+$ = 841.5757), m/z 421.2903 (calcd $[M+2/2]^+$ = 421.2878), m/z 1683.1426 (calcd $[2M+H]^+$ = 1683.2175) . (HPLC analysis at 220 nm). Retention time in HPLC: 9.702.

**LLDVLLK(CHO) peptide 2a.** LCMS: m/z 812.1937 (calcd $[M+H]^+$ = 812.5128), m/z 889.1364 (calcd $[M+2K]^+$ = 889.4318). (HPLC analysis at 220 nm). Retention time in HPLC: 18.409

**LLDVLLKme₁ peptide 3a.** LCMS: m/z 827.5642 (calcd $[M+H]^+$ = 827.5601), m/z 1676.0956 (calcd $[M+H]^+$ = 1676.0856). (HPLC analysis at 220 nm). Retention time in HPLC: 10.990.

### HPLC Trace of peptide 1a

## MS-Trace of peptide 1a



M+H
841.5729

M+2/2
421.2903

2M+H
1683.1426

Exact Mass: 840.5684

## HPLC Trace of Reaction



18.409

10.990

Retention time (min)

# MS-Trace of peak 10.597



M+H
827.5642

2M+Na
1676.0956

Exact Mass: 826.5528

# MS-Trace of peak 18.409



M+2K
889.1364

M+H
812.1937

Exact Mass: 811.5055
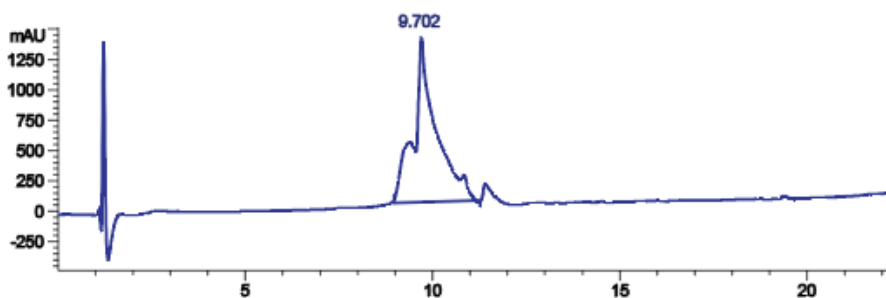
## XII. Supplementary Figure 4. Modification of dimethyllysine containing linear peptides to allysine peptide products.

**LILKme₂PF:**
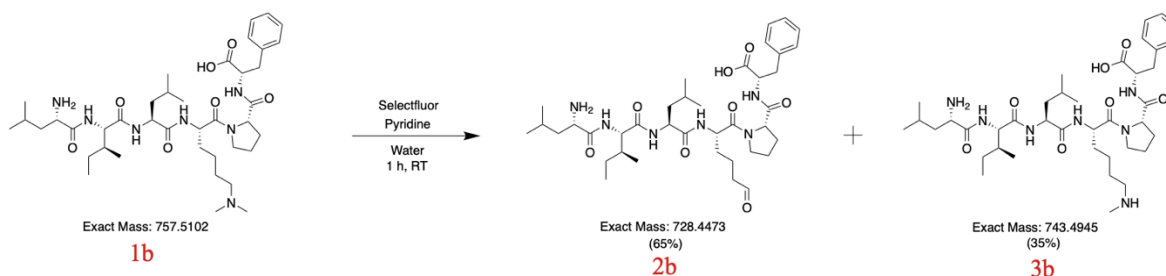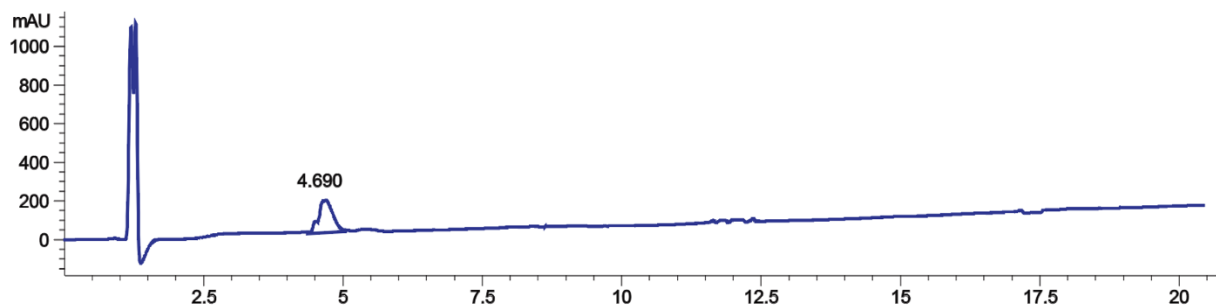


To 1.0 mg of dimethyllysine peptide LILKme₂PF (1b) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 64.4%.

**LILKme₂PF peptide 1b**. LCMS: m/z 758.5179 (calcd $[M+H]^+$ = 758.5175), m/z 379.7628 (calcd $[M+2/2]^+$ = 379.4545). (HPLC analysis at 220 nm). Retention time in HPLC: 9.607.

**LILK(CHO)PF peptide 2b**. LCMS: m/z 729.4550 (calcd $[M+H]^+$ = 729.4545). (HPLC analysis at 220 nm). Retention time in HPLC: 12.368.

**LILKme₁PF peptide 3b**. LCMS: m/z 744.5017 (calcd $[M+H]^+$ = 744.5018), m/z 372.7546 (calcd $[M+2/2]^+$ =372.7509). (HPLC analysis at 220 nm). Retention time in HPLC: 10.432.

### HPLC Trace of peptide 1b

## MS-Trace of peptide 1b



## HPLC Trace of Reaction
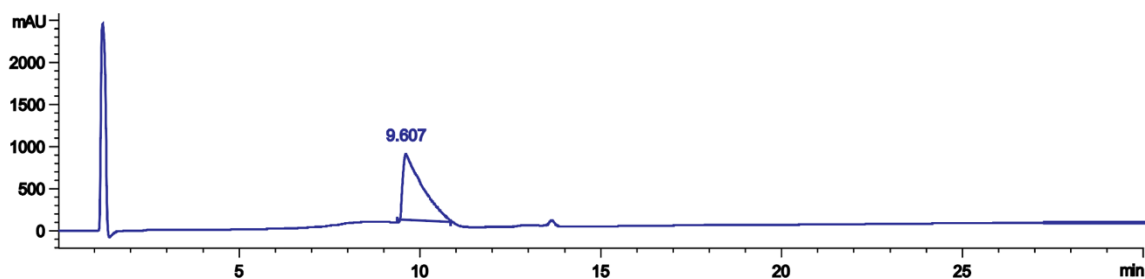
**MS-Trace of peak 10.432**



Exact Mass: 743.4945
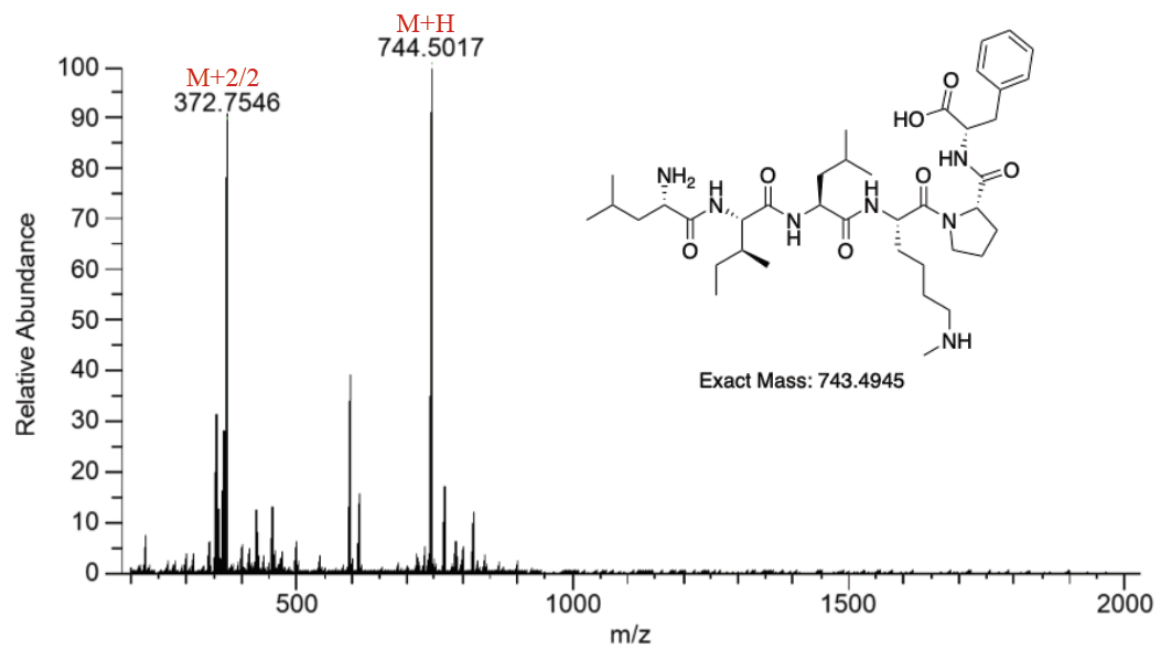
**MS-Trace of peak 12.368**



Exact Mass: 728.4473
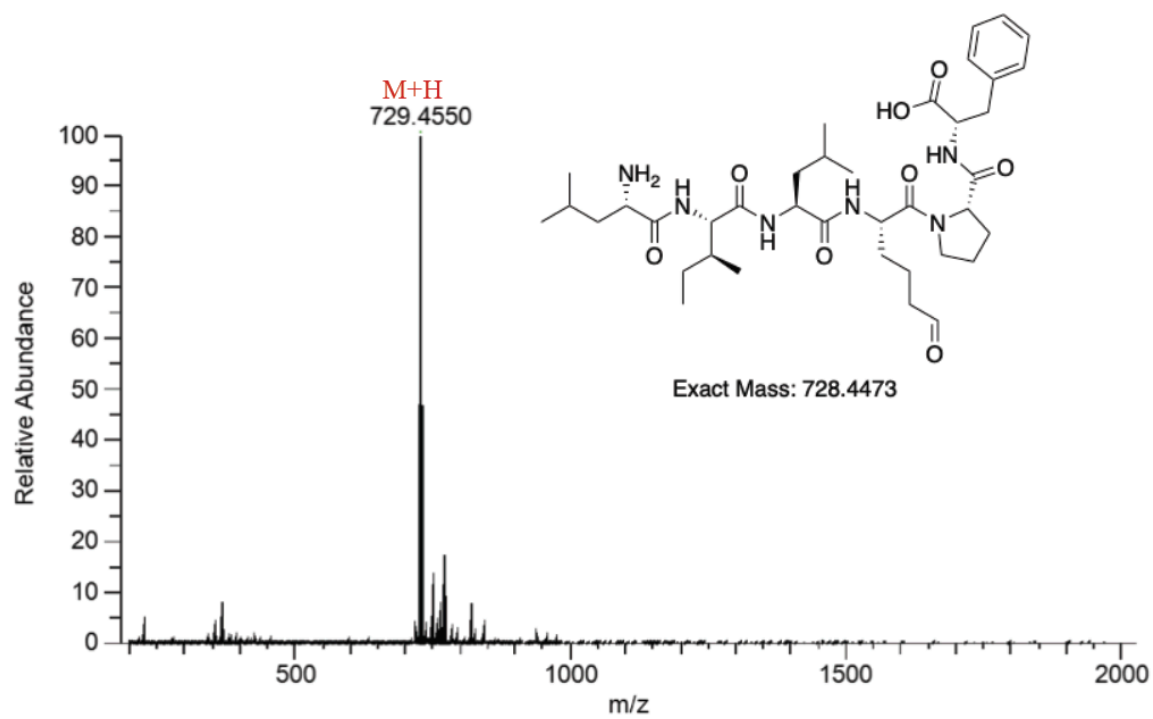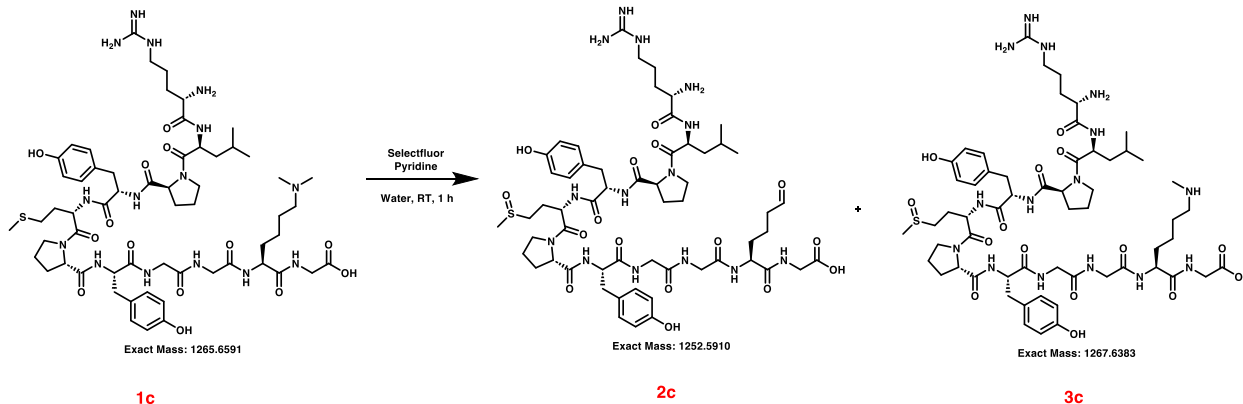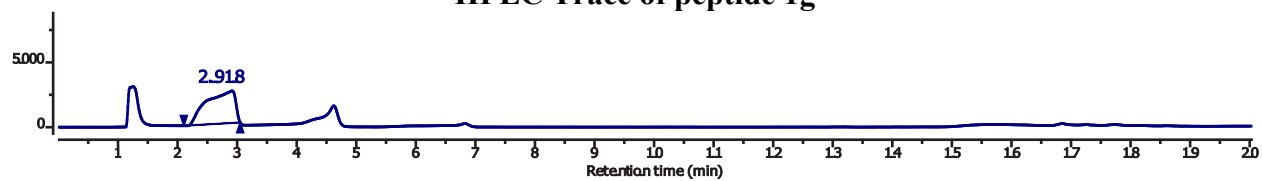
**RLPYMPYGGKme₂G:**



To 1.0 mg of dimethyllysine peptide RLPYMPYGGKme₂G (1c) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 87%. We observed the formation of sulfoxide from methionine under the reaction condition.

**RLPYMPYGGKme₂G peptide 1c**. LCMS: m/z 1266.6647 (calcd $[M+H]^+$ = 1266.6663), m/z 633.8289 (calcd $[M+2/2]^+$ = 633.8331), m/z 656.7443 (calcd $[(M+2/2)+Na]^+$ = 656.7331). (HPLC analysis at 220 nm). Retention time in HPLC: 5.972

**RLPYMPYGGK(CHO)G peptide 2c**. LCMS: m/z 1253.5971 (calcd $[M+H]^+$ = 1253.5983), m/z 627.2918 (calcd $[M+2/2]^+$ = 627.2991). (HPLC analysis at 220 nm). Retention time in HPLC: 7.162

**RLPYMPYGGKme₁G peptide 3c**. LCMS: m/z 1268.6483 (calcd $[M+H]^+$ = 1268.6456), m/z 634.8188 (calcd $[M+2/2]^+$ = 634.8228). (HPLC analysis at 220 nm). Retention time in HPLC: 5.569

**HPLC Trace of peptide 1c**
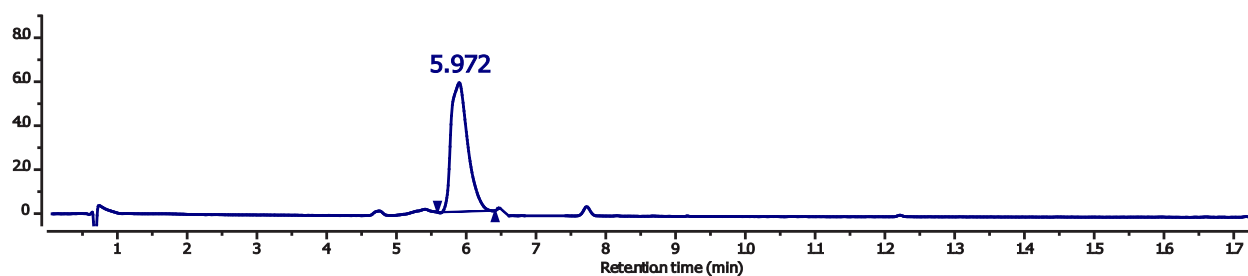
## MS Trace of peptide 1c
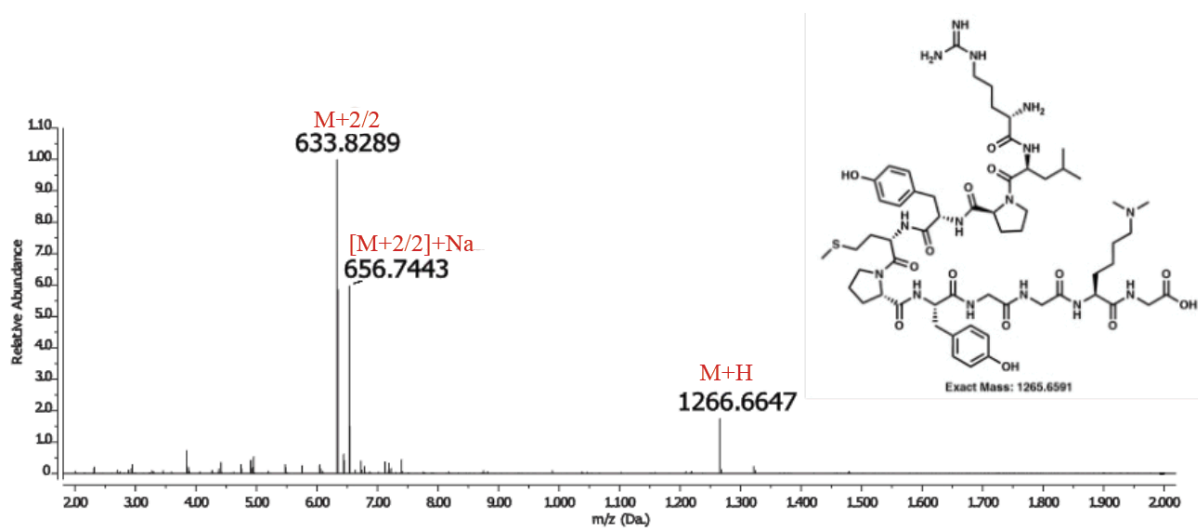


M+2/2
633.8289

[M+2/2]+Na
656.7443

M+H
1266.6647

Exact Mass: 1265.6591

## HPLC Trace of Reaction



7.162

5.569

Retention time (min)

## MS Trace of peak 5.569



M+2/2
634.8188

M+H
1268.6483

Exact Mass: 1267.6383

**MS Trace of peak 7.162**



Exact Mass: 1262.5910

**LDKme₂VNR:**



Exact Mass: 771.4603

**1d**

Selectfluor
Pyridine

Water
1 h, RT

Exact Mass: 742.3974

(58%)

**2d**

Exact Mass: 757.4446

(42%)

**3d**
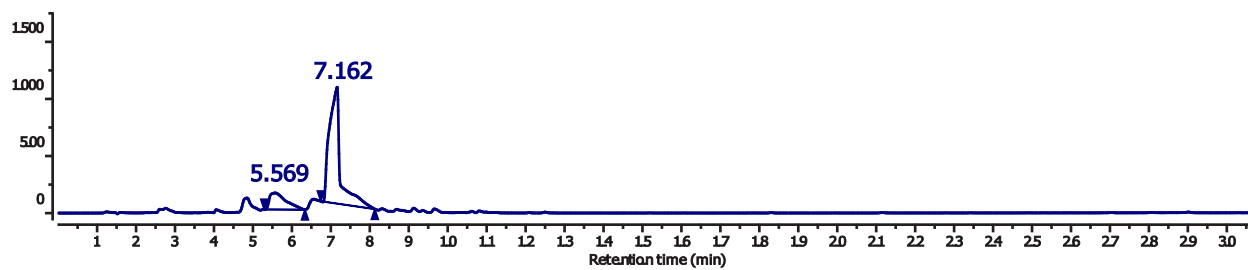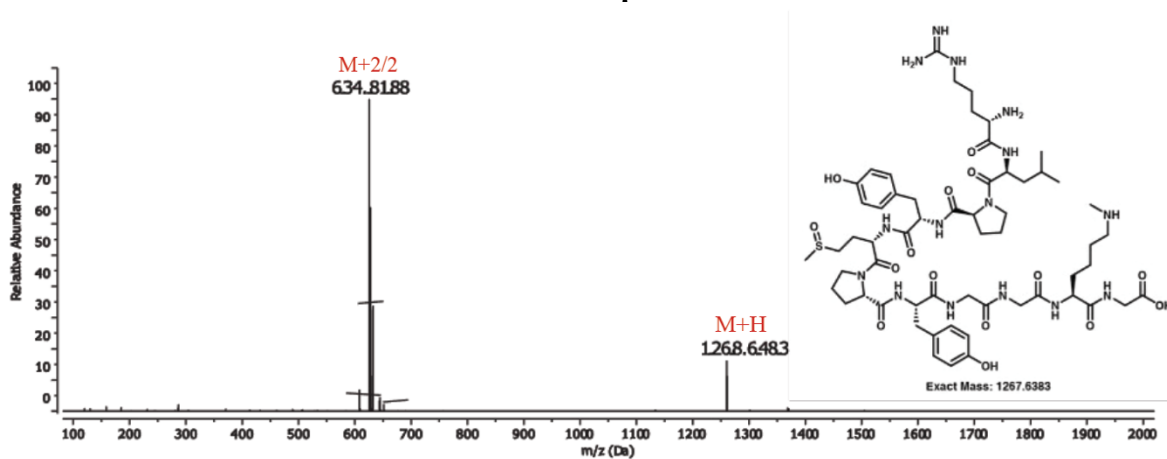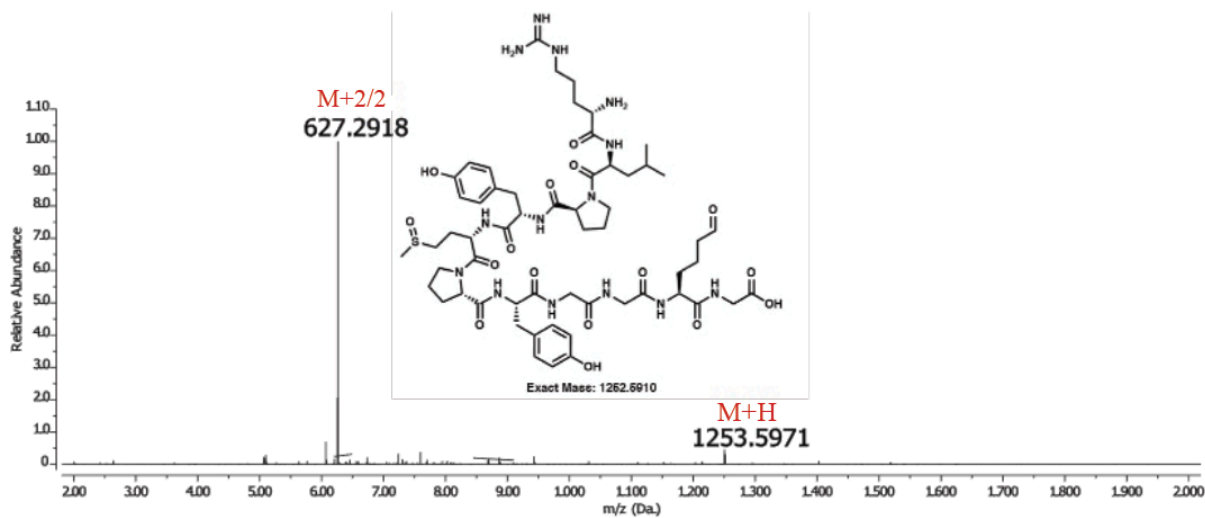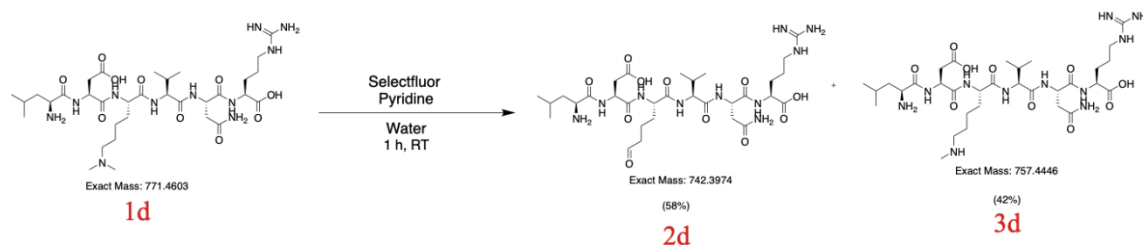
To 1.0 mg of dimethyllysine peptide LDKme₂VNR (1d) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 58%.

**LDKme₂VNR peptide 1d**. LCMS: m/z 772.4650 (calcd [M+H]+ = 772.4676), m/z 386.7363 (calcd [M+2/2]+ = 386.7338), ), m/z 1543.9222 (calcd [2M+H]$^+$ = 1543.9281). (HPLC analysis at 220 nm). Retention time in HPLC: 4.572.

**LDK(CHO)VNR peptide 2d**. LCMS: m/z 743.4043 (calcd [M+H]$^+$ = 743.3974), 372.2060 (calcd [M+2/2]$^+$ = 372.1987), 395.228 (calcd [M+2/2+Na]$^+$ = 395.1879), m/z 761.4148 (calcd [M+H+H₂O]$^+$ = 761.4152). (HPLC analysis at 220 nm). Retention time in HPLC: 5.285.

**LDKme₁VNR peptide 3d**. LCMS: m/z 758.4581 (calcd [M+H]$^+$ = 758.4519), m/z 379.7299 (calcd [M+2/2]$^+$ = 379.7259). (HPLC analysis at 220 nm). Retention time in HPLC: 4.210.
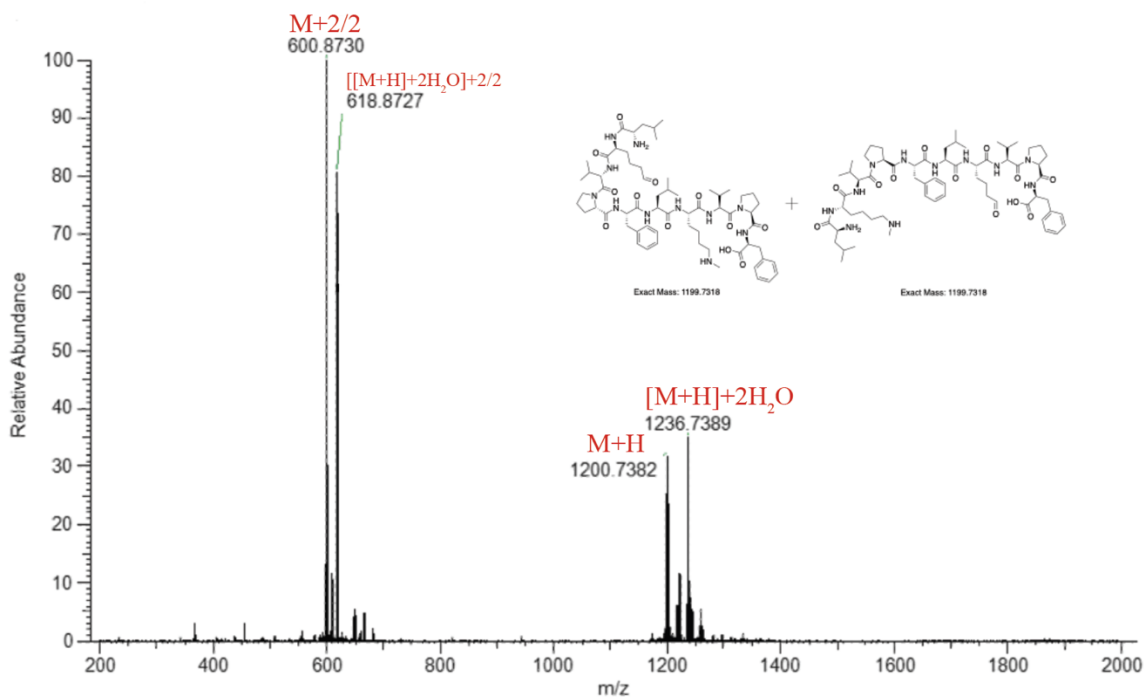
## HPLC Trace of peptide 1d



## MS-Trace of peptide 1d



Exact Mass: 771.4603

## HPLC Trace of Reaction

**MS-Trace of peak 4.210**



**MS-Trace of peak 5.285**

**FSDKKIKme₂K:**



Exact Mass: 1020.6332
**1e**

Selectfluor
Pyridine
Water
1 h, RT

Exact Mass: 991.5702
89%
**2e**
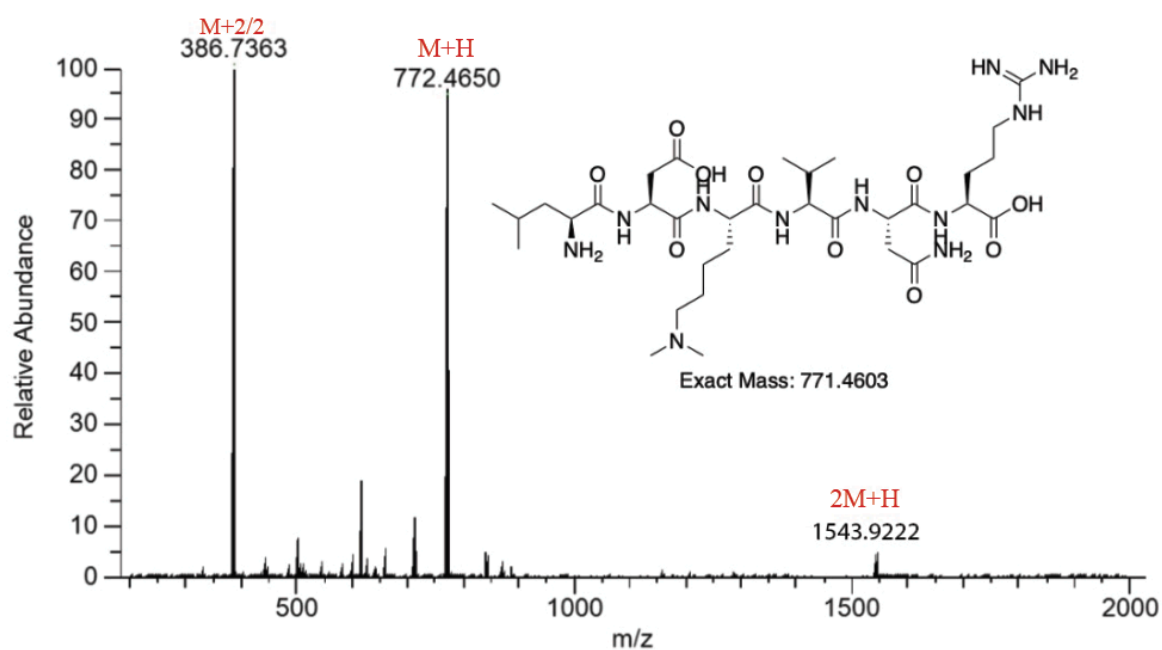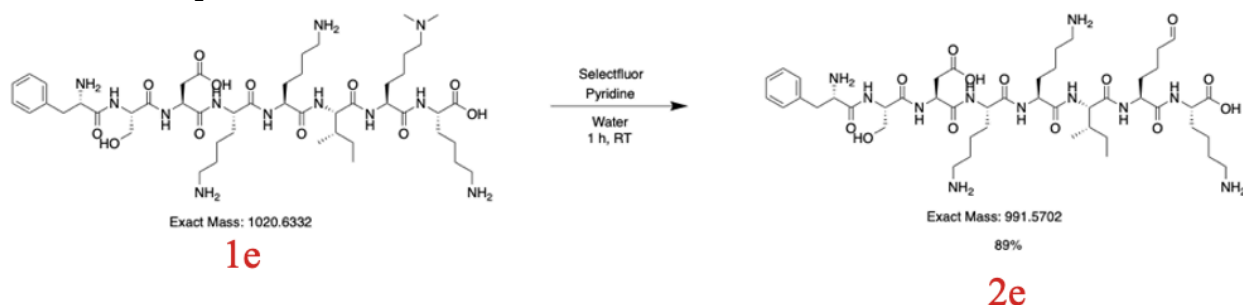
To 1.0 mg of dimethyllysine peptide FSDKKIKme₂K (1e) dissolved in 400 μL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 94%.

**FSDKKIKme₂K peptide 1e**. LCMS: m/z 1021.6397 (calcd $[M+H]^+ = 1021.6404$), m/z 511.3233 (calcd $[M+2/2]^+ = 511.3202$), m/z 341.2182 (calcd $[M+3/3]^+ = 341.5468$) . (HPLC analysis at 220 nm). Retention time in HPLC: 3.014.

**FSDKKIK(CHO)K peptide 2e**. LCMS: m/z 992.5628 (calcd $[M+H]^+= 992.5775$), m/z 496.7852 (calcd $[M+2/2]^+ = 496.7887$), m/z 1010.5767 (calcd $[M+H+H_2O]^+= 1010.5886$). (HPLC analysis at 220 nm). Retention time in HPLC: 3.645.

**HPLC Trace of peptide 1e**
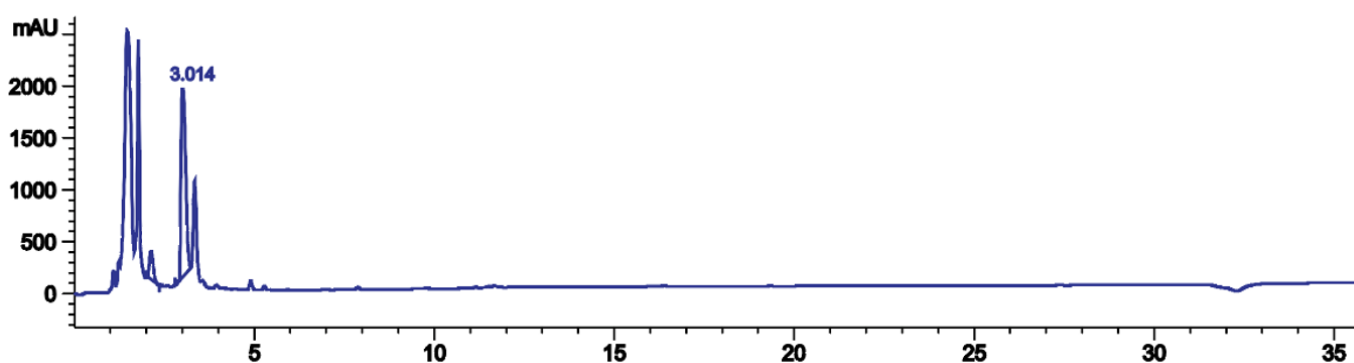
## MS-Trace of peptide 1e



## HPLC Trace of Reaction
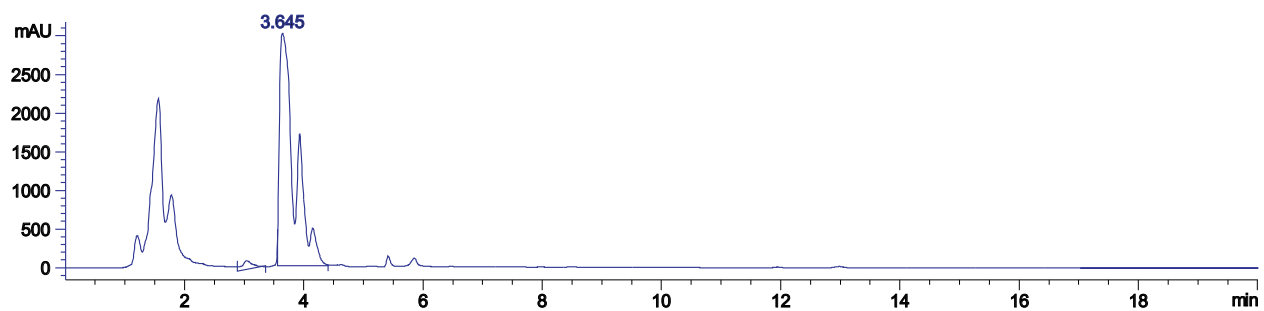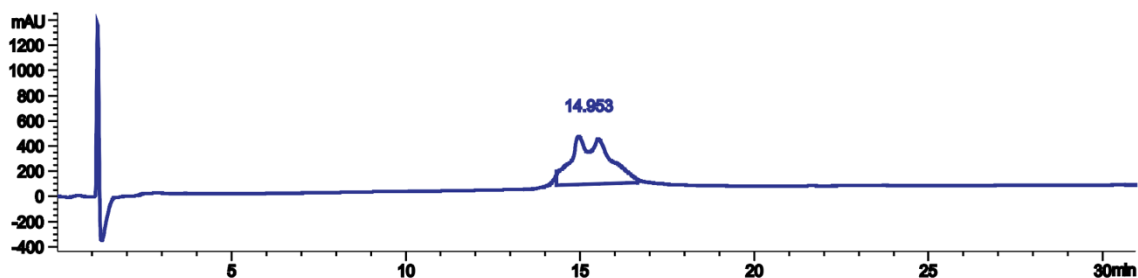
**MS-Trace of peak 3.651**
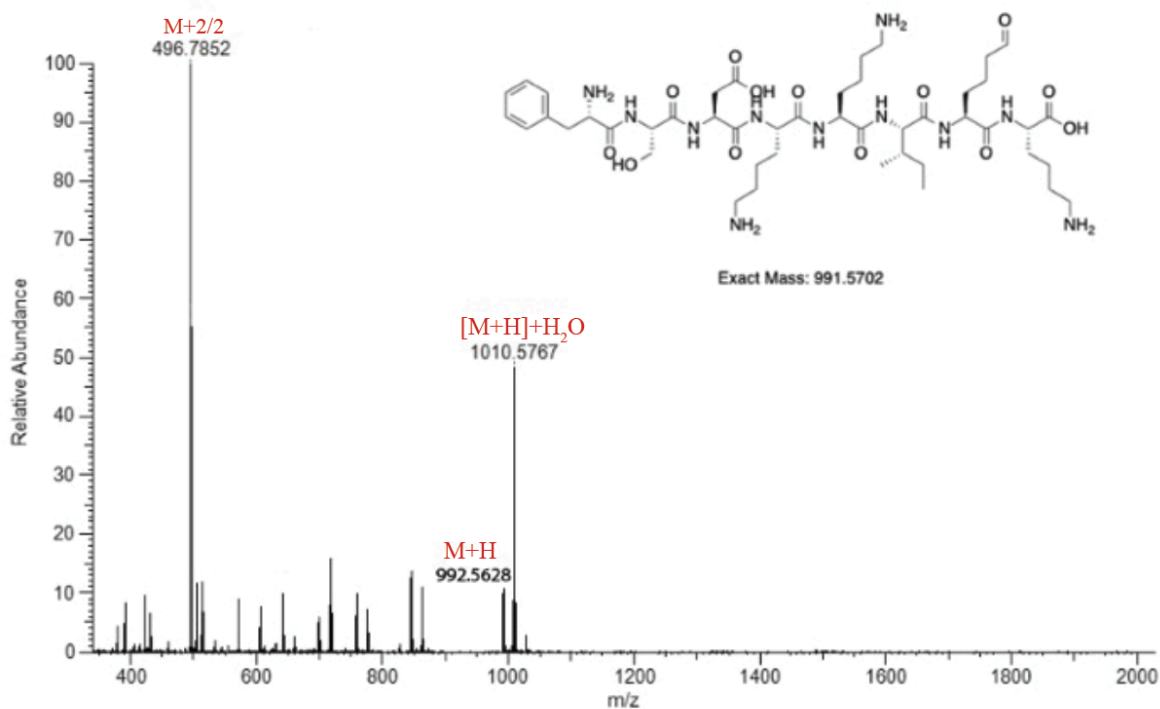


**AKme₂GSKAF(Pra)A:**
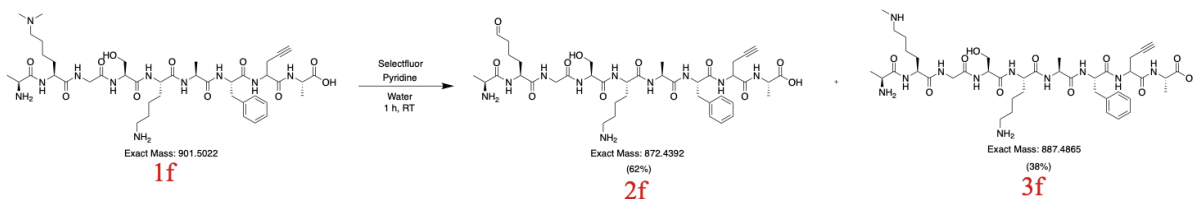


To 1.0 mg of dimethyllysine peptide AKme₂GSKAF(Pra)A (1f) dissolved in 400 μL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 96%.

**AKme₂GSKAF(Pra)A peptide 1f**. LCMS: m/z 902.5069 (calcd [M+H]+ = 902.5094), m/z 451.7568 (calcd [M+2/2]+ = 451.7583). (HPLC analysis at 220 nm). Retention time in HPLC: 4.690.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4474 (calcd [M+H]⁺ = 873.4465), m/z 437.2269 (calcd [M+2/2]⁺ = 437.2232). (HPLC analysis at 220 nm). Retention time in HPLC: 5.145.

**AKme₁GSKAF(Pra)A peptide 3f**. LCMS: m/z 888.4890 (calcd [M+H]$^+$ = 888.4938), m/z 444.7476 (calcd [M+2/2]$^+$ = 444.7445), m/z 296.8342 (calcd [M+3/3]$^+$ = 296.8296). (HPLC analysis at 220 nm). Retention time in HPLC: 3.503.

## HPLC Trace of 1f



## MS-Trace of peptide 1f

# HPLC Trace of Reaction



# MS-Trace of peak 3.503



M+2/2
444.7476

M+3/3
296.8342

M+H
888.4890

Exact Mass: 887.4865
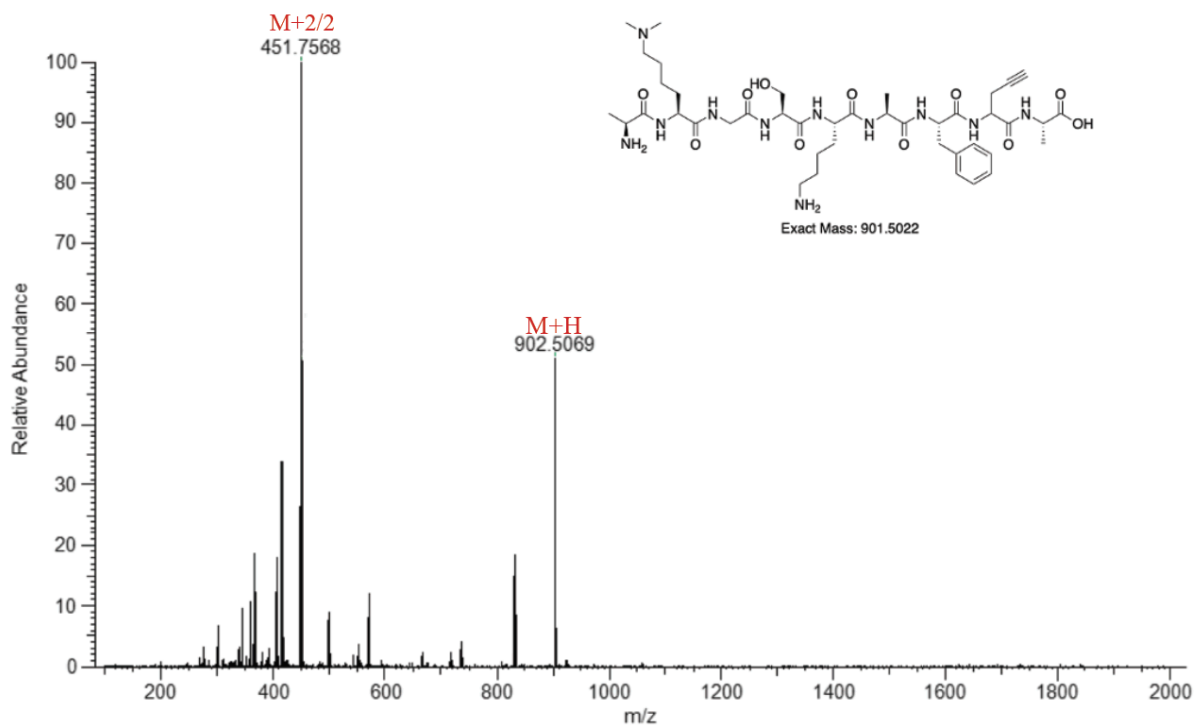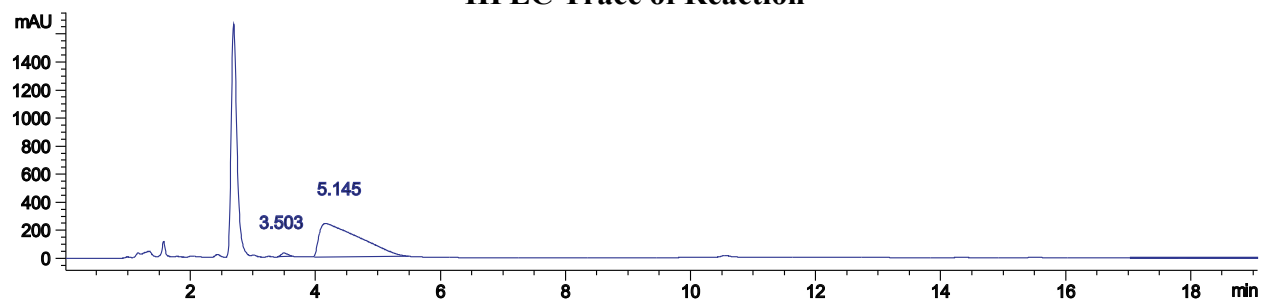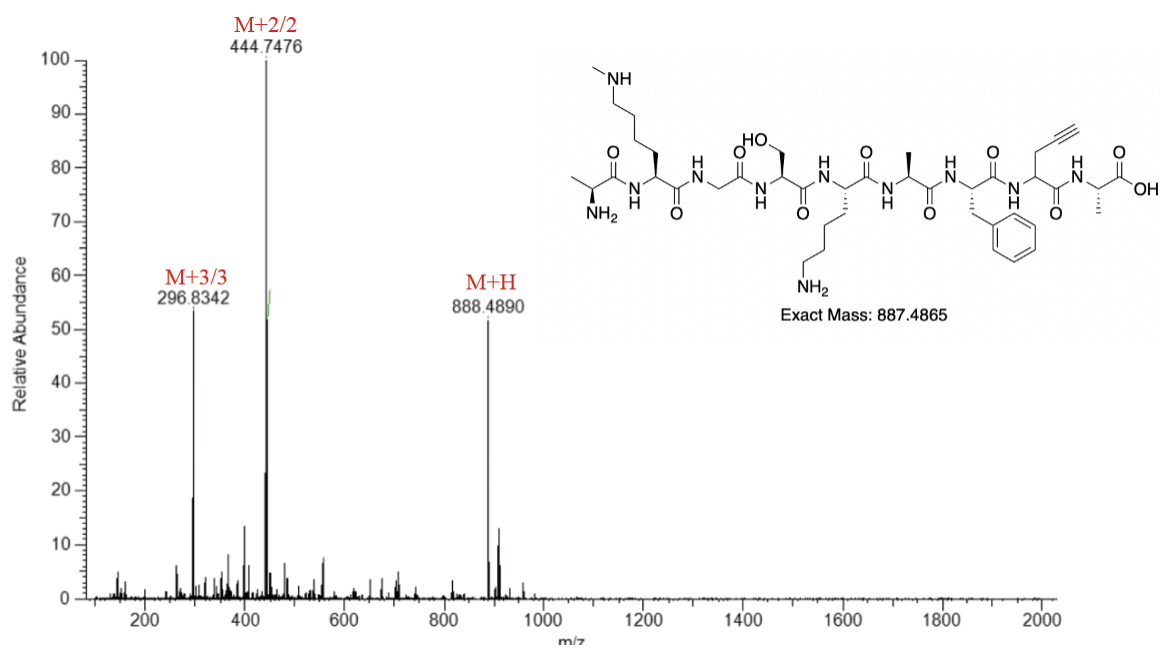
**MS-Trace of peak 5.145**
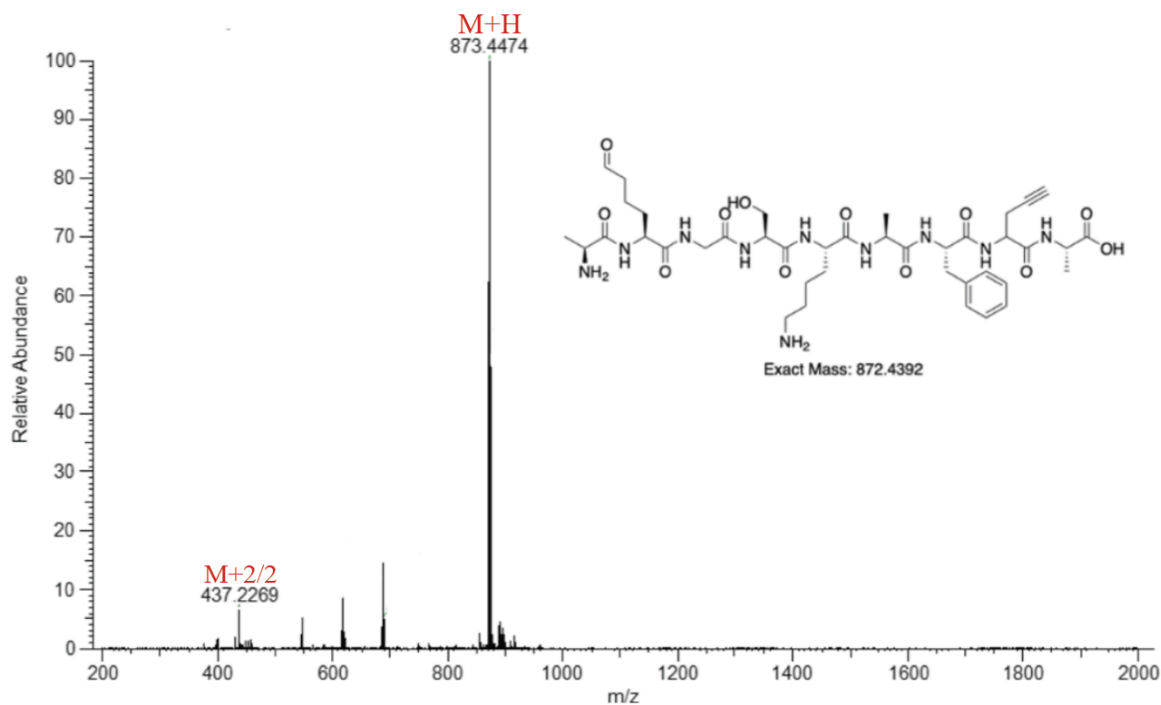


**VFKme₂RN:**
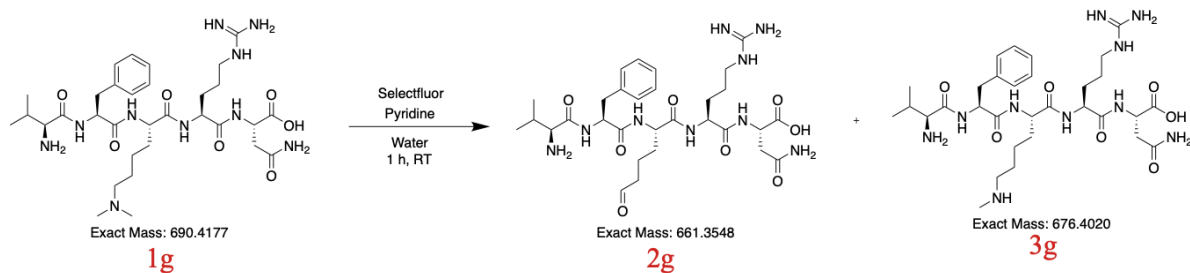


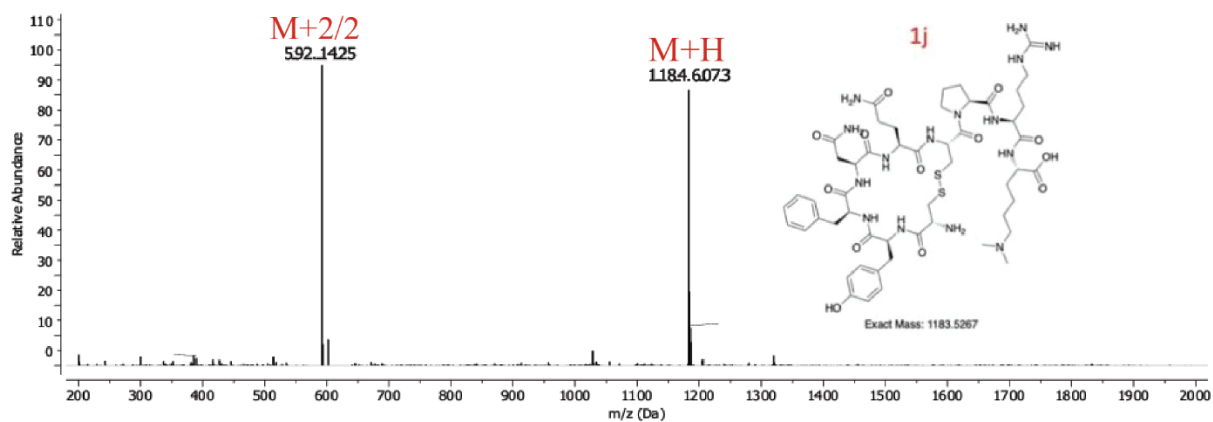To 1.0 mg of dimethyllysine peptide VFKme₂RN (2g) dissolved in 400 μL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was >95%.

**VFKme₂RN peptide 1g**. LCMS: m/z 691.4230 (calcd $[M+H]^+$ = 691.4255), m/z 346.2151 (calcd $[M+2/2]^+$ = 346.2128). (HPLC analysis at 220 nm). Retention time in HPLC: 2.918.

**VFK(CHO)RN peptide 2g.** LCMS: m/z 662.22275 (calcd $[M+H]^+$ = 662.3548), m/z 680.25673 (calcd $[M+H+H_2O]^+$ =1221.6948), m/z 698.26742 (calcd $[M+H+2H_2O]^+$ =1221.6948), m/z

331.62048 (calcd [M+2/2]$^+$ = 331.6774). (HPLC analysis at 220 nm). Retention time in HPLC: 3.346.

**HPLC Trace of peptide 1g**



**MS-Trace of peptide 1g**
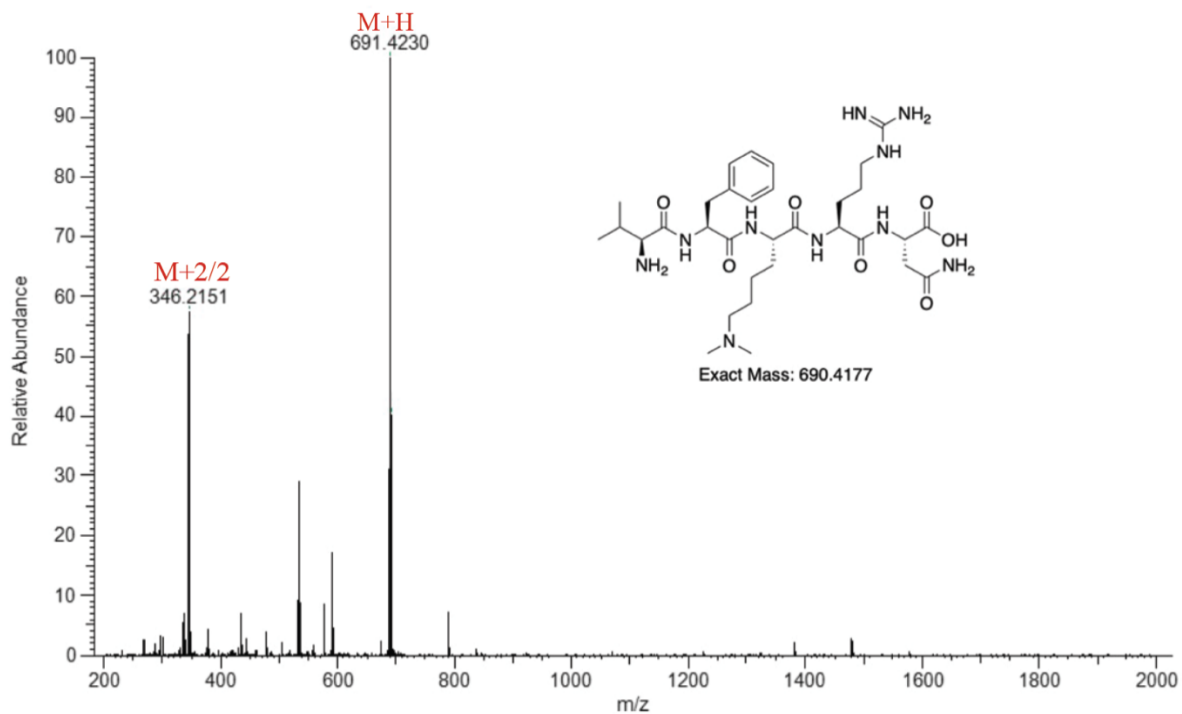


**HPLC Trace of Reaction**
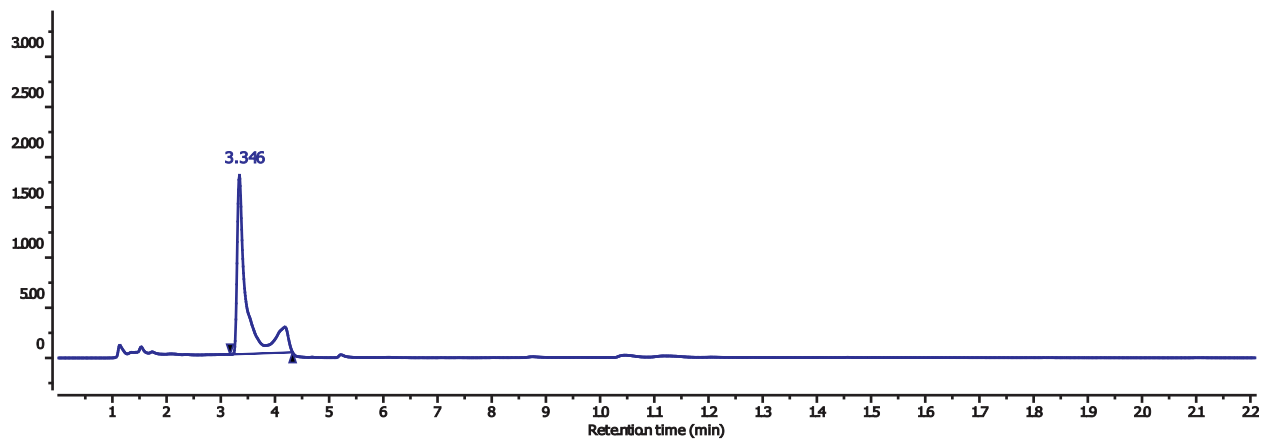
**MS-Trace of peak 3.346**
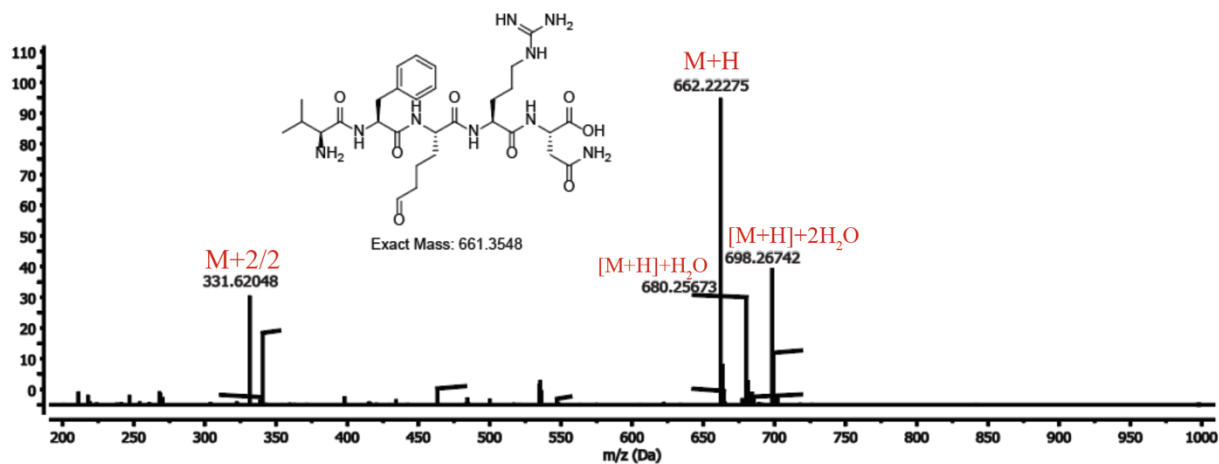


**LKme₂VPFLKme₂VPF:**
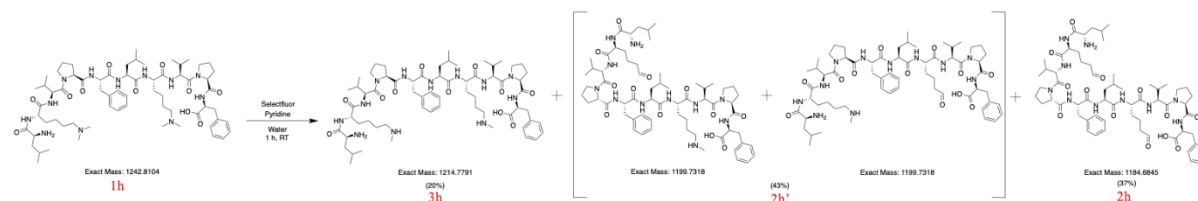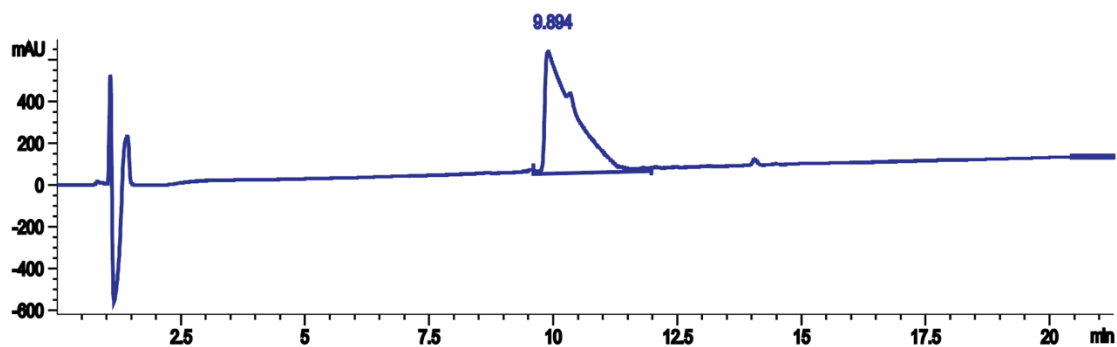


To 1.0 mg of dimethyllysine peptide LKme₂VPFLKme₂VPF (1h) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. The total % conversion to peptide aldehyde was 80%. The % conversion to peptide aldehyde with one modification is 43% and the % conversion to peptide aldehyde with two modifications is 37%.

**LKme₂VPFLKme₂VPF peptide 1h**. LCMS: m/z 1243.8174 (calcd $[M+H]^+$ =1243.8177), m/z 622.4124 (calcd $[M+2/2]^+$ = 622.4088), m/z 415.2775 (calcd $[M+3/3]^+$ = 415.2725). (HPLC analysis at 220 nm). Retention time in HPLC: 10.793.

**LK(CHO)VPFLK(CHO)VPF peptide 2h**. LCMS: m/z 1185.6948 (calcd $[M+H]^+$ = 1185.6918), m/z 1221.6958 (calcd $[M+H+2H_2O]^+$ =1221.6948), m/z 1257.6935 (calcd $[M+H+4H_2O]^+$ = 1257.7351. (HPLC analysis at 220 nm). Retention time in HPLC: 13.988.

**LKme₁VPFLK(CHO)VPF + LK(CHO)VPFLKme₁PF peptides 2h'**. LCMS: m/z 1200.7382 (calcd $[M+H]^+$ = 1200.7391), m/z 1236.7389 (calcd $[M+H+2H_2O]^+$ = 1236.7391), m/z 600.8730 (calcd $[M+2/2]^+$ = 600.8695), m/z 618.8727 (calcd $[[M+2H_2O]+2/2]^+$ = 618.8728). (HPLC analysis at 220 nm). Retention time in HPLC: 11.469.

**LKme₁VPFLKme₁VPF peptide 3h**. LCMS: m/z 1215.7854 (calcd [M+H]⁺ = 1215.7864), m/z 608.3966 (calcd [M+2/2]⁺ = 608.3932), m/z 405.9337 (calcd [M+3/3]⁺ = 405.9288). (HPLC analysis at 220 nm). Retention time in HPLC: 10.158.

**HPLC Trace of peptide 1h**



**MS-Trace of peptide 1h**

**HPLC Trace of Reaction**



**MS-Trace of Two Monomethyl Modifications (peak 10.158)**



Exact Mass: 1214.7791

## MS-Trace of One Monomethyl and One CHO Modification (peak 11.469)



## MS-Trace of di-aldehyde peptide product (peak 13.988)
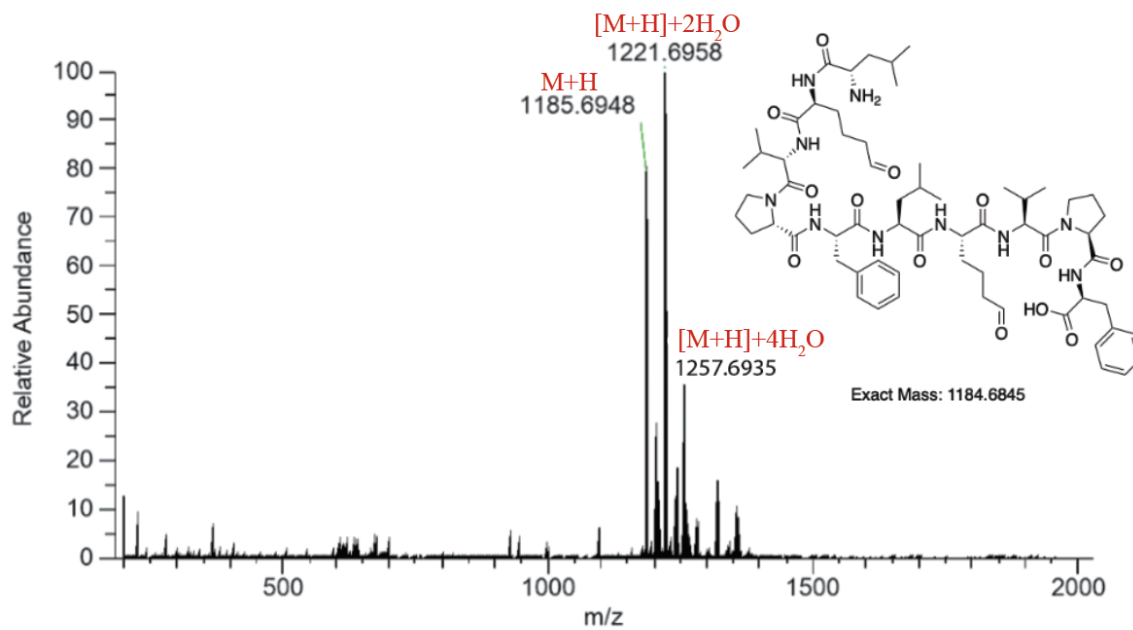
**XIII. Supplementary Figure 5. Modification of dimethyllysine containing cyclic peptides to allysine cyclic peptide products.**

**cycLILKme₂PF:**
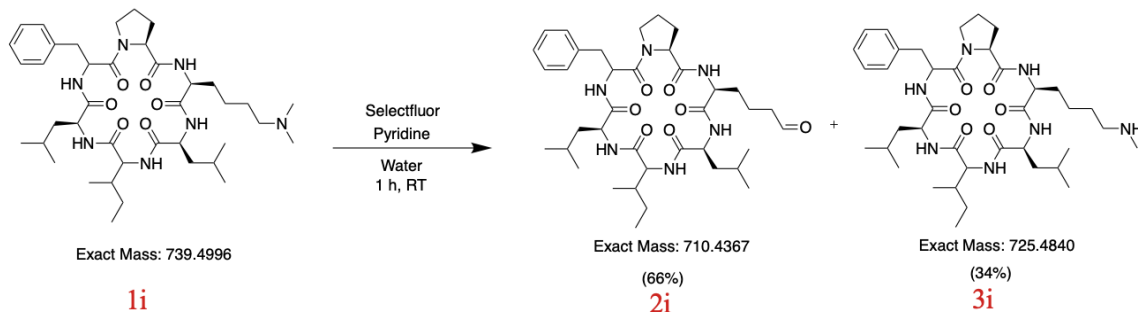


To 1.0 mg of dimethyllysine cyclic peptide LILKme₂PF (1i) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to cyclic peptide aldehyde was 66%.

**LILKme₂PF cyclic peptide 1i**. LCMS: m/z 740.5036 (calcd [M+H]$^+$ = 740.5069), m/z 763.5052 (calcd [M+Na]$^+$ = 763.5069), 1502.9341 (calcd [2M+Na]$^+$ = 1502.9962), 1517.9373 (calcd [2M+K]$^+$ = 1517.9710). (HPLC analysis at 220 nm). Retention time in HPLC: 14.953.

**LILK(CHO)PF cyclic peptide 2i**. LCMS: m/z 711.4445 (calcd [M+H]$^+$ = 711.4440), m/z 1443.8643 (calcd [2M++Na]$^+$ = 1443.8530), m/z 733.4229 (calcd [M+Na]$^+$ = 733.4265). (HPLC analysis at 220 nm). Retention time in HPLC: 18.200.

**LILKme₁PF cyclic peptide 3i**. LCMS: m/z 726.4906 (calcd [M+H]$^+$ = 726.4907), m/z 749.4926 (calcd [M+Na]$^+$ = 749.4805). (HPLC analysis at 220 nm). Retention time in HPLC: 15.299.

**HPLC Trace of cyclic peptide 1i**
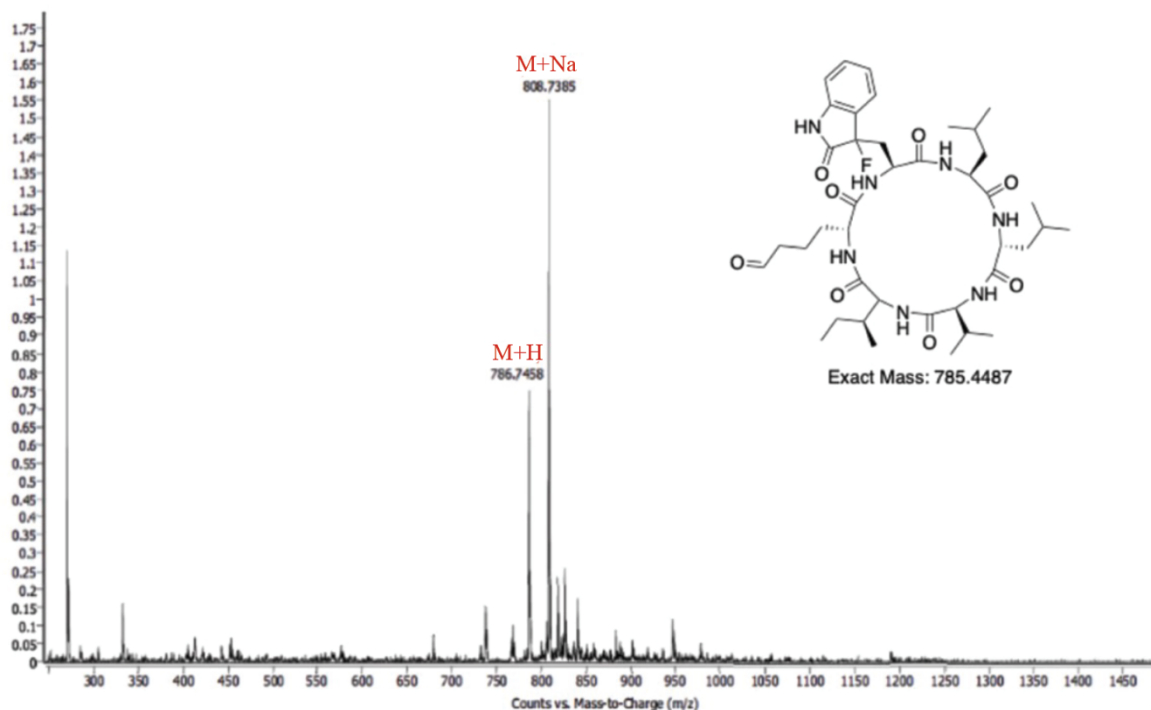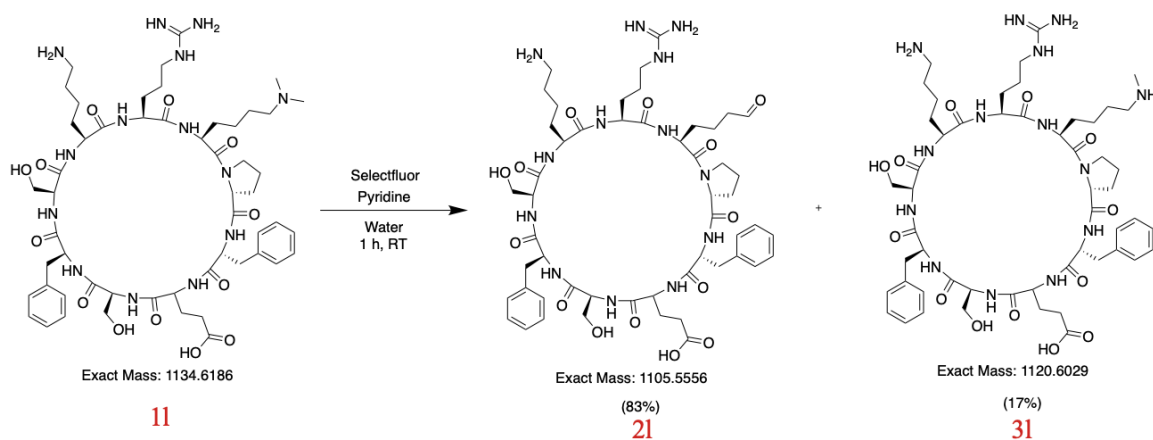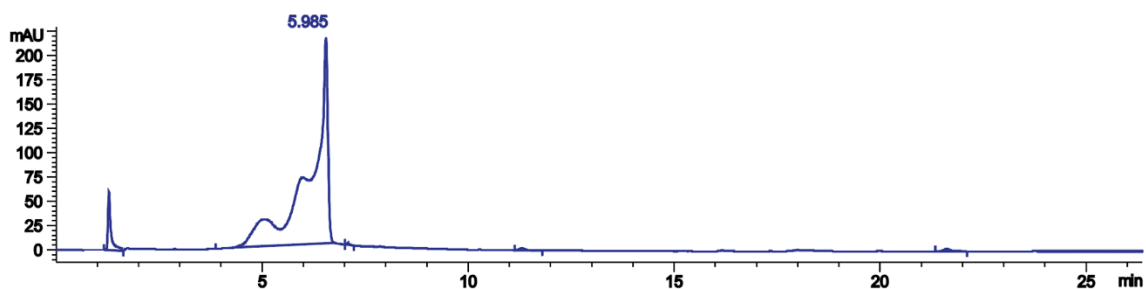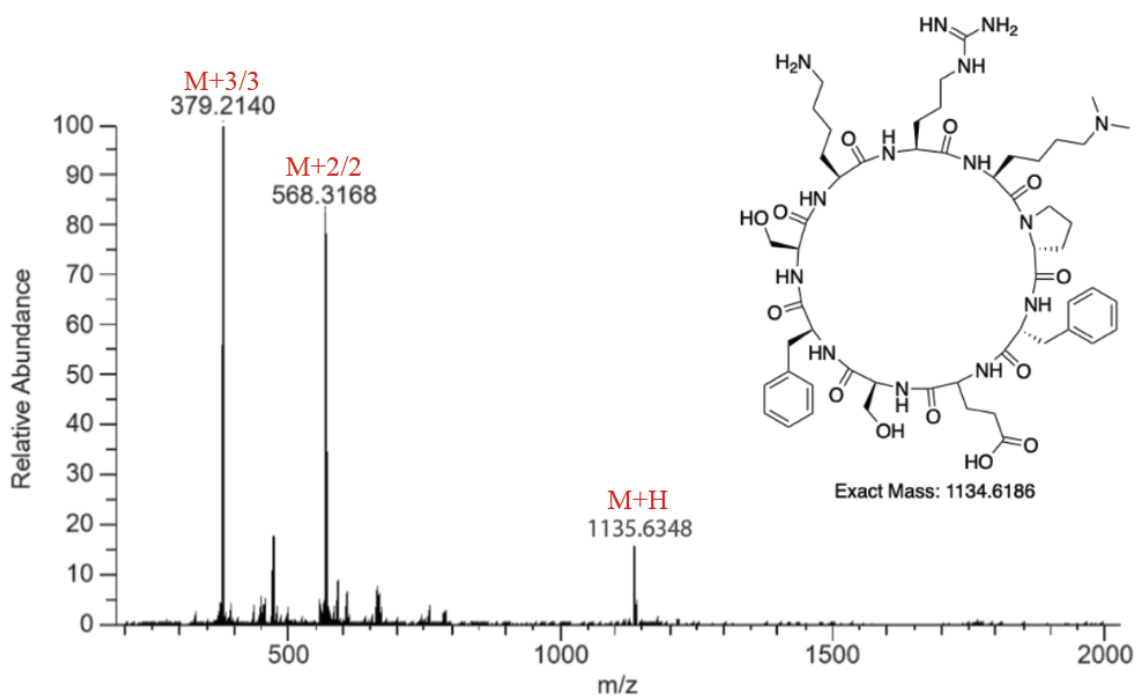
**MS-Trace of cyclic peptide 1i**
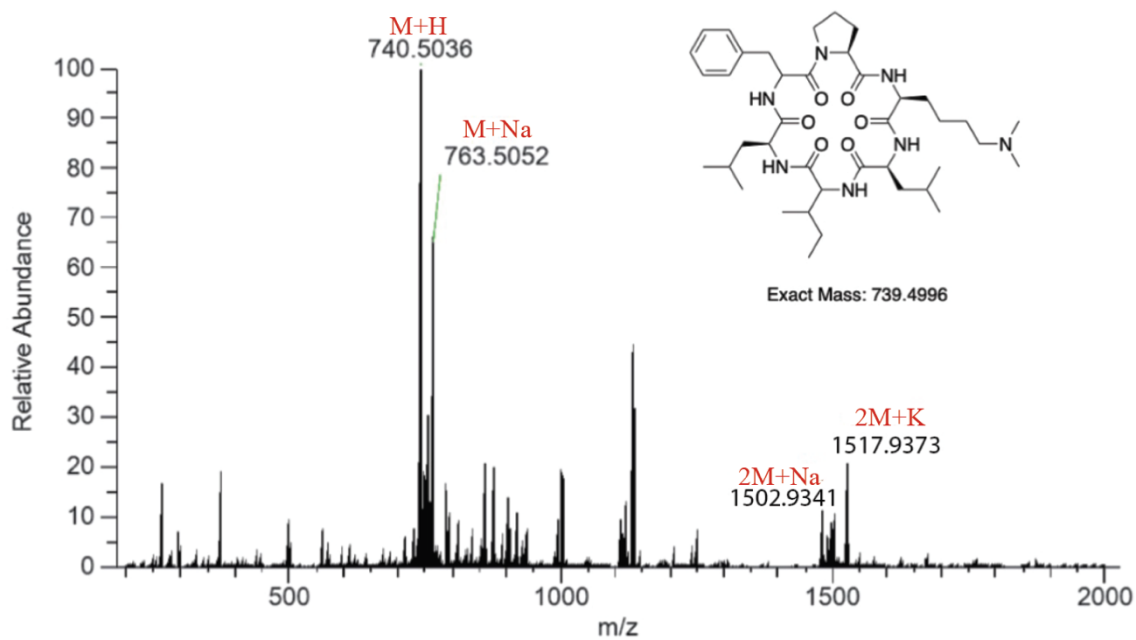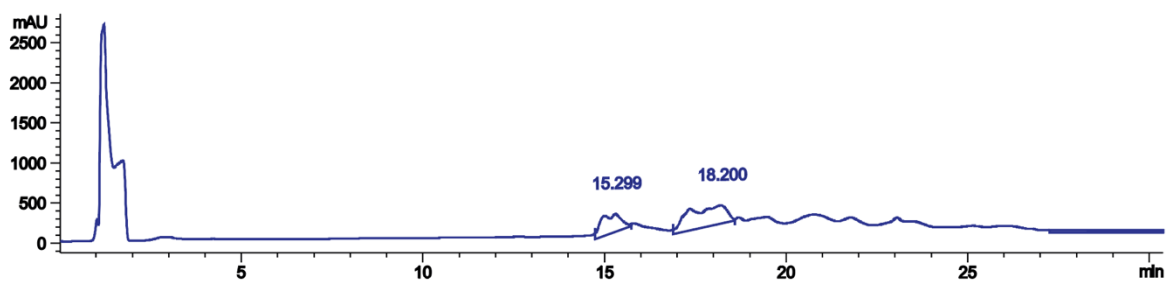


Exact Mass: 739.4996

**HPLC Trace of Reaction**

**MS-Trace of peak 15.299**
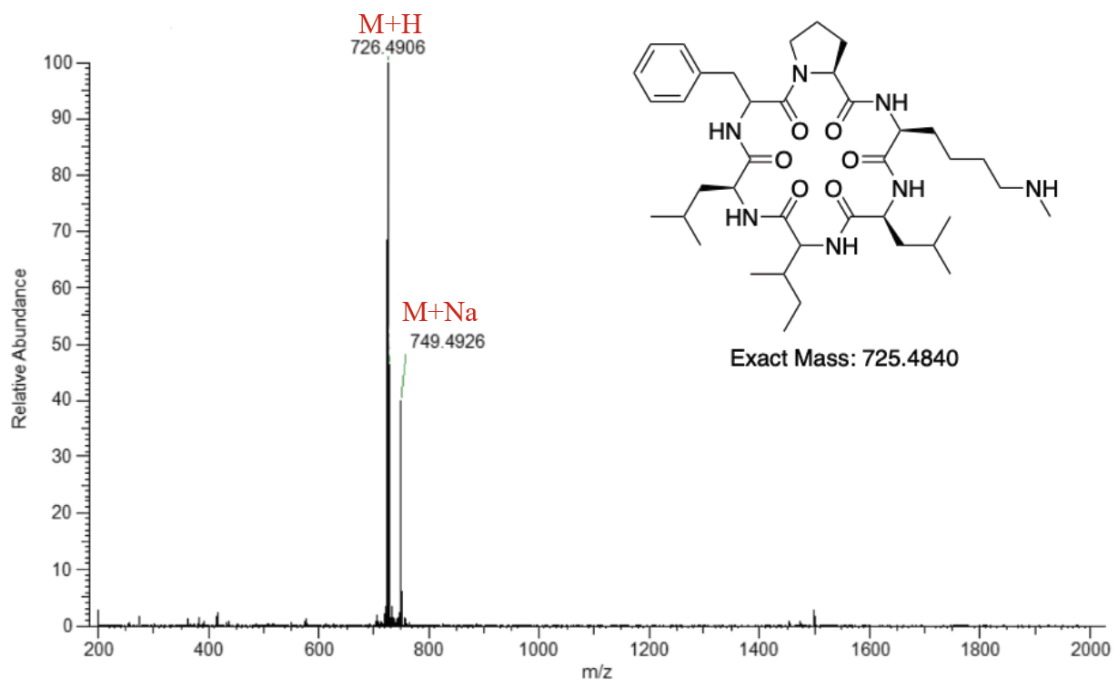


M+H
726.4906

M+Na
749.4926

Exact Mass: 725.4840

**MS-Trace of peak 18.200**



M+H
711.4445

M+Na
733.4229

2M+Na
1443.8643

Exact Mass: 710.4367

**cyc-CYFNQCPRKme₂-disulfide:**



Exact Mass: 1183.5267

**1j**

Selectfluor
Pyridine

Water
1 h, RT

Exact Mass: 1154.4637

(99%)

**2j**
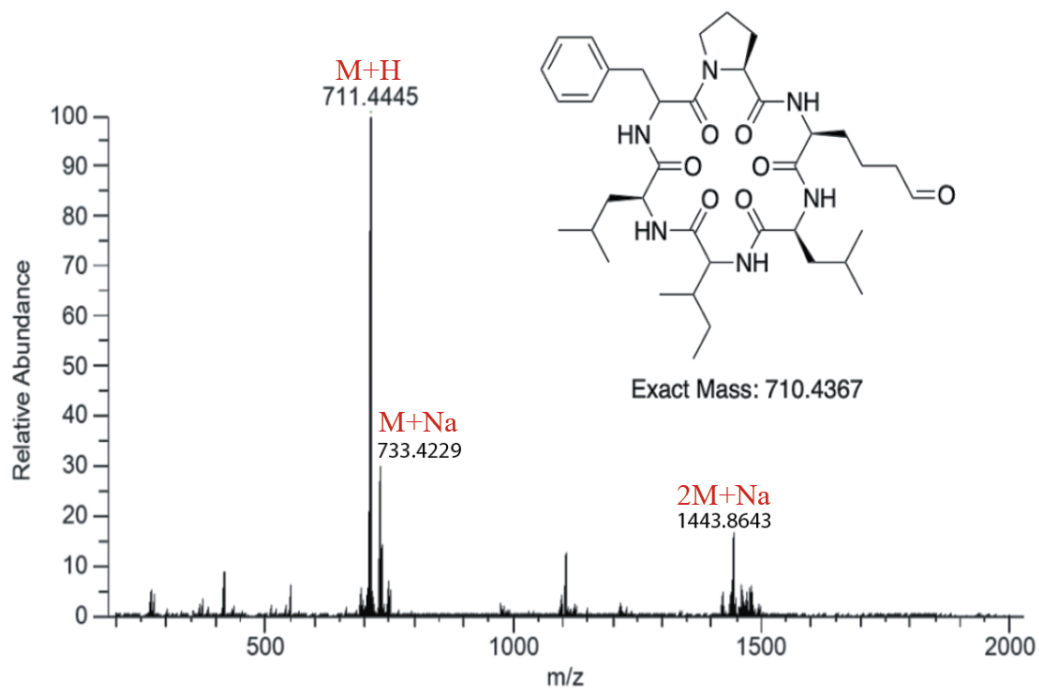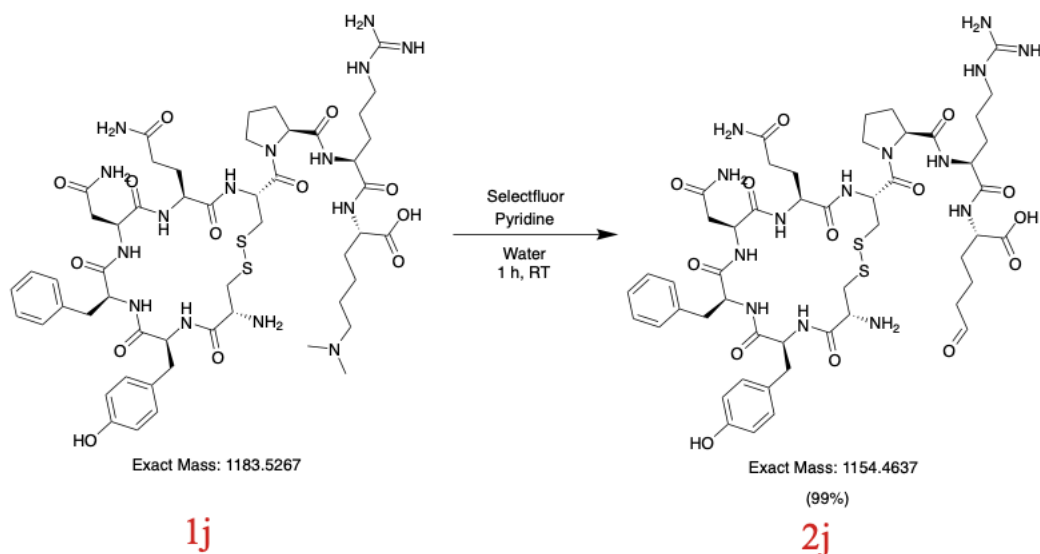
To 1.0 mg of dimethyllysine cyclic peptide CYFNQCPRKme₂ (1j) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to cyclic peptide aldehyde was 99%.

**CYFNQCPRKme₂ cyclic peptide 1j**. LCMS: m/z 1184.6073 (calcd $[M+H]^+ = 1183.5340$), m/z 592.1425 (calcd $[M+2/2]^+ = 592.7670$). (HPLC analysis at 220 nm). Retention time in HPLC: 6.089.

**CYFNQCPRK(CHO) cyclic peptide 2j**. LCMS: m/z 1155.68258 (calcd $[M+H]^+ = 1154.46441$), m/z 578.37614 (calcd $[M+2/2]^+ = 578.3750$), m/z 614.48191 (calcd $[(M+2/2)+2H_2O]^+ = 614.4740$). (HPLC analysis at 220 nm). Retention time in HPLC: 9.664.

**HPLC Trace of cyclic peptide 1j**
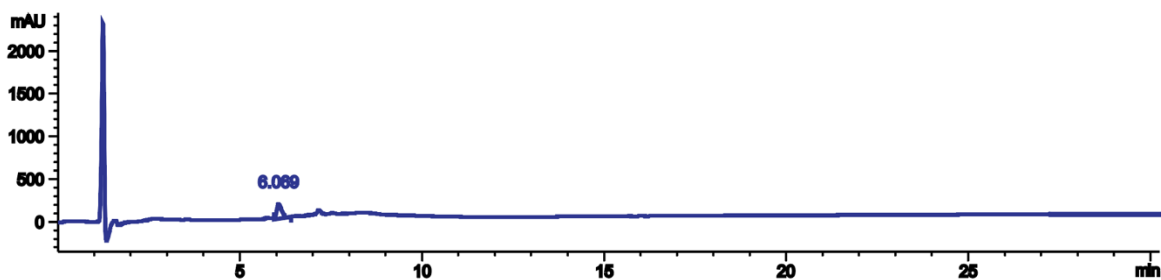
**MS-Trace of cyclic peptide 1j**



**HPLC Trace of Reaction**
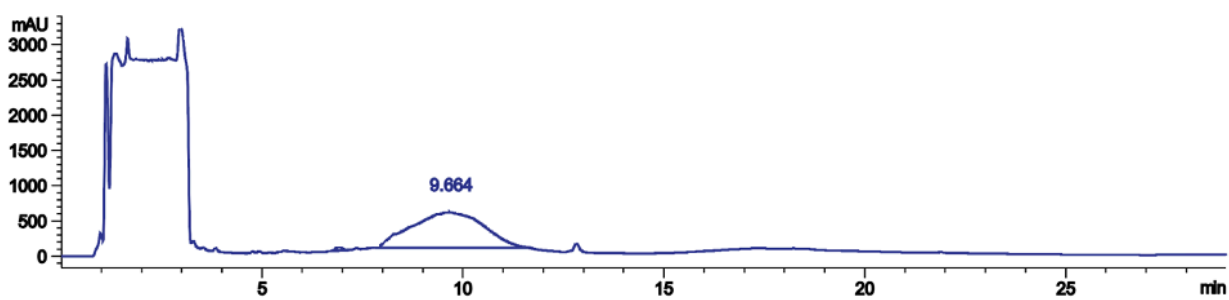


**MS-Trace of peak 9.664**

**cycIKme₂WLLV:**



Exact Mass: 780.5262
**1k**

Exact Mass: 785.4487
(60%)
**2k**

Exact Mass: 800.4960
(39%)
**3k**
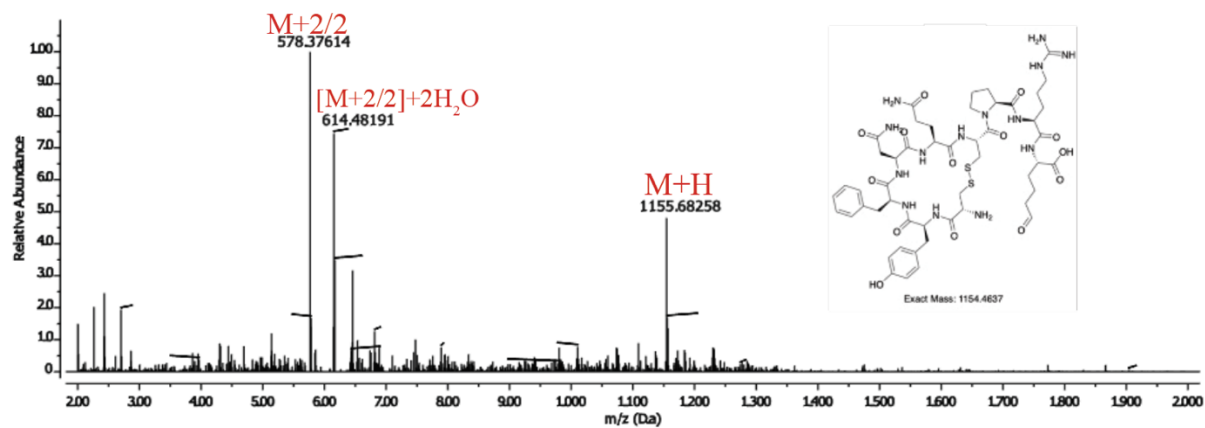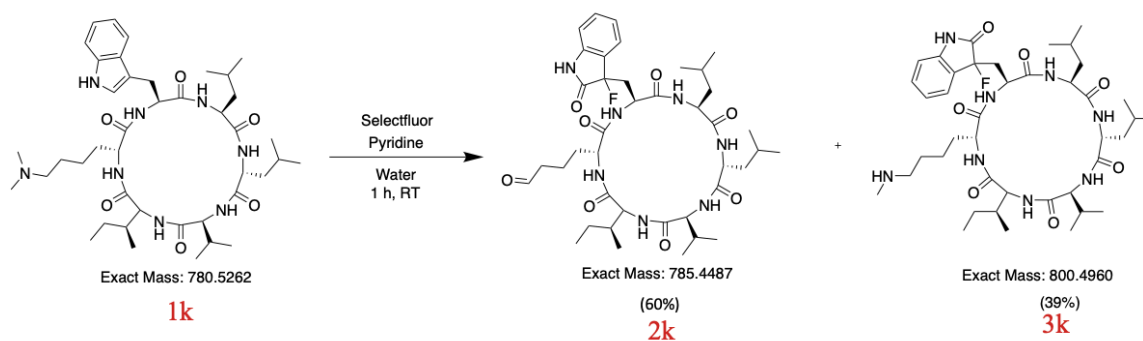
To 1.0 mg of dimethyllysine cyclic peptide IKme₂WLLV (1k) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to cyclic peptide aldehyde was 60%. We also observed oxidation and fluorination of tryptophan under the reaction condition.

**IKme₂WLLV cyclic peptide 1k.** LCMS: m/z 781.4158 (calcd $[M+H]^+ = 781.5335$), m/z 803.4025 (calcd $[M+Na]^+ = 803.5262$), m/z 1562.0595 (calcd $[2M+H]^+ = 1562.0597$), m/z 1583.9216 (calcd $[2M+Na]^+ = 1583.9216$). (HPLC analysis at 220 nm). Retention time in HPLC: 9.894.

**IK(CHO)WLLV cyclic peptide 2k.** LCMS: m/z 786.7458 (calcd $[M+H]^+ = 786.4560$), m/z 808.7385 (calcd $[M+Na]^+ = 808.4560$). (HPLC analysis at 220 nm). Retention time in HPLC: 14.204.

**IKme₁WLLV cyclic peptide 3k.** LCMS: m/z 801.54712 (calcd $[M+H]^+ = 801.5445$), m/z 401.27448 (calcd $[M+2/2]^+ = 401.2645$). (HPLC analysis at 220 nm). Retention time in HPLC: 10.560.

**HPLC Trace of cyclic peptide 1k**

**MS-Trace of cyclic peptide 1k**
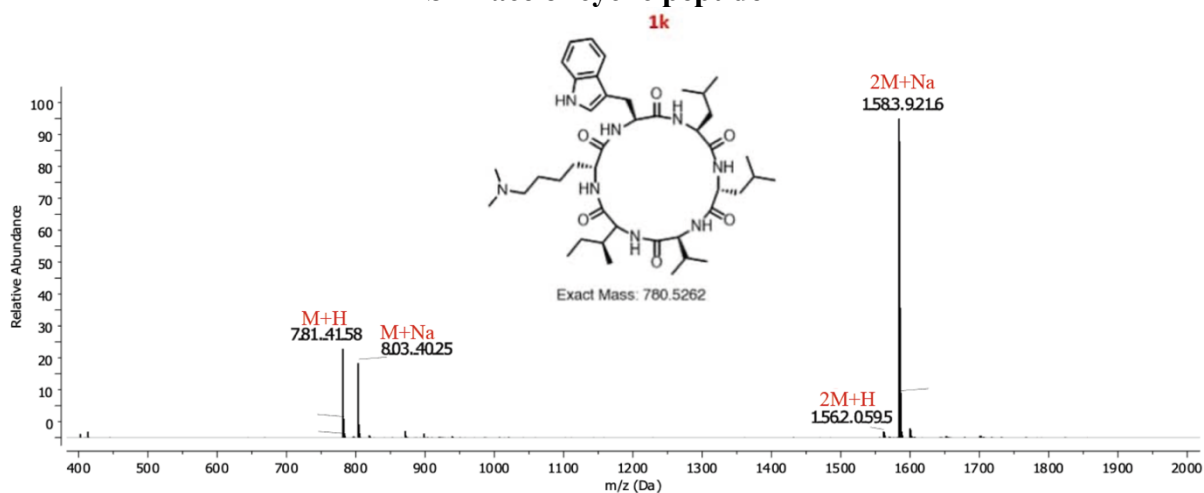


**HPLC Trace of cyclic peptide 1k**
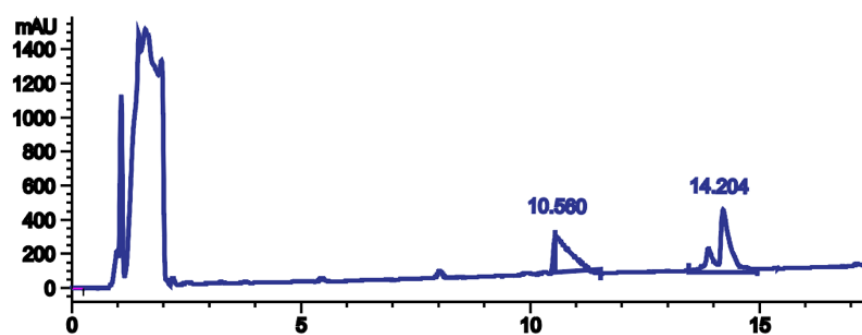


**MS-Trace of peak 10.560**

**MS-Trace of peak 14.204**



M+Na
808.7385

M+H
786.7458

Exact Mass: 785.4487

**cycFSKRKme2(D-p)(D-f)ES:**



Exact Mass: 1134.6186
**1l**

Selectfluor
Pyridine

Water
1 h, RT

Exact Mass: 1105.5556
(83%)
**2l**

+

Exact Mass: 1120.6029
(17%)
**3l**
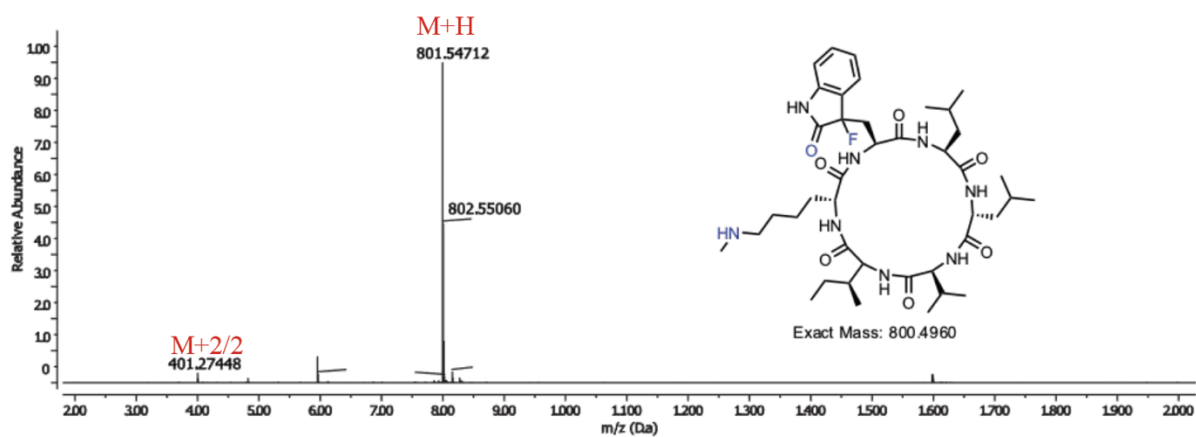
To 1.0 mg of dimethyllysine  cyclic peptide FSKRKme2(D-p)(D-f)ES (1l) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.   % conversion to cyclic peptide aldehyde was 83%.

**FSKRKme₂(D-p)(D-f)ES cyclic peptide 1l**. LCMS: m/z 379.2140 (calcd [M+3/3]⁺ = 379.2062), m/z 1135.6348 (calcd [M+H]⁺ = 1135.6259), m/z 568.3168 (calcd [M+2/2]⁺ = 568.3129). (HPLC analysis at 220 nm). Retention time in HPLC: 5.985.

**FSKRK(CHO)(D-p)(D-f)ES cyclic peptide 2l**. LCMS: m/z 1106.5598 (calcd [M+H]⁺ = 1106.5629), m/z 553.7853 (calcd [M+2/2]⁺ = 553.7814), m/z 369.4252 (calcd [M+3/3]⁺ = 369.5185), m/z 1129.5490 (calcd [M+Na]⁺ = 1129.5527). (HPLC analysis at 220 nm). Retention time in HPLC: 8.918.

**HPLC Trace of cyclic peptide 1l**



**MS-Trace of cyclic peptide 1l**

## HPLC Trace of Reaction



6.558    8.918

Retention time (min)

## MS-Trace of peak 8.918



M+3/3
369.4252

M+2/2
553.7853

M+H
1106.5598
M+Na
1129.5490

Exact Mass: 1105.5556

Relative Abundance

m/z

## cycLKme₂VPFLKme₂VPF:



Exact Mass: 1224.7998
1m

Selectfluor
Pyridine
Water
1 h, RT

Exact Mass: 1196.7685
(10%)
3m

Exact Mass: 1181.7213
(38%)
2m'

Exact Mass: 1181.7213

Exact Mass: 1166.6740
(52%)
2m
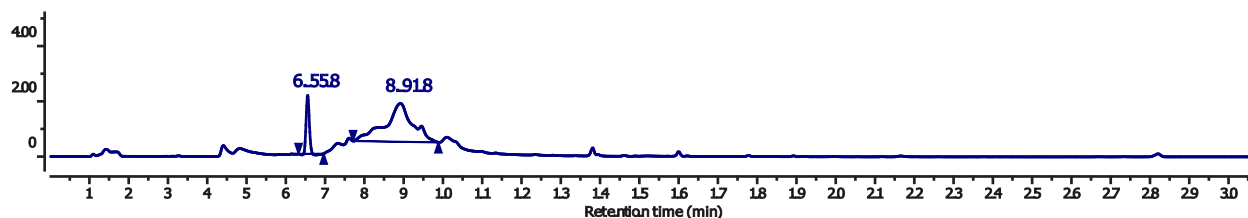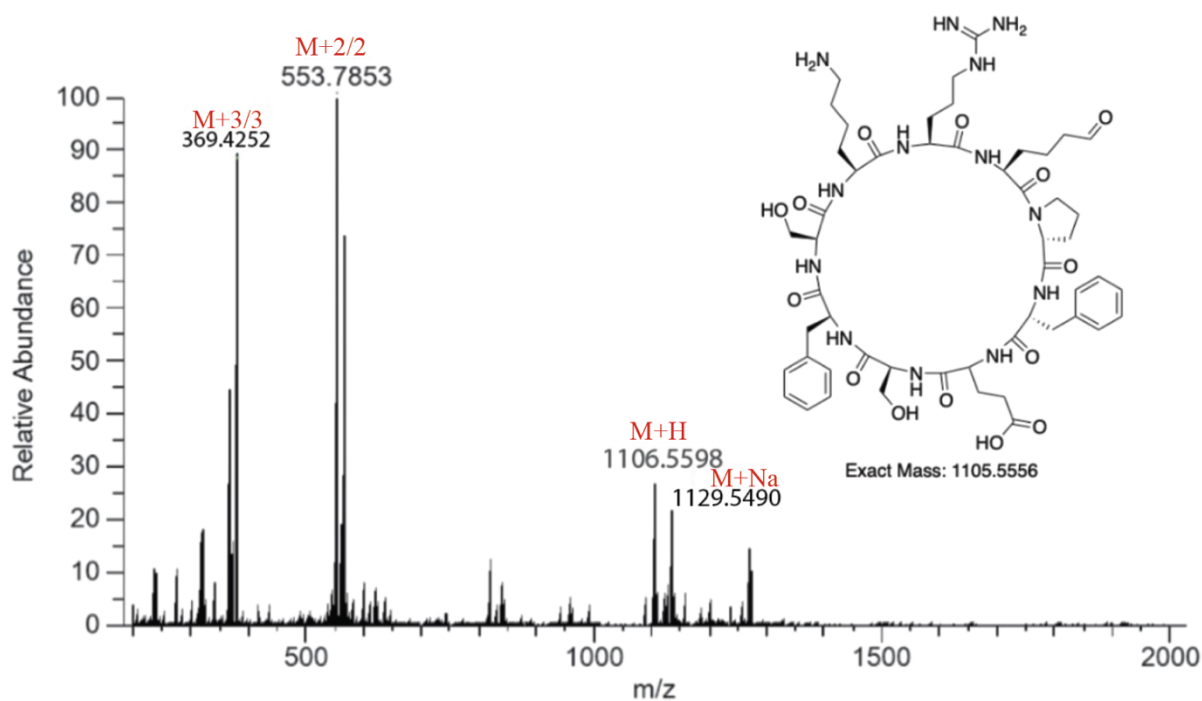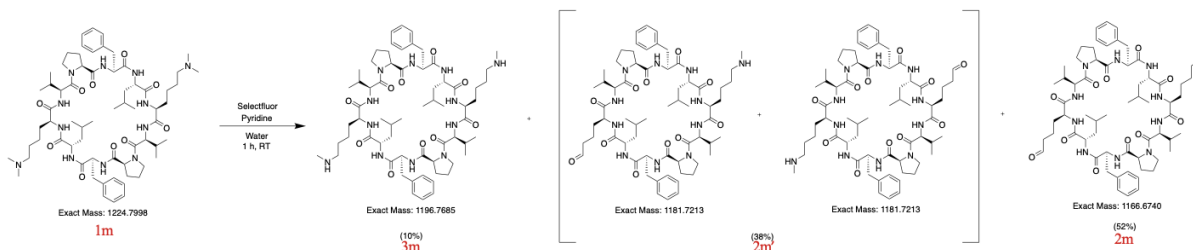
To 1.0 mg of dimethyllysine  cyclic peptide LKme₂VDFLKme₂VPF (1m) dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.  %

conversion to peptide aldehyde was 90%. The % conversion to peptide aldehyde with one modification is 38% and the % conversion to cyclic peptide aldehyde with two modifications is 52%.

**LKme₂VPFLKme₂VPF cyclic peptide 1m**. LCMS: m/z 1225.8007 (calcd [M+H]+ = 1225.8071), m/z 613.4041(calcd [M+2/2]+ = 613.4035). (HPLC analysis at 220 nm). Retention time in HPLC: 14.160.

**LK(CHO)VPFLK(CHO)VPF cyclic peptide 2m**. LCMS: m/z 1167.6793 (calcd [M+H]$^+$ = 1167.6812), m/z 584.3432 (calcd [M+H]$^+$ = 584.3397). (HPLC analysis at 220 nm). Retention time in HPLC: 21.967.

**LKme₁VPFLKme₁VPF cyclic peptide 3m**. LCMS: m/z 1197.7699 (calcd [M+H]$^+$ = 1197.5855), m/z 599.3886 (calcd [M+2/2]$^+$ = 599.2927). (HPLC analysis at 220 nm). Retention time in HPLC: 13.944.

**LK(CHO)VPFLKme₁VPF + LKme₁VDFLK(CHO)VPF cyclic peptides 3m'**. LCMS: m/z 1182.7278 (calcd [M+H]$^+$ = 1182.7285), m/z 1218.7291 (calcd [M+H+2H2O]$^+$ = 1218.7285). (HPLC analysis at 220 nm). Retention time in HPLC: 16.234.

### HPLC Trace of cyclic peptide 1m

**MS-Trace of cyclic peptide 1m**



M+2/2
613.4041

M+H
1225.8007

Exact Mass: 1224.7998

**HPLC Trace of Reaction**
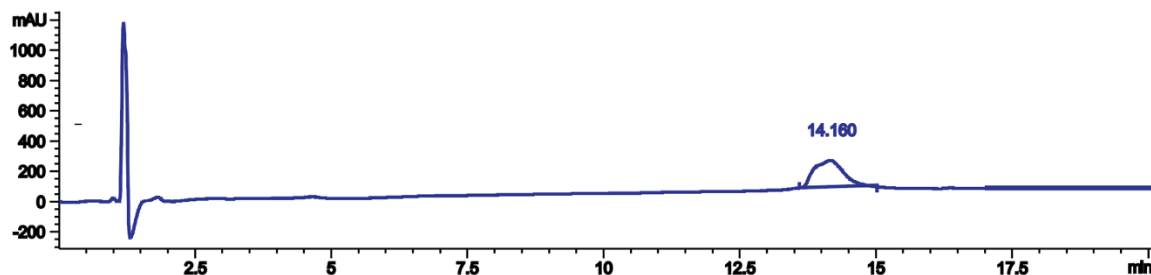


13.944

16.234

21.967

**MS-Trace of Two Monomethyls 3m (peak 13.944)**
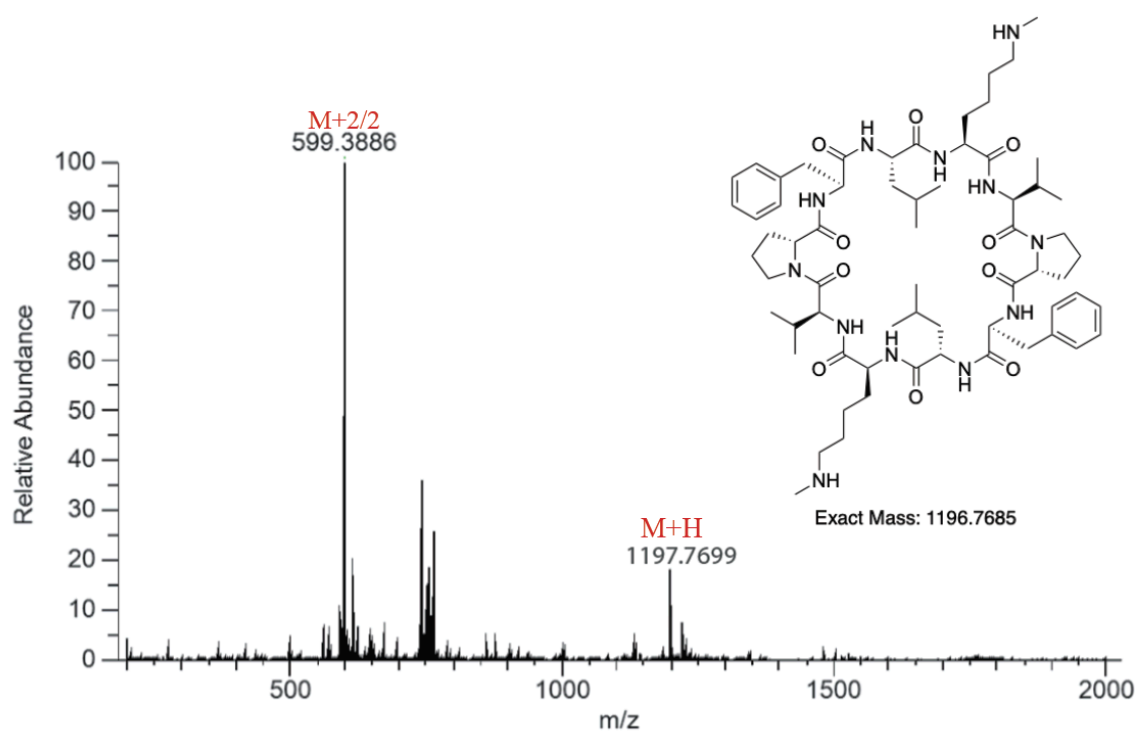


**MS-Trace of One Modification + One Monomethyl 3m' (peak 16.234)**
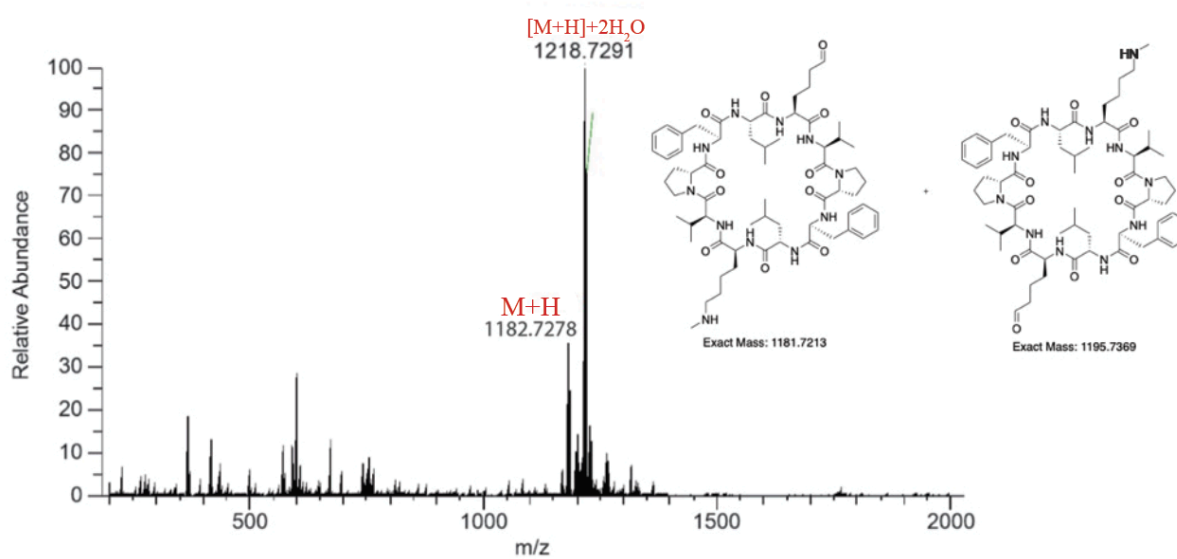
**MS-Trace of Reaction Two aldehyde modification 2m (peak 21.967)**
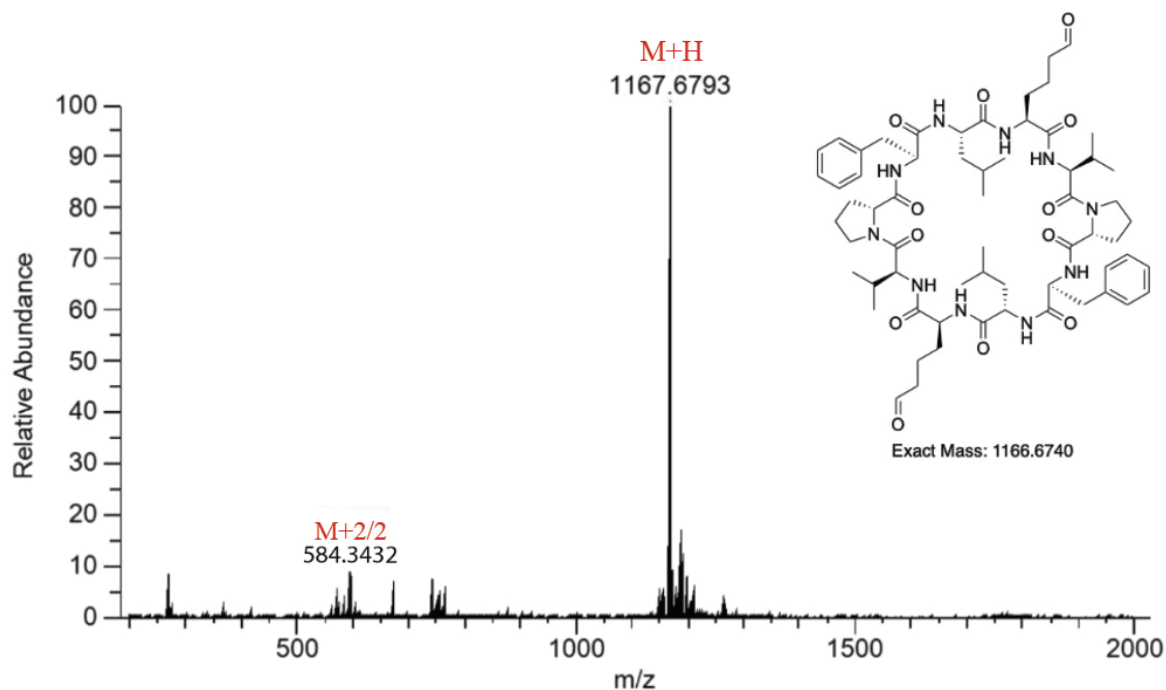


**Control Experiment with Free thiol containing FCV peptide:**
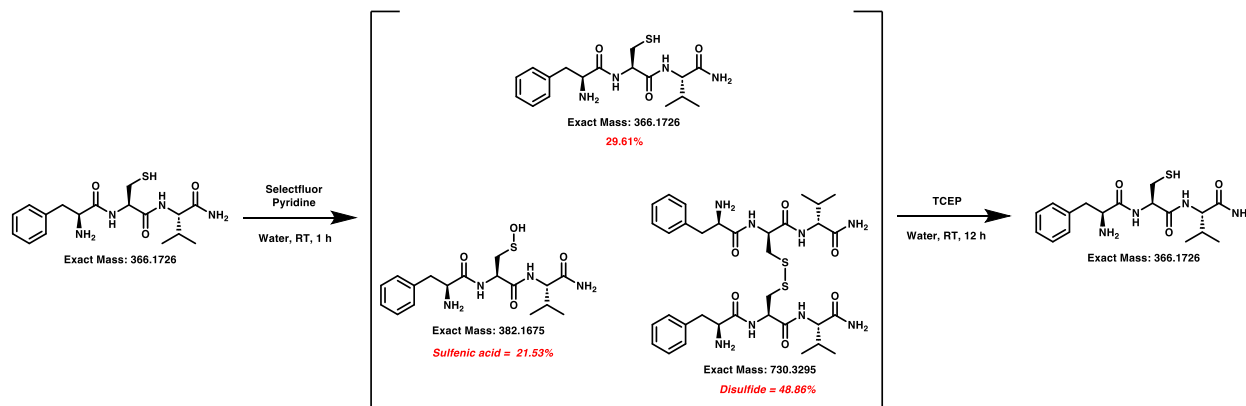


To 1.0 mg of FCV peptide dissolved in 400 µL of water, was added 2 eq of selectfluor and 5 eq of pyridine. The reaction mixture was stirred for 1 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. 48.86% of disulfide was observed, 21.53% sulfenic acid was observed, and 29.61% of FCV peptide observed. The addition of 50 equiv. of TCEP to the reaction for 12 h led to the regeneration of FCV peptide.

**FCV**. LCMS: m/z 367.1788 (calcd $[M+H]^+$ = 367.1798), m/z 733.3524 (calcd $[2M+H]^+$ = 733.3524). (HPLC analysis at 220 nm). Retention time in HPLC: 6.785.

**FCV-sulfenic acid**. LCMS: m/z 383.1750 (calcd [M+H]$^+$ = 383.1748), m/z 405.1566 (calcd [M+Na]$^+$ =405.1567), m/z 765.3423 (calcd [2M+H]$^+$ =765.3425). (HPLC analysis at 220 nm). Retention time in HPLC: 5.840.

**FCV-disulfide**. LCMS: m/z 731.3361 (calcd [M+H]$^+$ = 731.3368), m/z 366.1625 (calcd [M+2/2]$^+$ =366.1684), m/z 753.3179 (calcd [(M+Na]$^+$ =753.3187), m/z 1461.6590 (calcd [2M+H]$^+$ = 1461.6663). (HPLC analysis at 220 nm). Retention time in HPLC: 7.806.

### HPLC Trace of FCV peptide
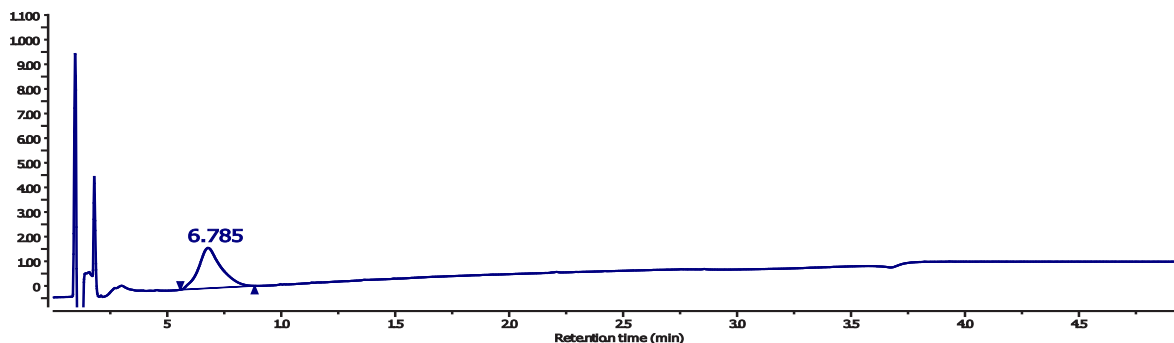


### MS Trace of FCV peptide
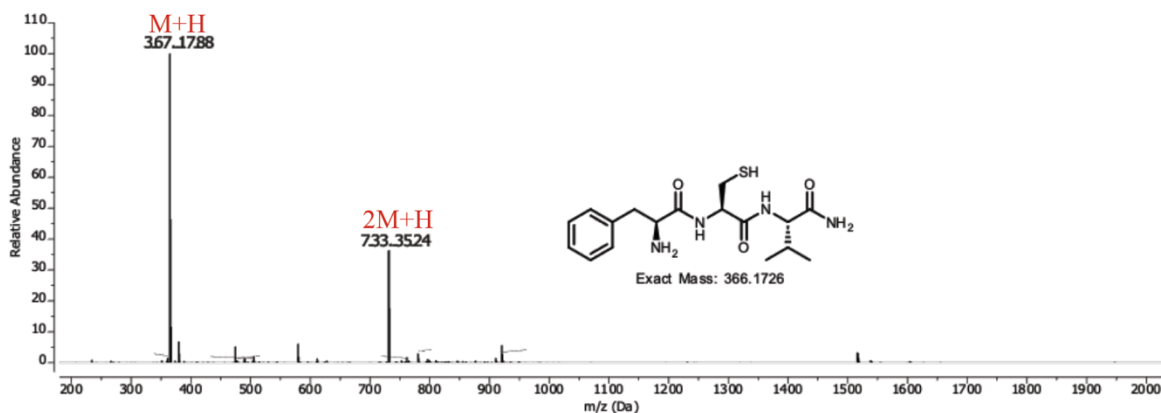


### HPLC Trace of reaction
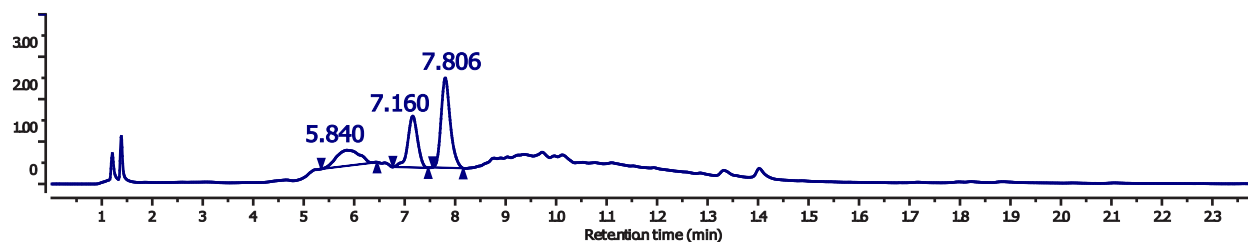
**MS Trace of FCV-sulfenic acid peptide (5.840)**
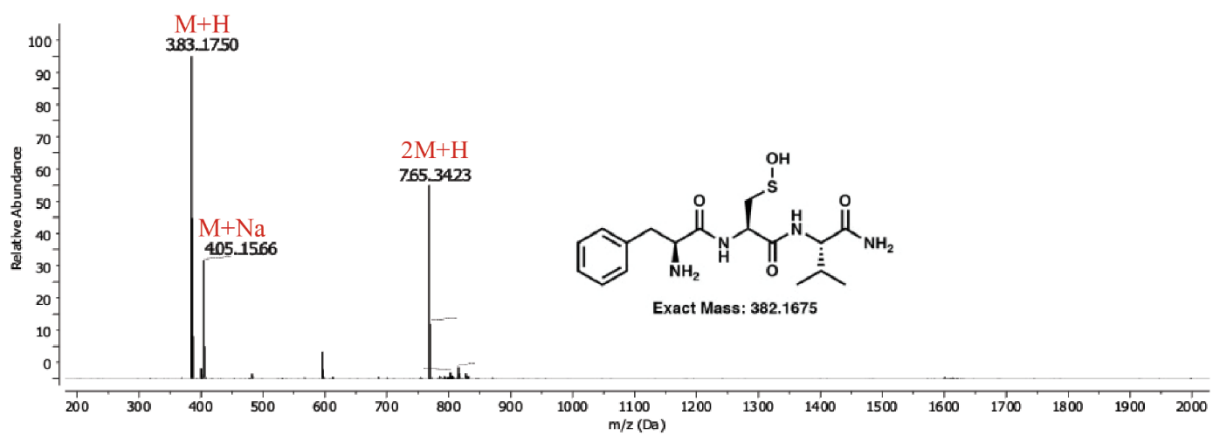


**MS Trace of FCV peptide (7.160)**
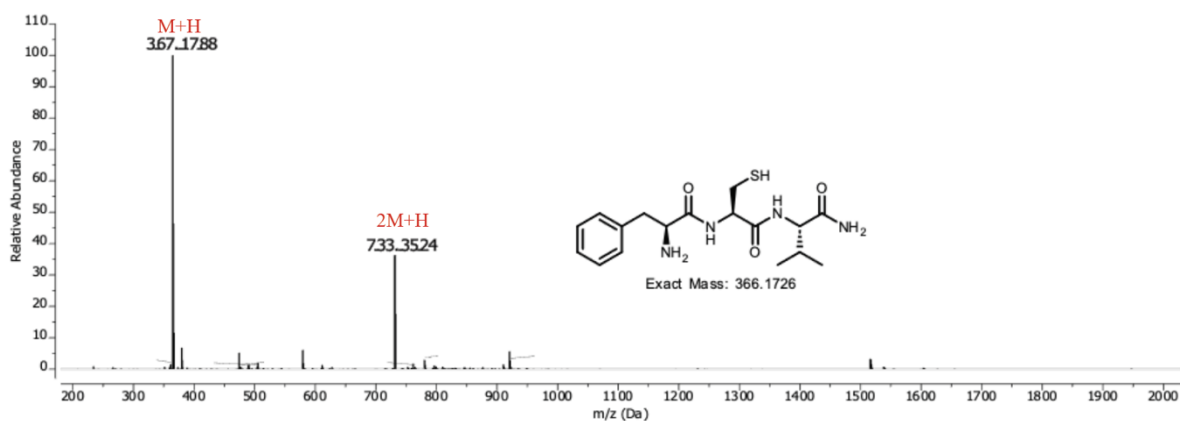


**MS Trace of FCV-disulfide peptide (7.806)**
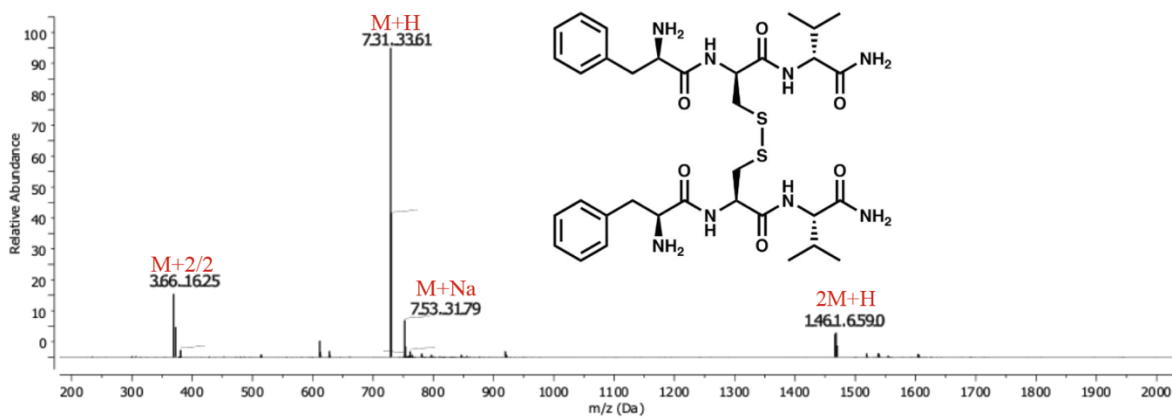
**HPLC Trace of TCEP reaction**
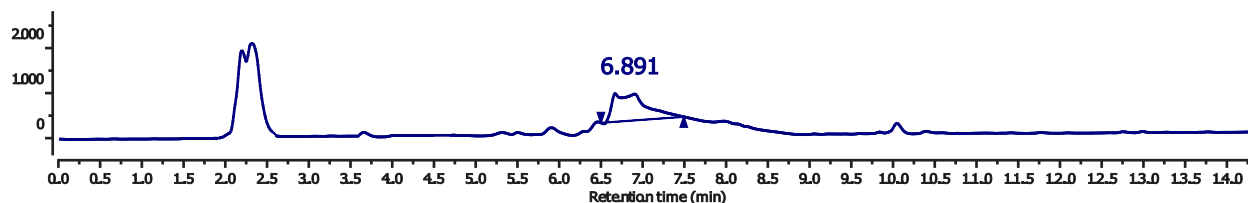


**MS Trace of FCV peptide (6.891)**
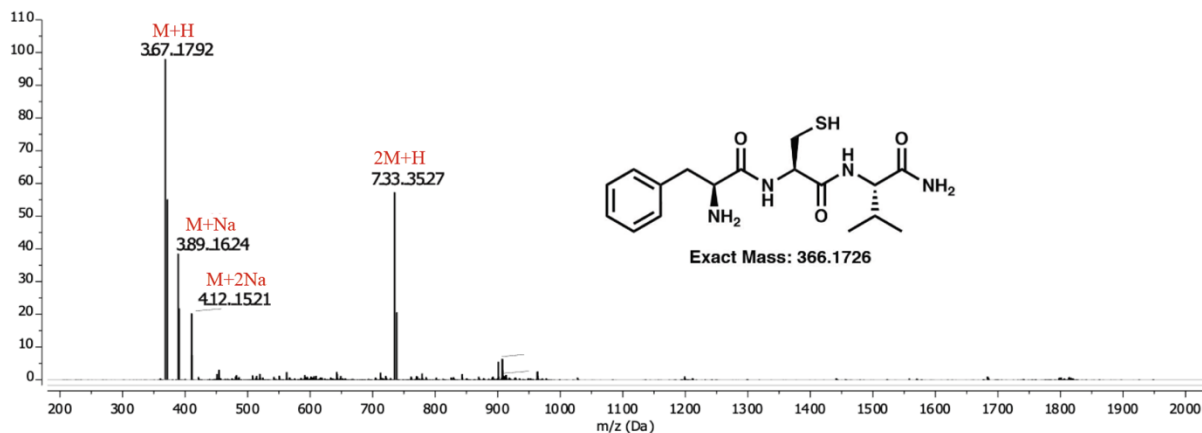


## XIV. Supplementary Figure 6. Modification of tertiary amine small molecules to aldehyde analogs.

**4-phenyl N,N-dimethyl Butylamine:**
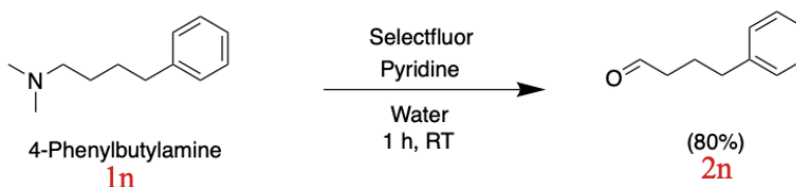


500.0 mg of tertiary amine 4-PhenylButylamine (1n) was dissolved in 3 mL of 1:1 water:ACN. Pyridine (5 eq) and selectfluor (2 eq) were then added to the reaction mixture. The reaction mixture was stirred for 1 h. The mixture was concentrated and purified by silica gel column chromatography (eluent: 5% EtOAc in Hexane) to generate the aldehyde product 4-phenyl-butanal (2n) in 80% yield. Products were characterized by NMR.

**4-phenylbutanal (2n):** $^1$**H NMR** (600 MHz, CDCl$_3$) δ 9.74 (s, 1H), 7.28 (t, $J$ = 7.7 Hz, 2H), 7.19 (d, $J$ = 7.3 Hz, 1H), 7.16 (d, $J$ = 6.9 Hz, 2H), 2.65 (t, $J$ = 7.6 Hz, 2H), 2.44 (td, $J$ = 7.4, 1.7 Hz, 2H), 1.95 (p, $J$ = 7.4 Hz, 2H).
**4-phenylbutanal (2n):** $^{13}$**C NMR** (151 MHz, CDCl$_3$) δ 202.33, 172.49, 141.20, 128.45, 126.08, 43.11, 34.98, 23.62.

# <sup>1</sup>H NMR of 4-phenylbutanal (2n)



# <sup>13</sup>C NMR of 4-phenylbutanal (2n)

**N,N-dimethyl benzylamine:**



N,N-dimethyl-1-phenylmethanamine

**1o**

Selectfluor
Pyridine
Water
1 h, RT

(88%)

**2o**

500.0 mg of tertiary amine N,N-dimethyl-1phenylmethanamine (1o) was dissolved in 3 mL of 1:1 water:ACN. Pyridine (5 eq) and selectfluor (2 eq) were then added to the reaction mixture. The reaction mixture was stirred for 1 h. The mixture was concentrated and purified by silica gel column chromatography (eluent: 5% EtOAc in Hexane) to generate the aldehyde product benzaldehyde (2o) in 88% yield. Products were characterized by NMR.

**benzaldehyde (2o): $^1$H NMR** (600 MHz, CDCl$_3$) δ 9.96 (s, 1H), 7.82 (d, $J$ = 7.0, 1.4 Hz, 2H), 7.57 (t, $J$ = 7.5, 1.3 Hz, 1H), 7.47 (t, $J$ = 7.8 Hz, 2H).
**benzaldehyde (2o): $^{13}$C NMR** (151 MHz, CDCl$_3$) δ 192.28, 136.38, 134.39, 129.67, 128.94.

**$^1$H NMR of benzaldehyde (2o)**

## <sup>13</sup>C NMR of benzaldehyde (2o)



**Vanillin:**
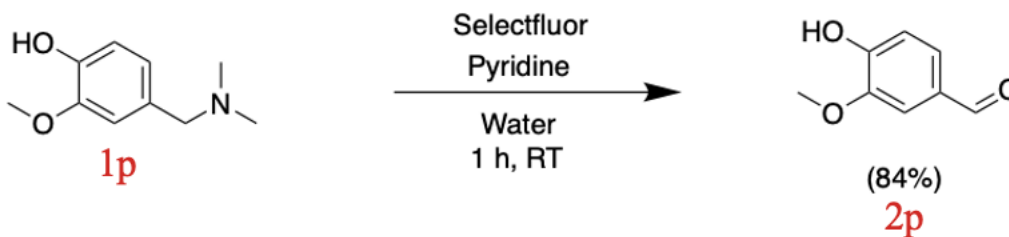


500.0 mg of tertiary amine Vanillin (1p) was dissolved in 3 mL of 1:1 water:ACN. Pyridine (5 eq) and selectfluor (2 eq) were then added to the reaction mixture. The reaction mixture was stirred for 1 h. The mixture was concentrated and purified by silica gel column chromatography (eluent: 10% EtOAc in Hexane) to generate the aldehyde product 4-hydroxy-3-methoxybenzaldehyde (2p) in 84% yield. Products were characterized by NMR.

**Vanillin aldehyde analog (2p):** **<sup>1</sup>H NMR** (600 MHz, CDCl₃) δ 9.79 (s, 1H), 7.42 – 7.34 (m, 2H), 7.06 – 6.94 (m, 1H), 6.27 (s, 1H), 3.95 (s, 3H).
**Vanillin aldehyde analog (2p):** **<sup>13</sup>C NMR** (151 MHz, CD₃OD) δ 191.43, 153.24, 148.22, 129.22, 126.47, 114.86, 110.39, 54.93.

# $^1$H NMR of vanillin aldehyde analog (2p)
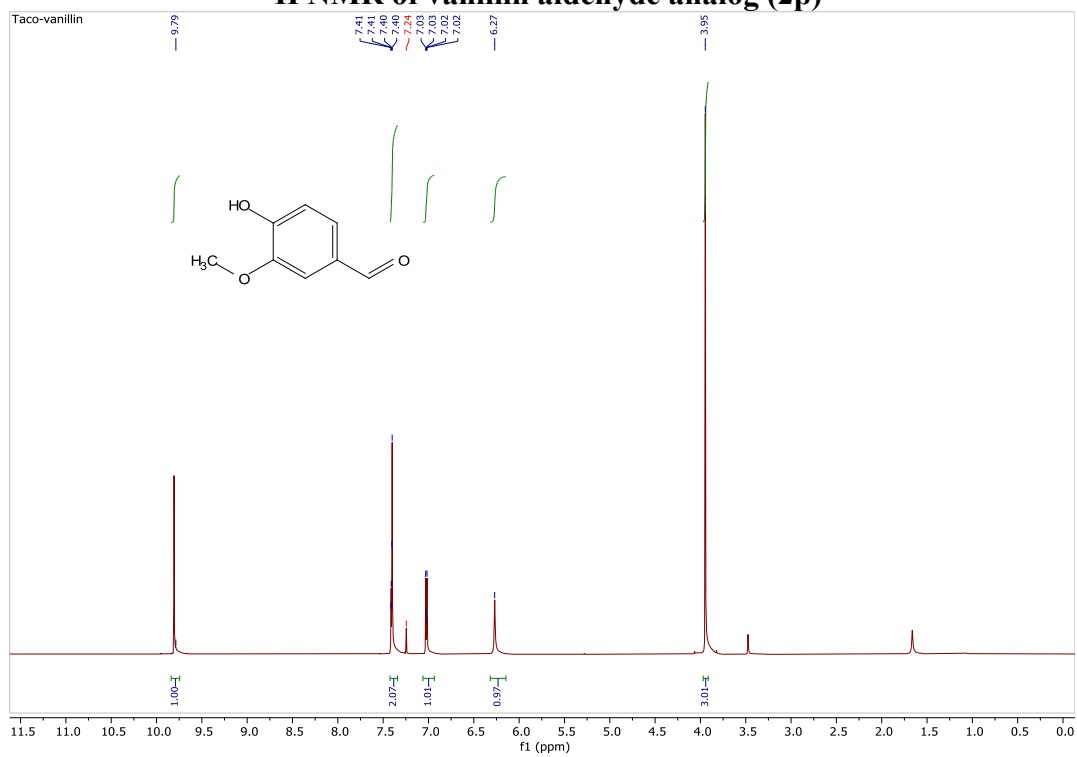


# $^{13}$C NMR of vanillin aldehyde analog (2p)

**Citalopram hydrogen bromide:**



Citalopram Hydrobromide
**1q**

(93%)
**2q**
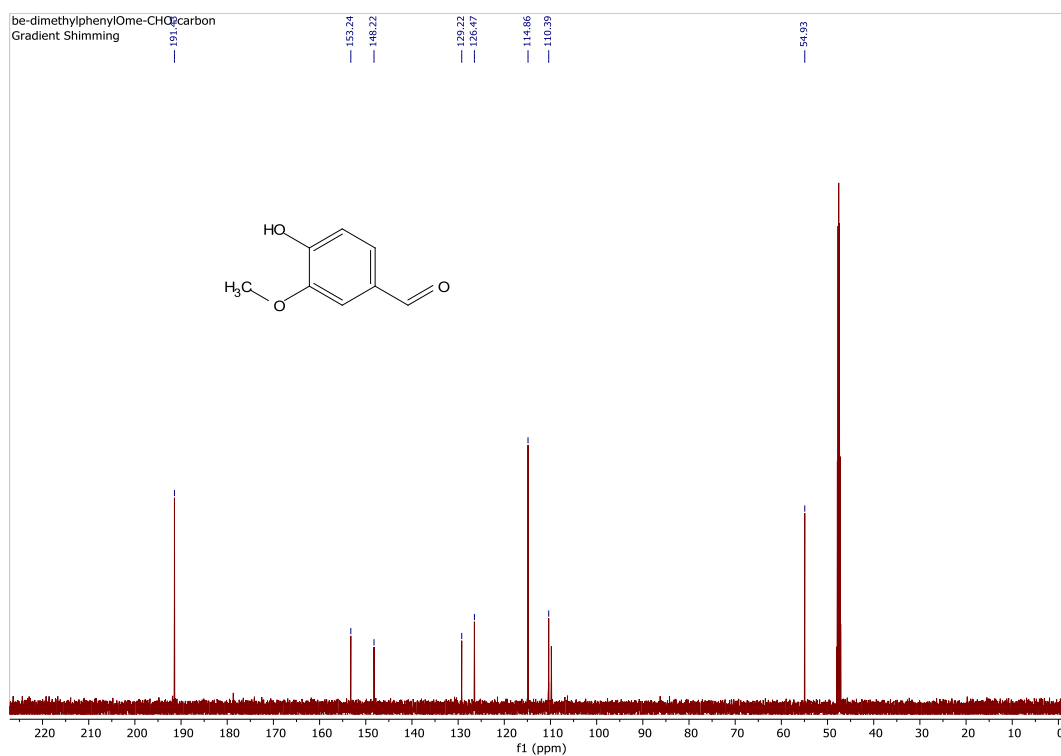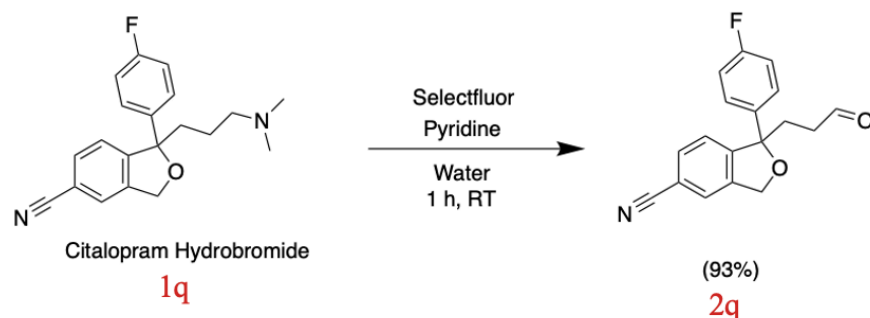
500.0 mg of tertiary amine citalopram hydrobromide (1q) was dissolved in 3 mL of 1:1 water:ACN. Pyridine (5 eq) and selectfluor (2 eq) were then added to the reaction mixture. The reaction mixture was stirred for 1 h. The mixture was concentrated and purified by silica gel column chromatography (eluent: 10% EtOAc in Hexane) to generate the aldehyde product 1-(4-fluorophenyl)-2-(3-oxopropyl)-1,3-dihydroisobenzofuran-5-carbonnitrile (2q) in 93% yield. Products were characterized by NMR.

**Citalopram aldehyde analog (2q):** **$^1$H NMR** (400 MHz, CDCl$_3$) δ 8.84 (s, 1H), 7.63 (dd, $J$ = 8.0, 1.4 Hz, 1H), 7.52 (d, $J$ = 8.0 Hz, 1H), 7.50 – 7.45 (m, 1H), 7.44 – 7.37 (m, 2H), 7.10 – 7.01 (m, 2H), 5.04 (q, $J$ = 13.0 Hz, 2H), 3.71 (d, $J$ = 1.2 Hz, 2H), 1.25 (s, 1H).

**Citalopram aldehyde analog (2q):** **$^{13}$C NMR** (101 MHz, CDCl$_3$) δ 181.91, 163.63, 161.17, 146.69, 139.42, 139.38, 131.82, 126.09, 126.01, 125.43, 123.53, 118.29, 116.03, 115.81, 112.58, 90.14, 71.44, 65.61, 55.09.

# <sup>1</sup>H NMR of citalopram aldehyde analog (2q)
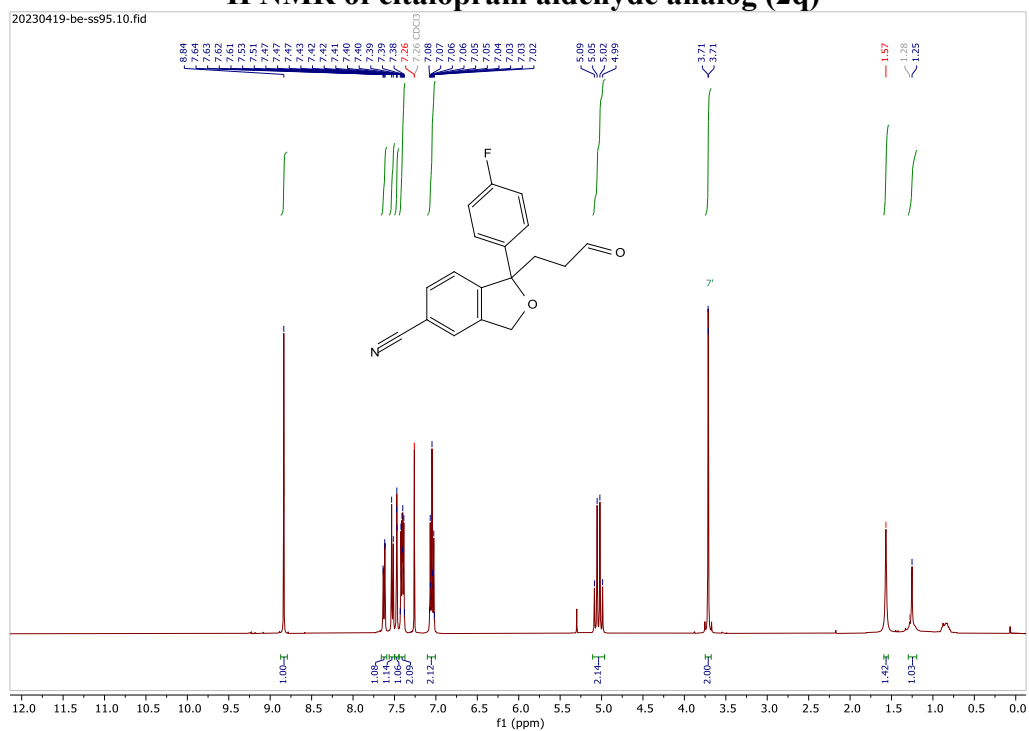


# <sup>13</sup>C NMR of citalopram aldehyde analog (2q)
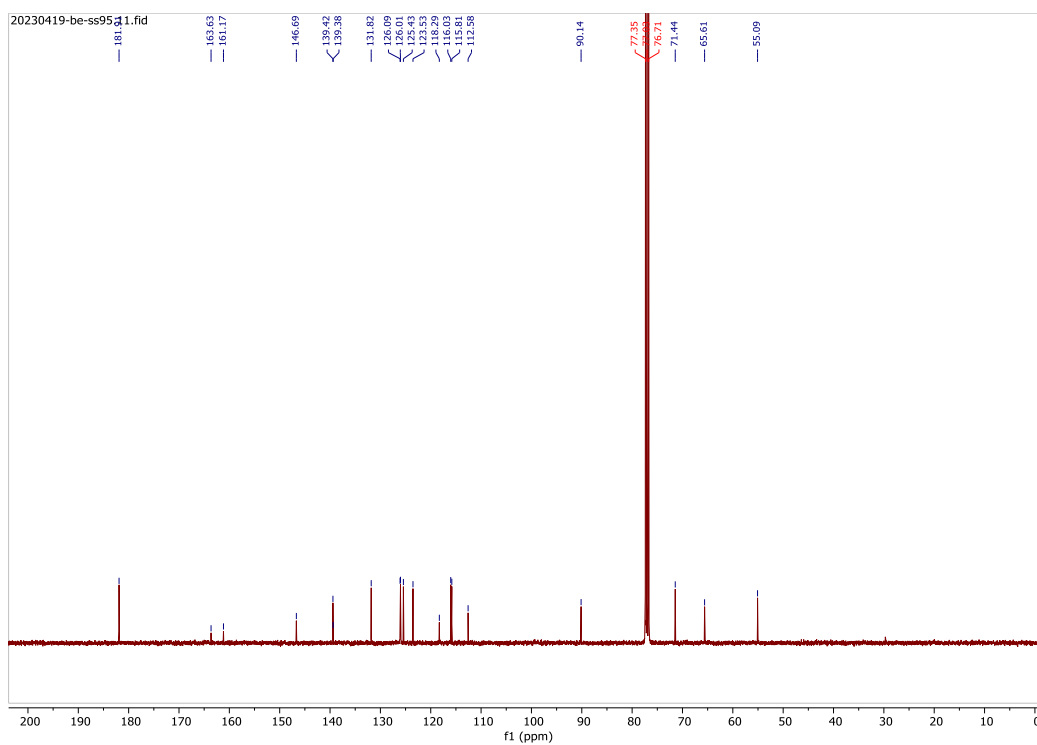
## XV. Supplementary Figure 7. Diversification of allysine containing peptides.

**Oxime Chemistry:**
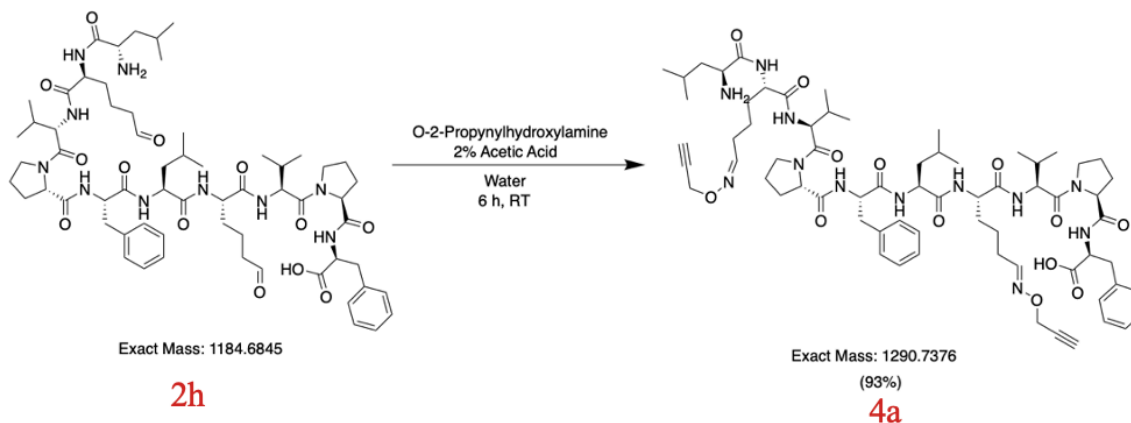


To 1.0 mg of peptide aldehyde LK(CHO)VPFLK(CHO)VPF (2h) dissolved in 400 µL of water, was added 5 eq of *o*-2-propynylhydroxylamine. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to oxime peptide was 93%.
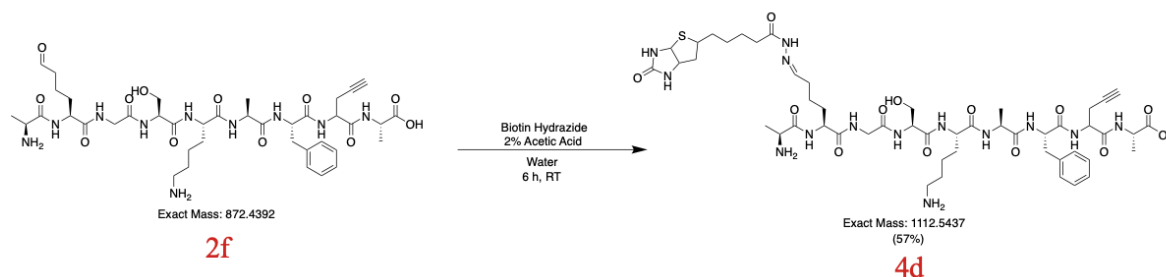
**LK(CHO)VPFLK(CHO)VPF peptide 2h (oxime reaction)**. LCMS: m/z 1185.6948 (calcd $[M+H]^+$ = 1185.6918), m/z 1221.6958 (calcd $[M+H+2H_2O]^+$ =1221.6948), m/z 1257.6935 (calcd $[M+H+4H_2O]^+$ = 1257.6935. (HPLC analysis at 220 nm). Retention time in HPLC: 14.816.

**LK(CHO)VDFLK(CHO)VPF Oxime peptide 4a (oxime reaction)**. LCMS: m/z 1327.7236 (calcd $[M+H+2H_2O]^+$ = 1327.7449). (HPLC analysis at 220 nm). Retention time in HPLC: 21.269.

### HPLC Trace of oxime forming reaction
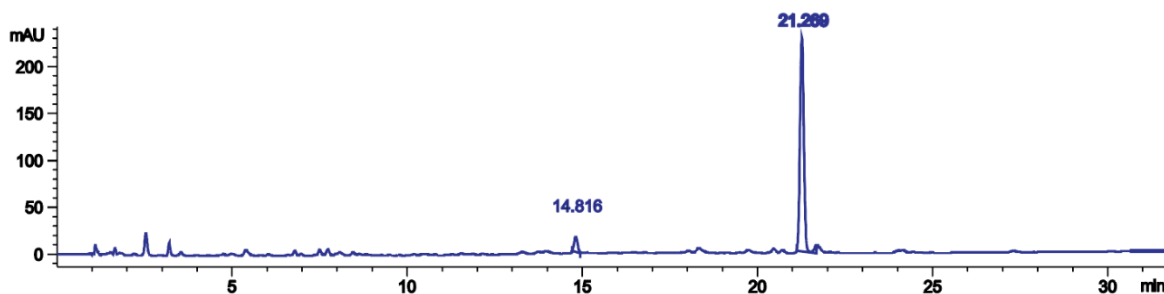
## MS-Trace of peak 14.816
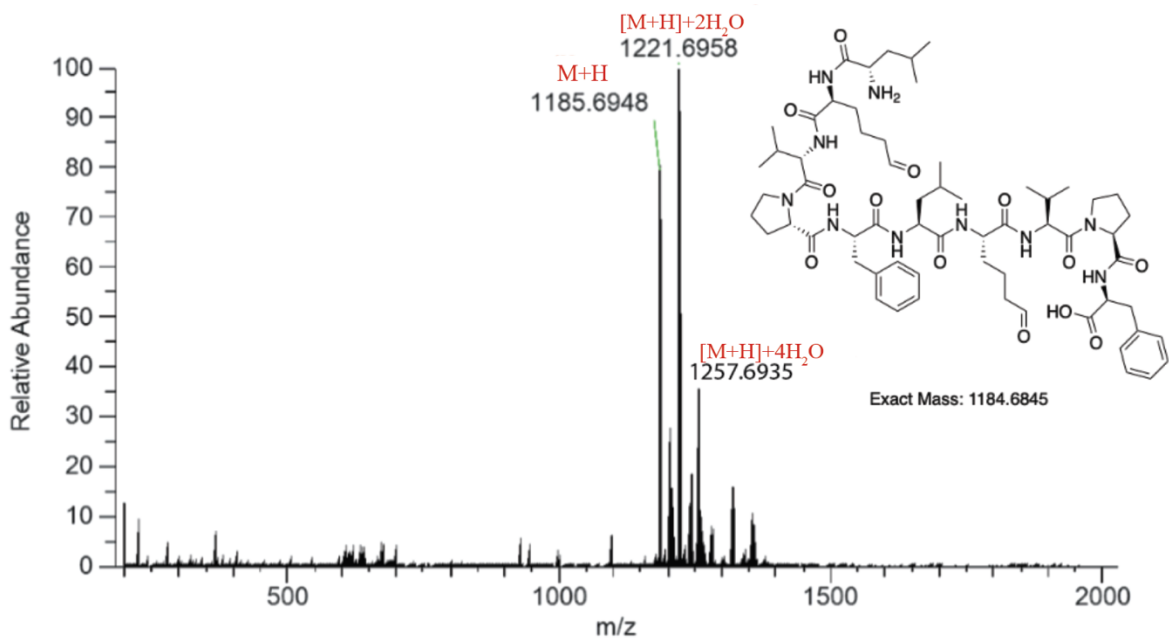


## MS-Trace of peak 21.269
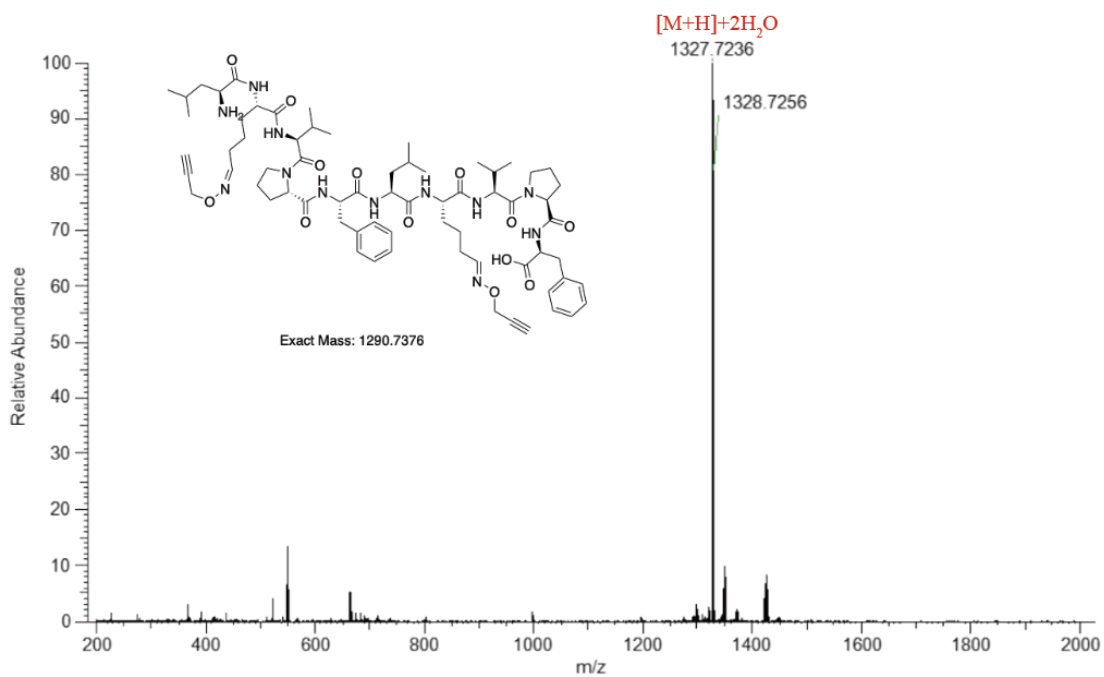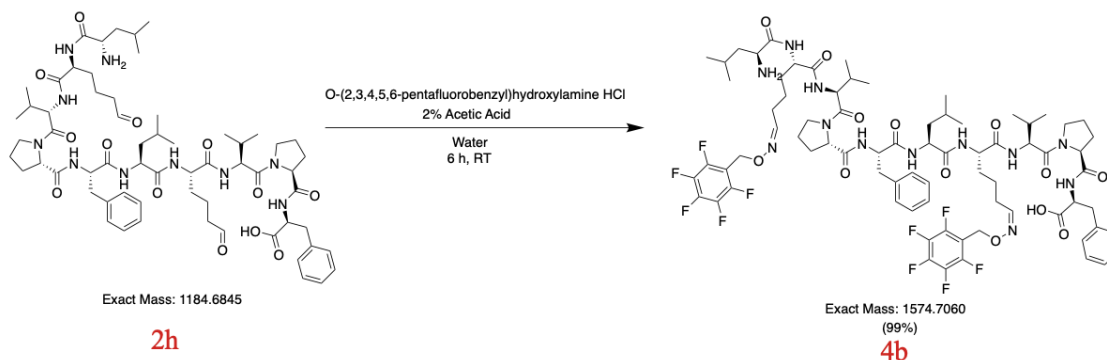
Exact Mass: 1184.6845

2h

Exact Mass: 1574.7060
(99%)

4b

To 1.0 mg of peptide aldehyde LK(CHO)VPFLK(CHO)VPF (2h) dissolved in 400 µL of water, was added 5 eq of O-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine hydrochloride. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.  % conversion to oxime peptide was 99%.

**LK(CHO)VDFLK(CHO)VPF Oxime peptide 4b**. LCMS: m/z 1611.6866 (calcd [M+H+2H$_2$O]$^+$ = 1611.7133), m/z 1633.6697 (calcd [M+H+2H$_2$O+Na]$^+$ = 1633.6933), (HPLC analysis at 220 nm). Retention time in HPLC: 13.494.

## HPLC Trace of Reaction
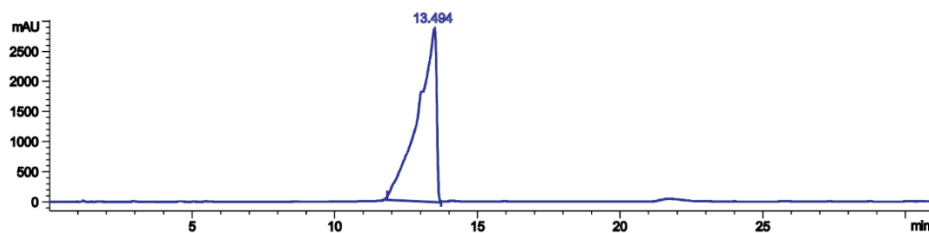
**MS-Trace of peak 13.494**
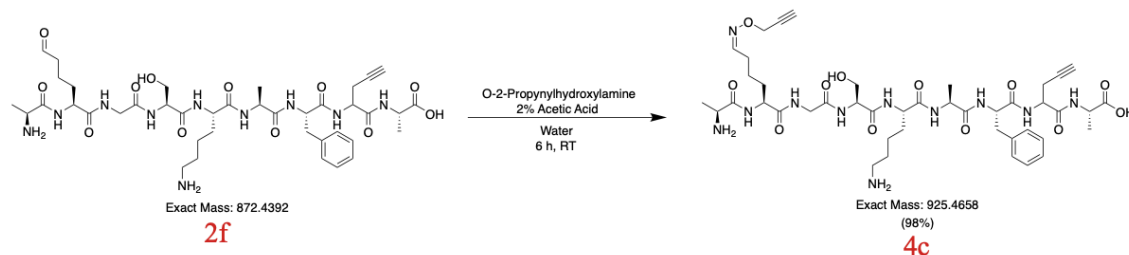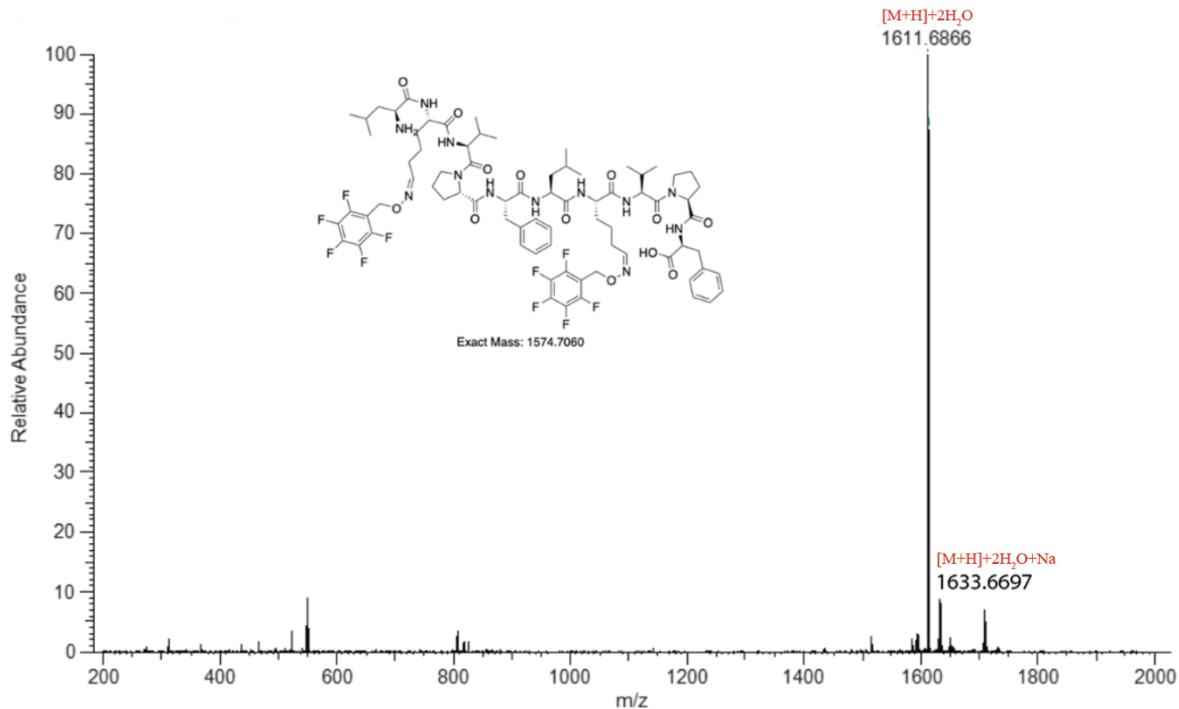




To 1.0 mg of peptide aldehyde AK(CHO)GSKAF(Pra)A (2f) dissolved in 400 µL of water, was added 5 eq of *o*-2-propynylhydroxylamine hydrochloride. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to oxime peptide was 98%.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4474 (calcd [M+H]$^+$ = 873.4465), m/z 437.2269 (calcd [M+2/2]$^+$ = 437.2232). (HPLC analysis at 220 nm). Retention time in HPLC: 5.145.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4474 (calcd [M+H]$^+$ = 873.4465), m/z 895.4271 (calcd [M+Na]$^+$ = 895.4192). (HPLC analysis at 220 nm). Retention time in HPLC: 5.820.

**AK(CHO)GSKAF(Pra)A Oxime peptide 4c.** LCMS: m/z 926.4707(calcd [M+H]$^+$ = 926.4730), m/z 948.4500 (calcd [M+Na]$^+$ = 948.4458). (HPLC analysis at 220 nm). Retention time in HPLC: 8.393.

### HPLC trace of AK(CHO)GSKAF(Pra)A [2f] aldehyde forming reaction



### MS-Trace of aldehyde containing peptide 2f (peak 5.145)

**HPLC Trace of Reaction**



**MS-Trace of peak 5.820**

**MS-Trace of peak 8.393**
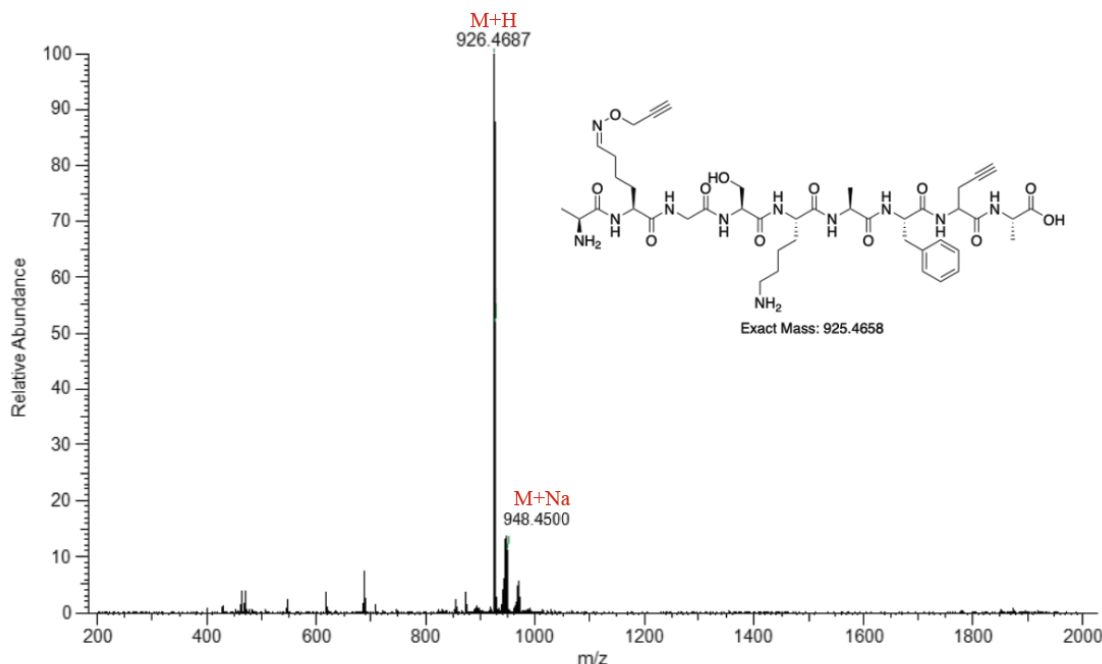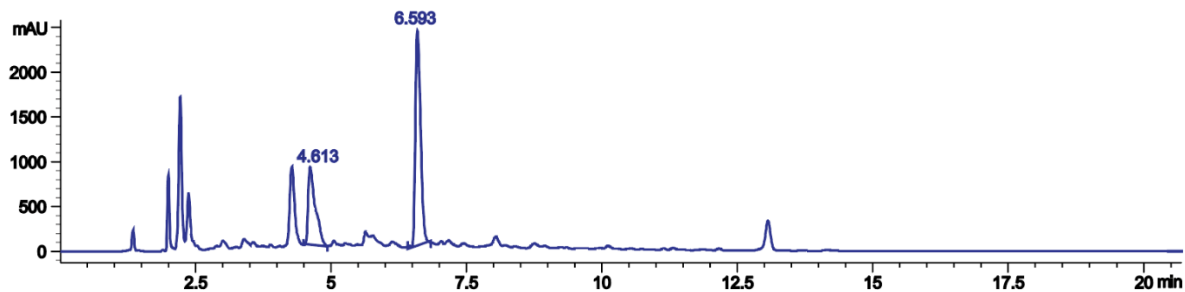


## Hydrazone Chemistry



To 1.0 mg of peptide aldehyde AK(CHO)GSKAF(Pra)A (2f) dissolved in 400 µL of water, was added 5 eq of Biotin Hydrazide. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.  % conversion to hydrazone peptide was 57%.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4474 (calcd [M+H]$^+$ = 873.4465), m/z 437.2264 (calcd [M+2/2]$^+$ = 437.2232), m/z 895.4278 (calcd [M+Na]$^+$ = 895.4192). (HPLC analysis at 220 nm). Retention time in HPLC: 3.940.

**AK(CHO)GSKAF(Pra)A Hydrazone peptide 4d**. LCMS: m/z 1113.5487 (calcd [M+H]$^+$ = 1113.5510), m/z 557.2775  (calcd [M+2/2]$^+$ = 557.2755). (HPLC analysis at 220 nm). Retention time in HPLC: 7.864.

# HPLC Trace of Reaction
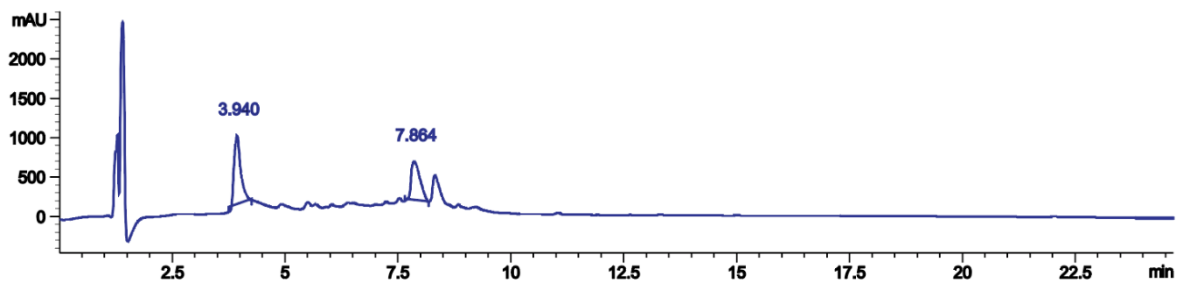


# MS-Trace of peak 3.940
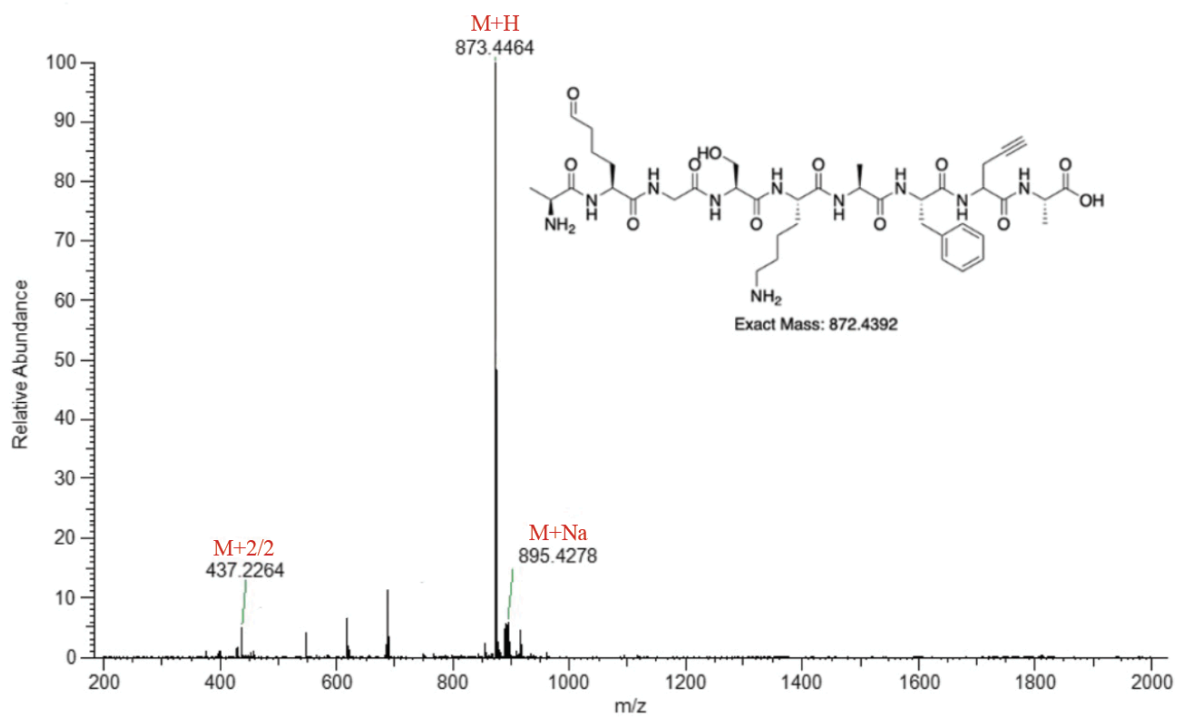
**MS-Trace of peak 7.864**
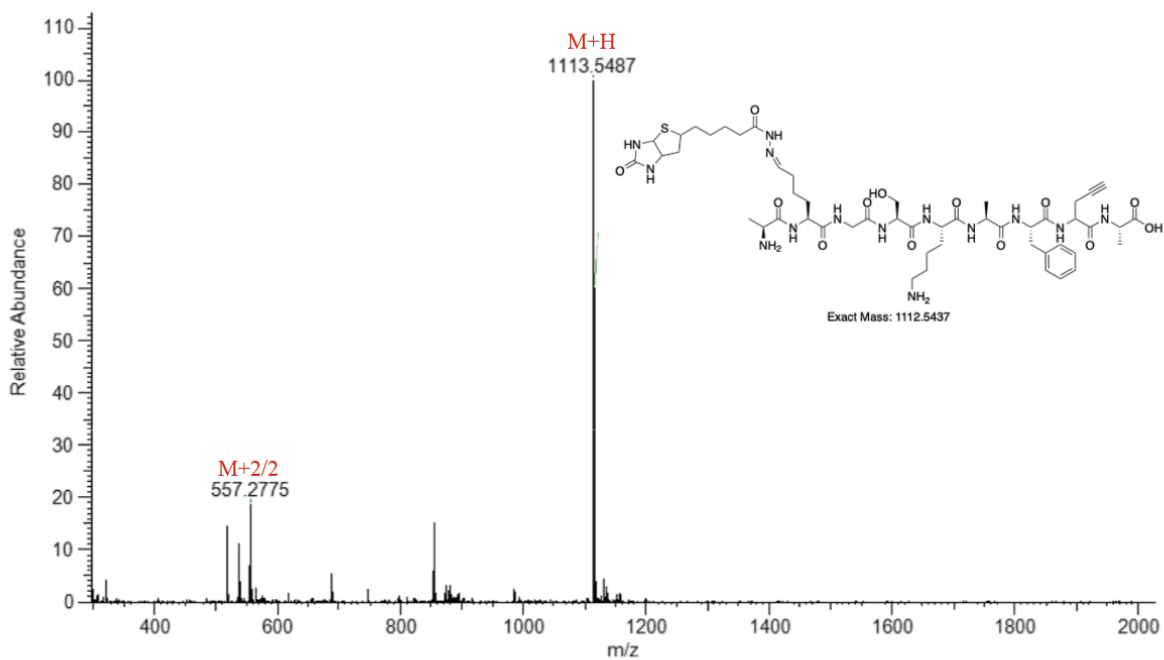


## Sulfonylhydrazone Modification:



2f → 4e
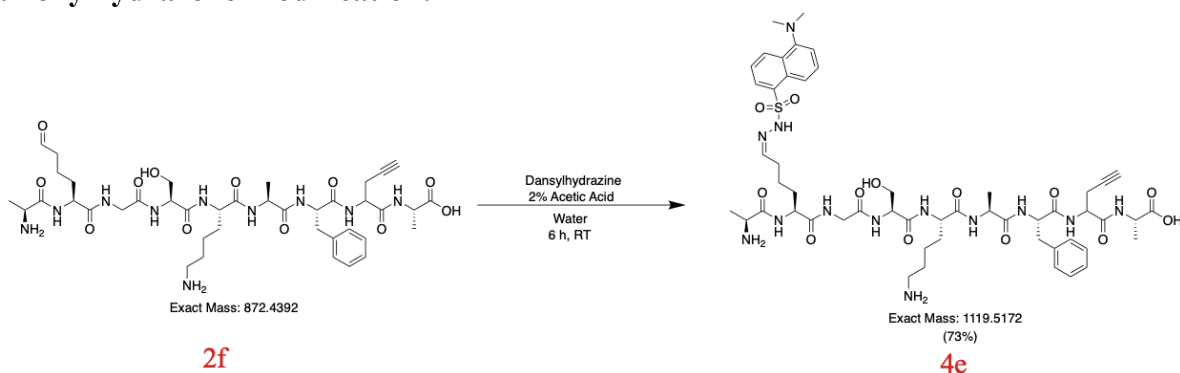
To 1.0 mg of peptide aldehyde AK(CHO)GSKAF(Pra)A (2f) dissolved in 400 µL of water, was added 5 eq of dansylhydrazine. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to sulfonylhydrazone peptide was 73%.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4465 (calcd $[M+H]^+$ = 873.4465), m/z 437.2265 (calcd $[M+2/2]^+$ = 437.2232), m/z 895.4279 (calcd $[M+H]^+$ = 895.4192). (HPLC analysis at 220 nm). Retention time in HPLC: 4.632.

**AK(CHO)GSKAF(Pra)A Sulfonylhydrazone peptide 4e**. LCMS: m/z 1120.5242 (calcd [M+$H_2O$+2/2]$^+$ = 569.7622). (HPLC analysis at 220 nm). Retention time in HPLC: 12.613.

## HPLC Trace of Reaction
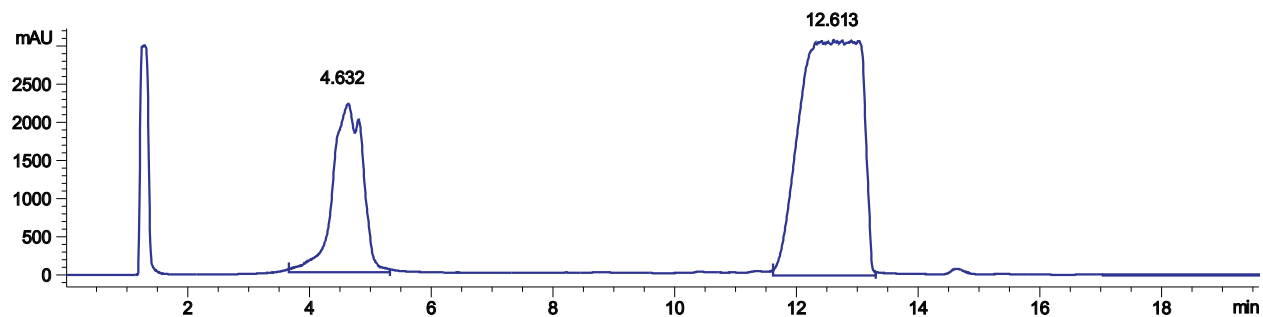


## MS-Trace of peak 4.632
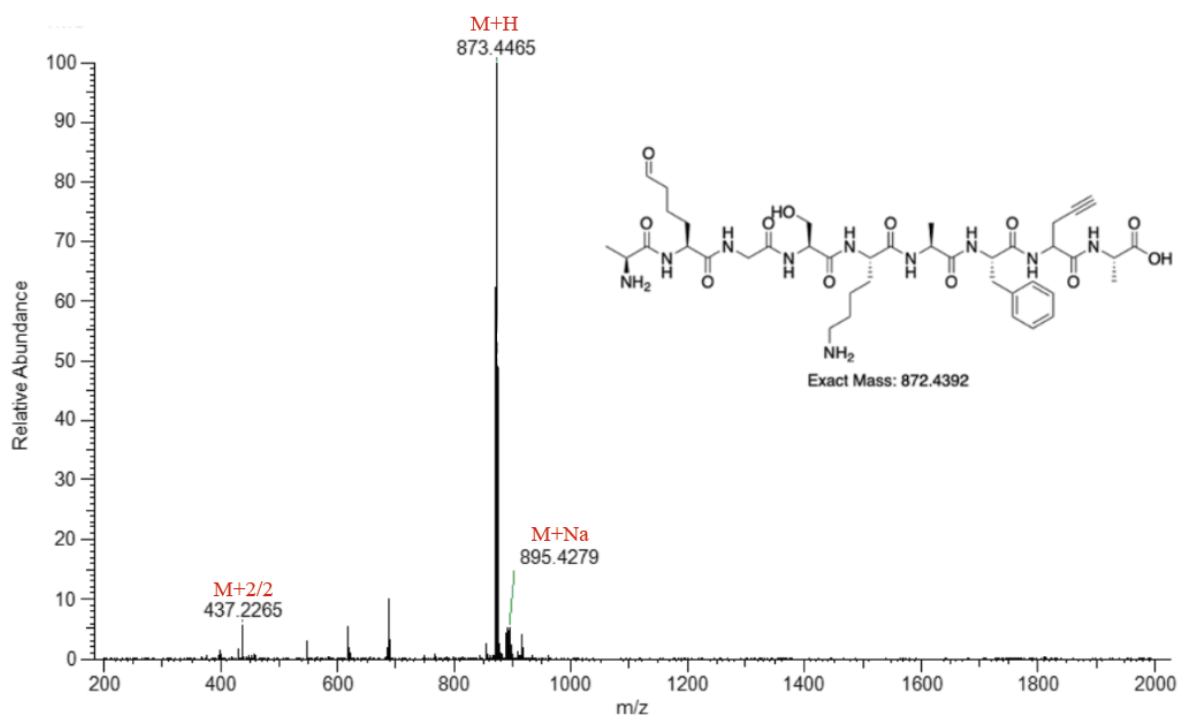
**MS-Trace of peak 12.613**
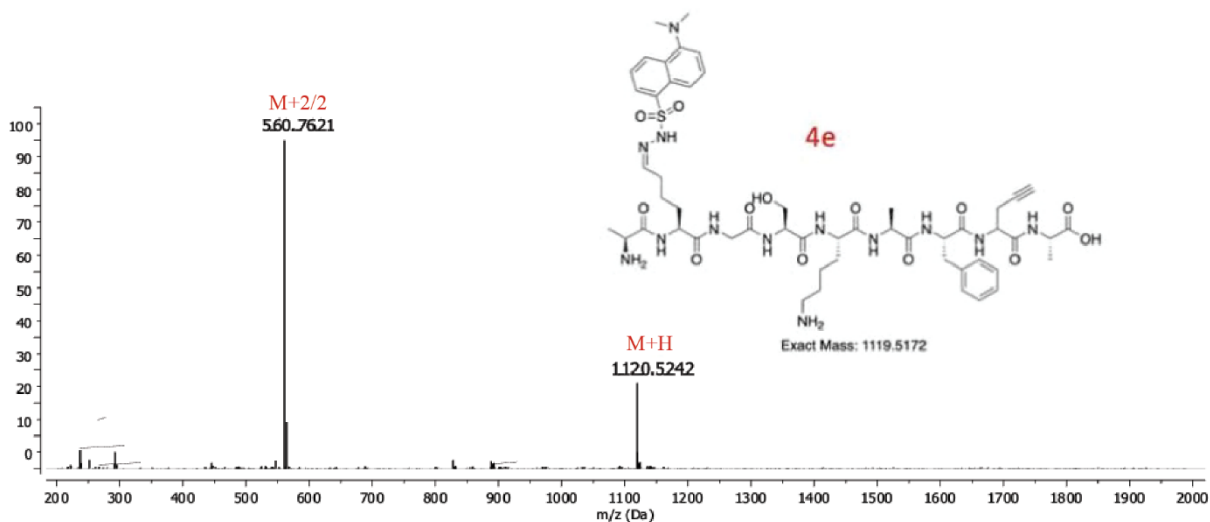


**Propargyl Amine Modification:**
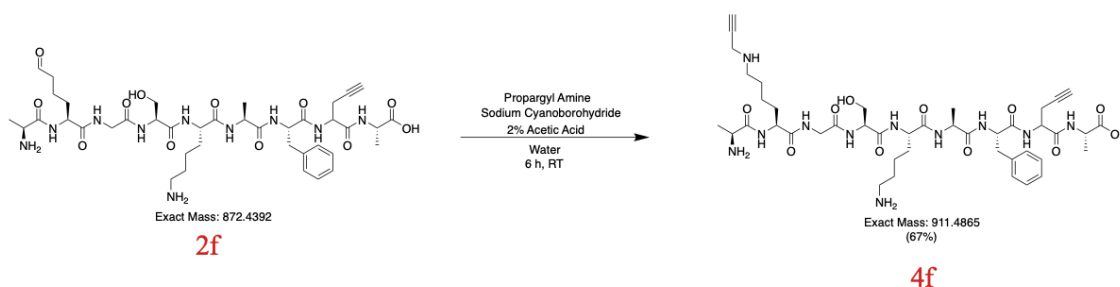


To 1.0 mg of peptide aldehyde AK(CHO)GSKAF(Pra)A (2f) dissolved in 400 µL of water, was added 5 eq of propargyl amine. The reaction mixture was stirred for 1 h, followed by the addition of 10 eq of sodium cyanoborohydride. Reaction was then stirred for 5 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to propargyl peptide was 67%.
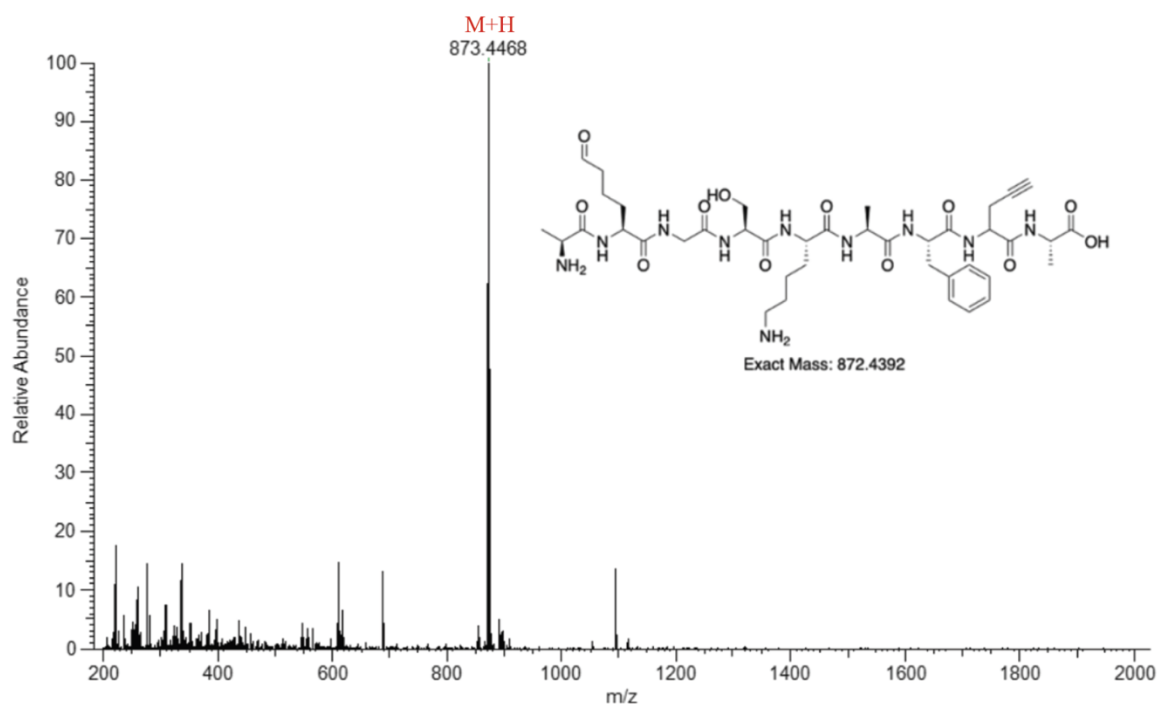
**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4463 (calcd [M+H]$^+$ = 873.4465), m/z 437.2265 (calcd [M+2/2]$^+$ = 437.2232), . (HPLC analysis at 220 nm). Retention time in HPLC: 4.613.

**AK(CHO)GSKAF(Pra)A Propargyl peptide 4f**. LCMS: m/z 912.4942 (calcd [M+H]$^+$ = 912.4940), m/z 456.7506 (calcd [M+2/2]$^+$ = 456.5332), m/z 895.4278 (calcd [M+Na]$^+$ =

895.4192), m/z 917.4094 (calcd [M+2Na]$^+$ = 917.3992). (HPLC analysis at 220 nm). Retention time in HPLC: 6.593.

**HPLC Trace of Reaction**



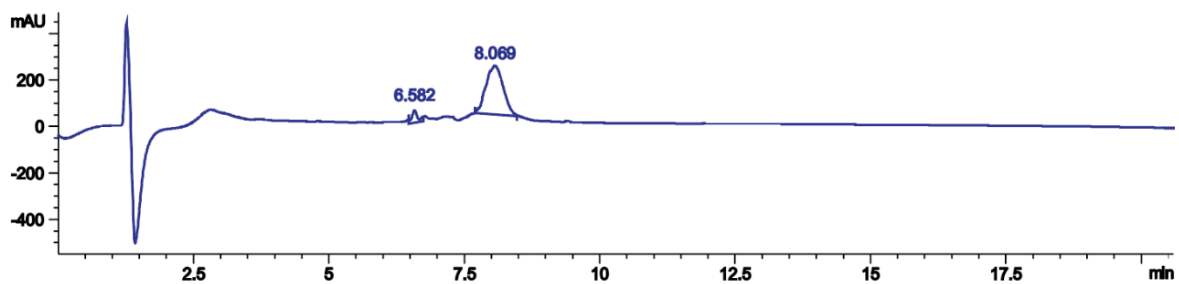**MS-Trace of peak 4.613**
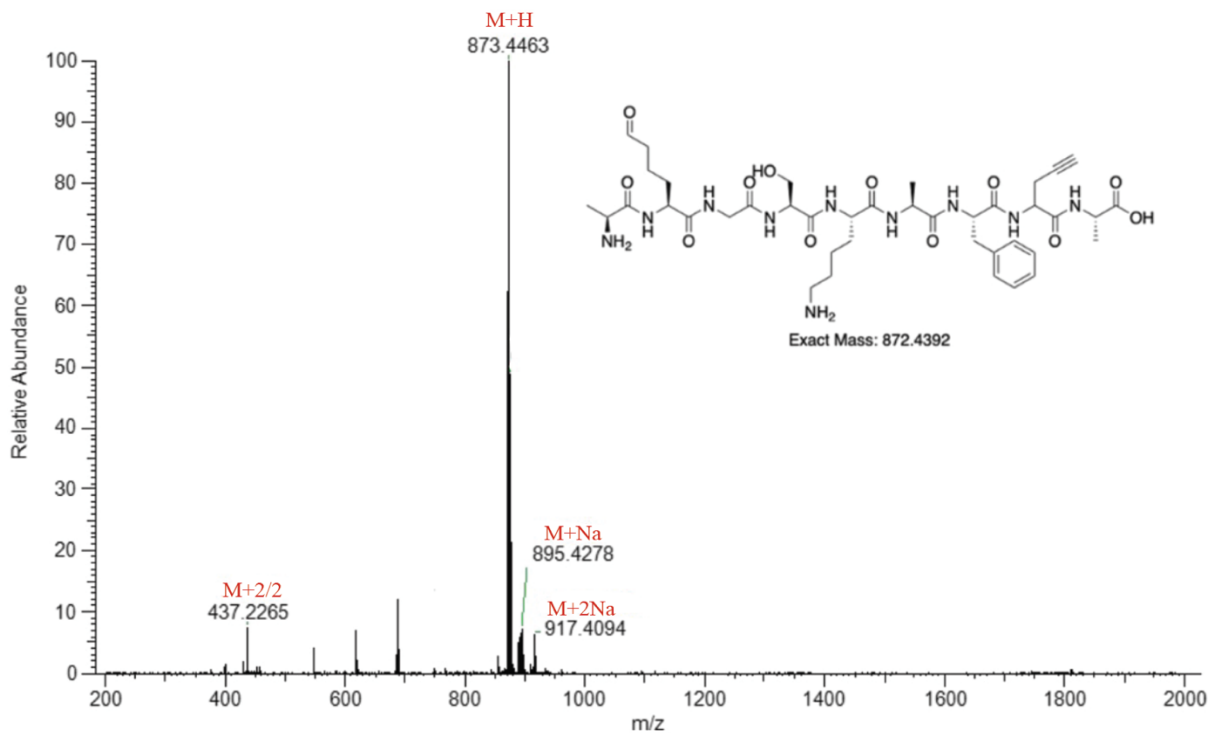
**MS-Trace of peak 6.593**
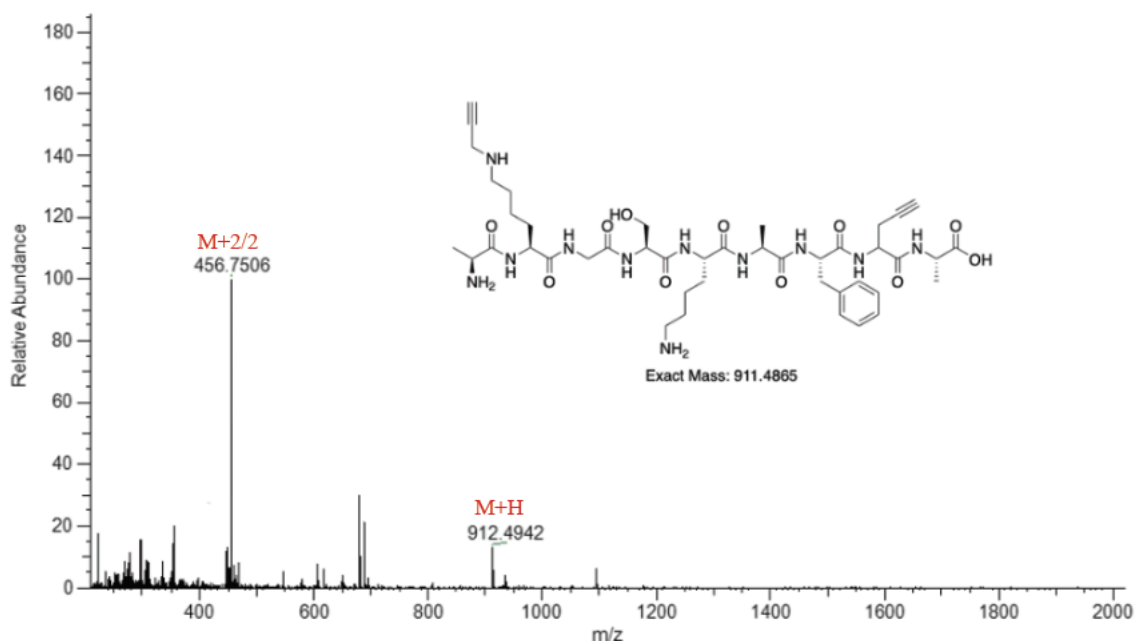


**Thiazolidine Chemistry**



2f → 4g
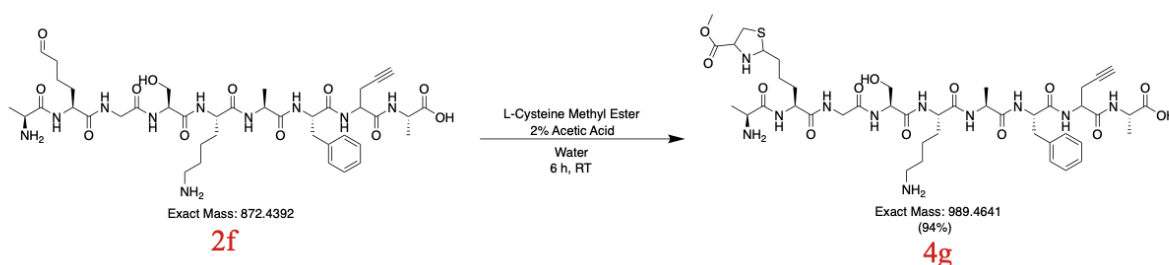
To 1.0 mg of peptide aldehyde AK(CHO)GSKAF(Pra)A (2f) dissolved in 400 µL of water, was added 2 eq of cysteine methyl ester. The reaction was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to thiazolidine peptide was 94%.

**AK(CHO)GSKAF(Pra)A peptide 2f**. LCMS: m/z 873.4468 (calcd $[M+H]^+ = 873.4465$). (HPLC analysis at 220 nm). Retention time in HPLC: 5.370.

**AK(CHO)GSKAF(Pra)A Thiazolidine peptide 4g**. LCMS: m/z 990.4692(calcd $[M+H]^+ = 990.4713$). (HPLC analysis at 220 nm). Retention time in HPLC: 6.443.

## HPLC Trace of Reaction
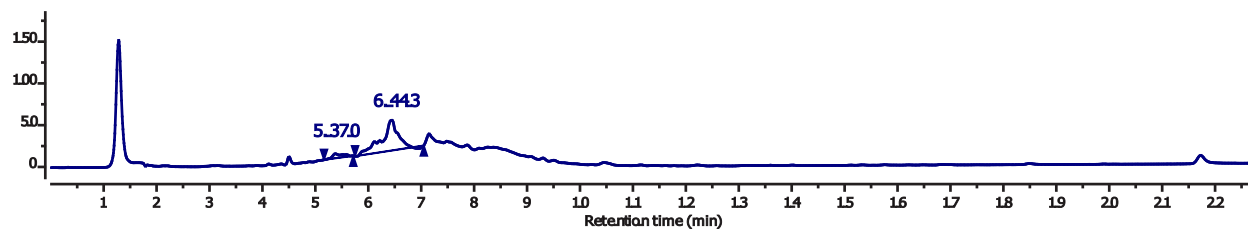


## MS-Trace of peak 5.392

**MS-Trace of peak 6.605**



M+H
990.4689

Exact Mass: 989.4641



Exact Mass: 742.3974
**2d**

L-Cysteine Methyl Ester
Water
6 h, RT

Exact Mass: 859.4222
94%
**4h**
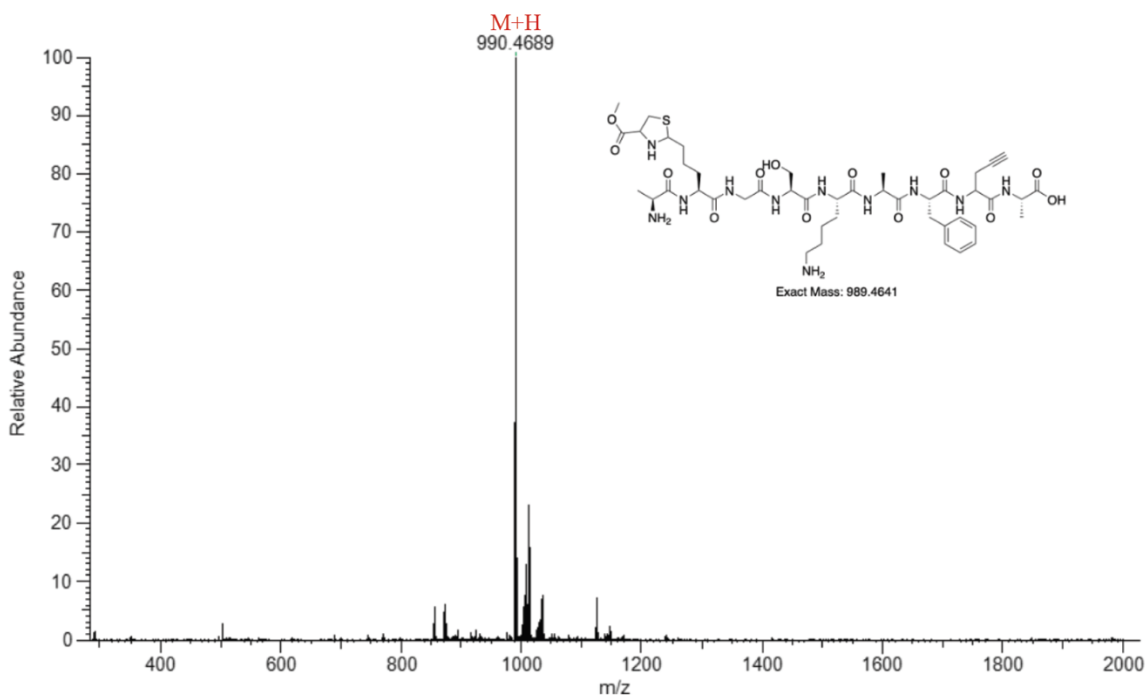
1.0 mg of peptide aldehyde LDKme₂VNR (2d) was dissolved in 200 μL of water and added to L-Cysteine Methyl Ester (2 eq.). The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to hydroxylamine peptide was 94%.

**LDK(CHO)VNR peptide 2d**. LCMS: m/z 743.4043 (calcd $[M+H]^+$ = 743.3974), 372.2060 (calcd $[M+2/2]^+$ = 372.1987), 395.2281 (calcd $[M+2/2+Na]^+$ = 395.1787), (HPLC analysis at 220 nm). Retention time in HPLC: 6.582.

**LDKme₂VNR Thiazolidine peptide 4h**. LCMS: m/z 878.4205 (calcd $[M+H+H_2O]^+$ = 878.4295), m/z 439.7134 (calcd $[M+H+2H_2O+2/2]^+$ = 439.2147). (HPLC analysis at 220 nm). Retention time in HPLC: 8.069.

# HPLC Trace of Reaction



# MS-Trace of peak 6.582
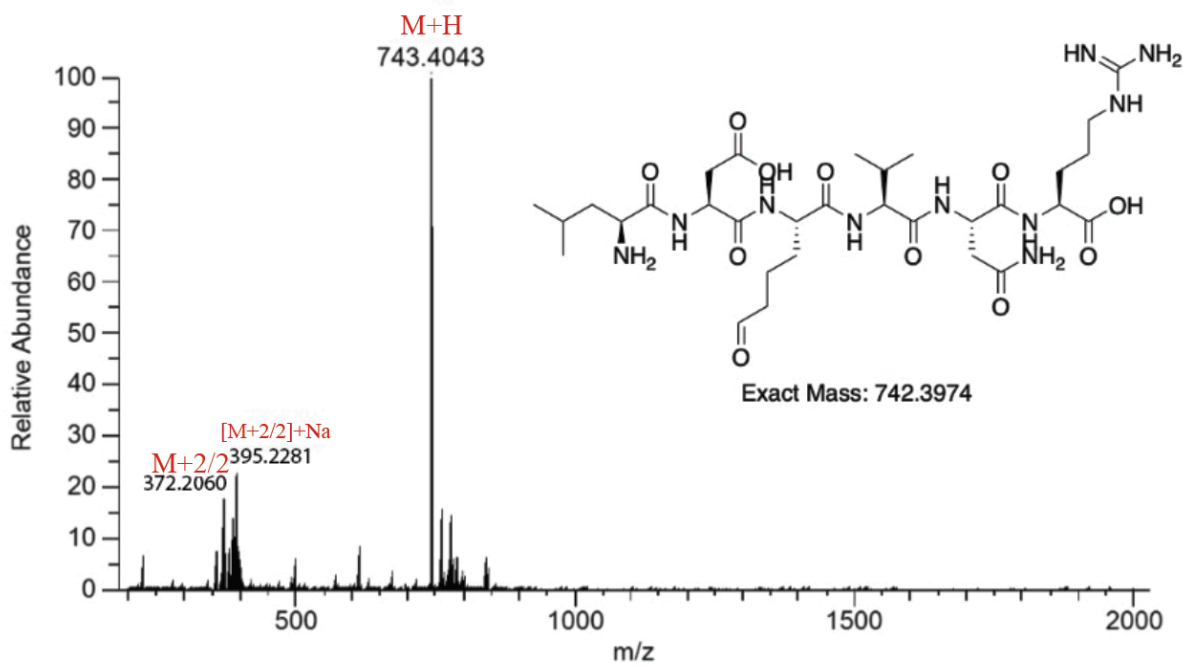
**MS-Trace of peak 8.069**
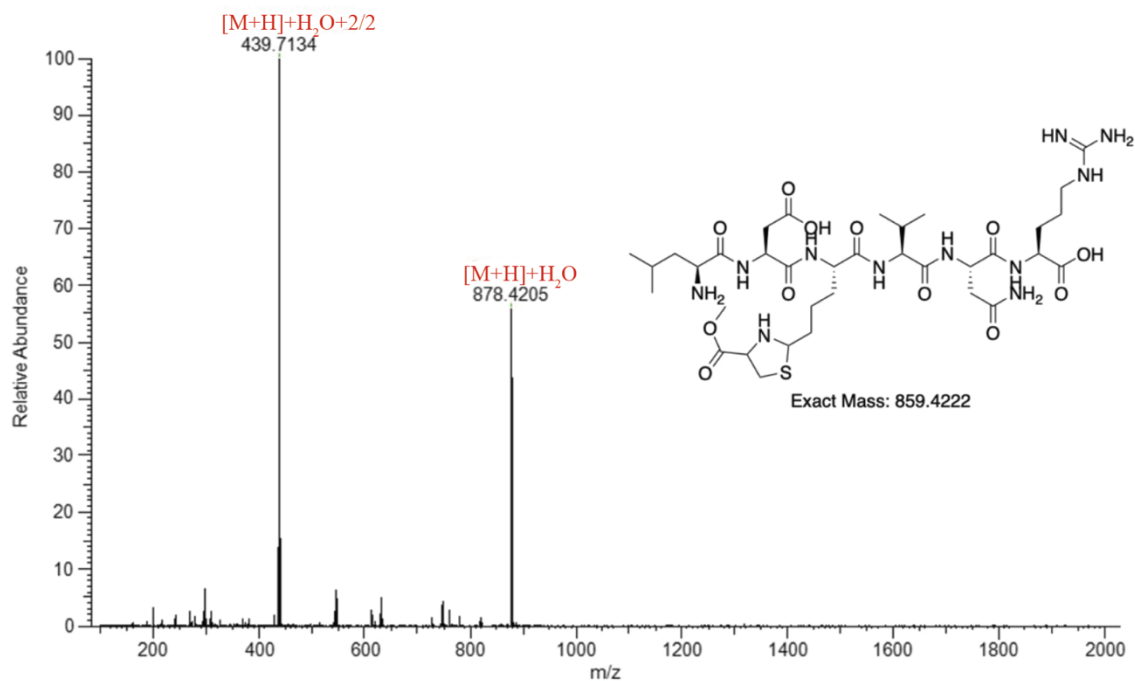


## XVI. Supplementary figure 8: Solid-support mediated generation of allysine peptides

**Optimization of solid support allysine generation using FKme₂HV**
**FKme₂HV:**
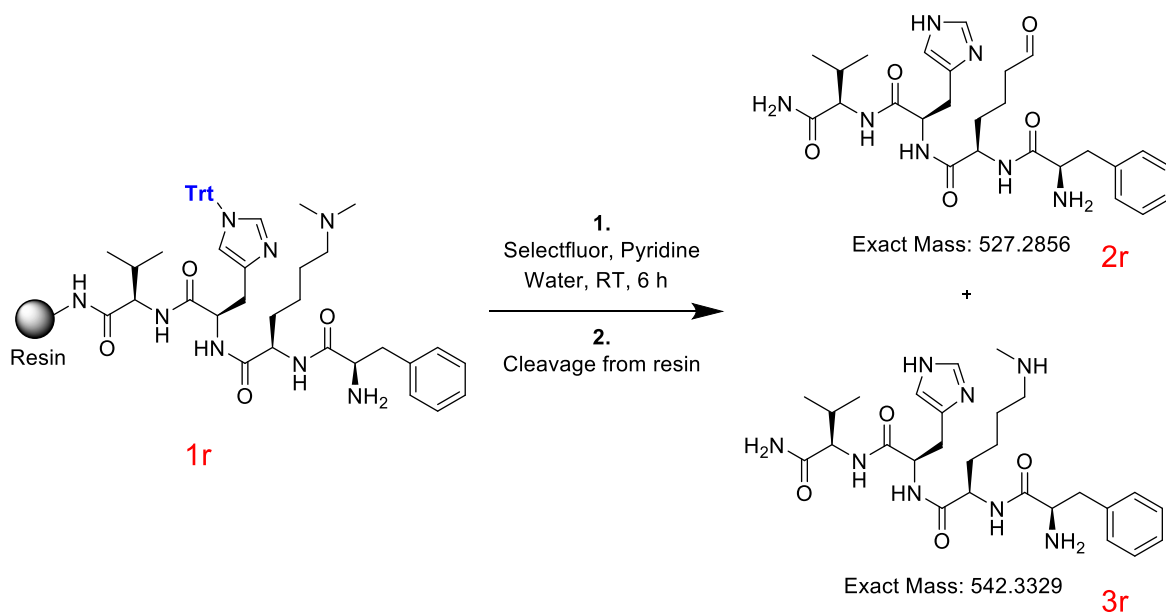
To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing tertiary amine peptide FKme₂H(Trt)V (1r) dissolved in 18 mL of water, was added 100 mg of selectfluor and 50 µL of pyridine (step 1). Two reactions mixtures were stirred for 2 and 6 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA (step 2). Following cleavage, peptide was dried under air and subsequently injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde at 2 h was 74% and at 6 h was 87%.

**Resin cleaved FKme₂HV peptide 1r**. LCMS: m/z 606.3788 (calcd [M+H]$^+$ = 557.3558), m/z 628.3115 (calcd [M+Na]$^+$ = 628.3582), m/z 1135.6862 (calcd [2M+H]$^+$ = 1135.6863). (HPLC analysis at 220 nm). Retention time in HPLC: 4.801

**Resin cleaved FK(CHO)HV peptide 2r**. LCMS: m/z 528.2431 (calcd [M+H]$^+$ = 528.2929), m/z 550.2743 (calcd [M+Na]$^+$ = 550.2748), m/z 1055.5327 (calcd [2M+H]$^+$ = 1055.5785). (HPLC analysis at 220 nm). Retention time in HPLC: 5.304

**Resin cleaved FKme₁HV peptide 3r**. LCMS: m/z 543.3416 (calcd [M+H]$^+$ = 543.3402), m/z 565.3128 (calcd [M+Na]$^+$ = 565.3221), m/z 1085.6222 (calcd [2M+H]$^+$ = 1085.6731). (HPLC analysis at 220 nm). Retention time in HPLC: 4.584
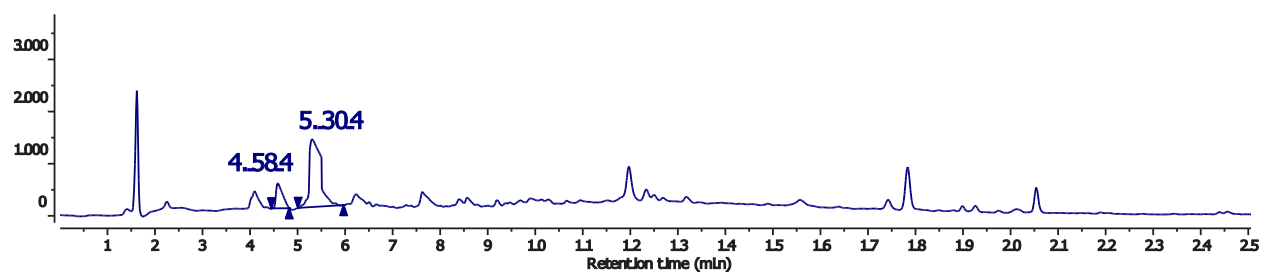
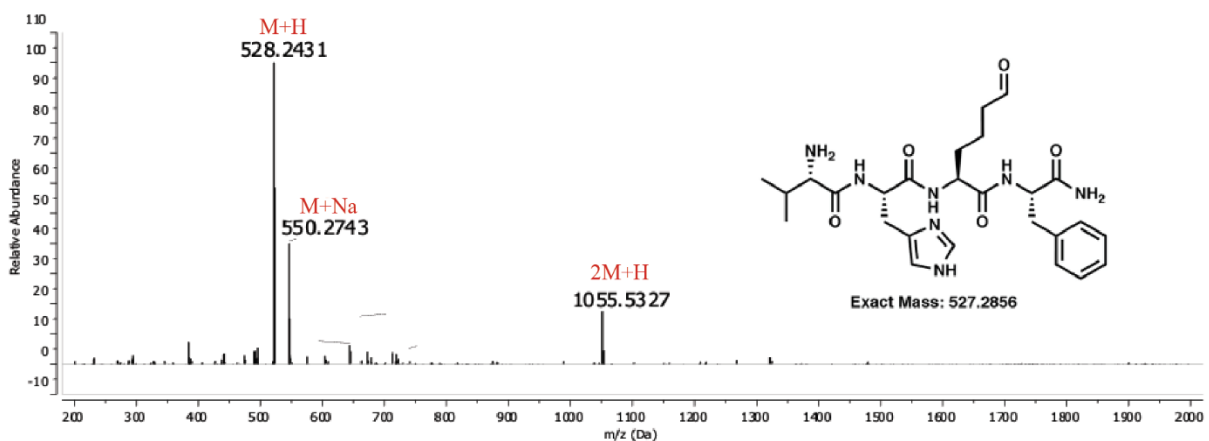**HPLC Trace of resin cleaved and deprotected 1r**
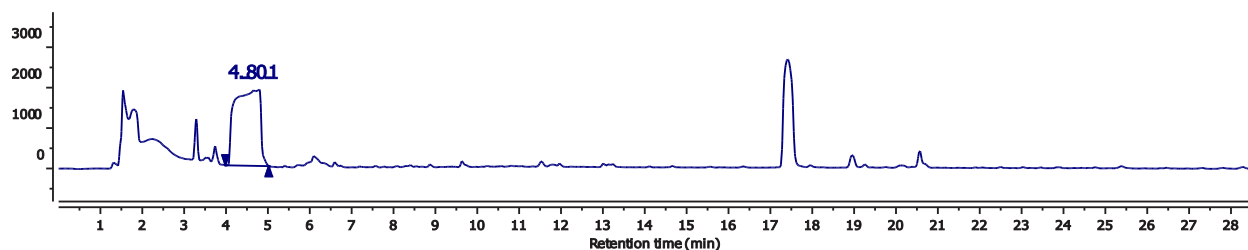


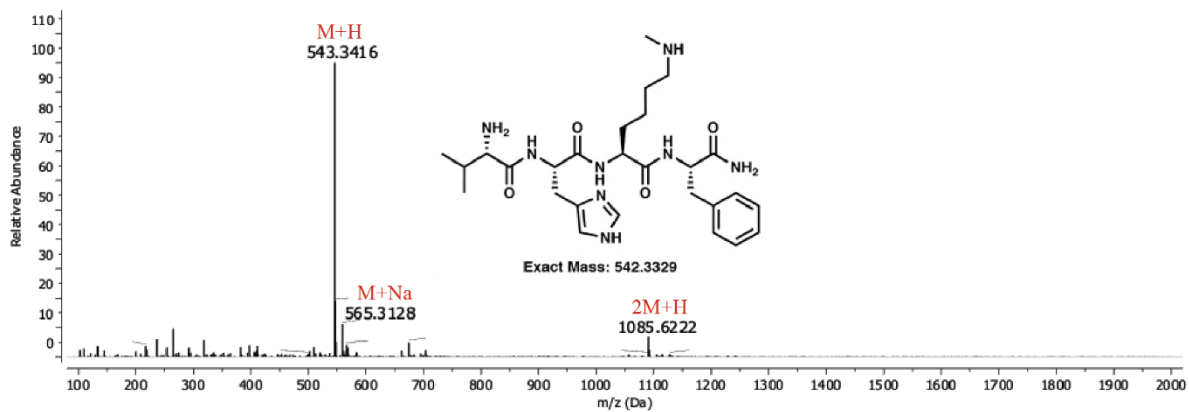**MS Trace of resin cleaved and deprotected 1r (peak 4.801)**
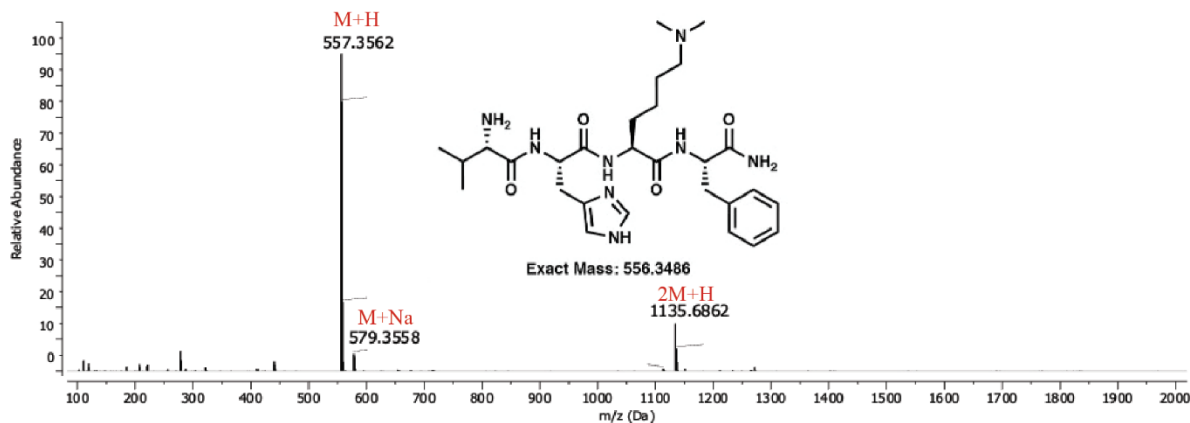
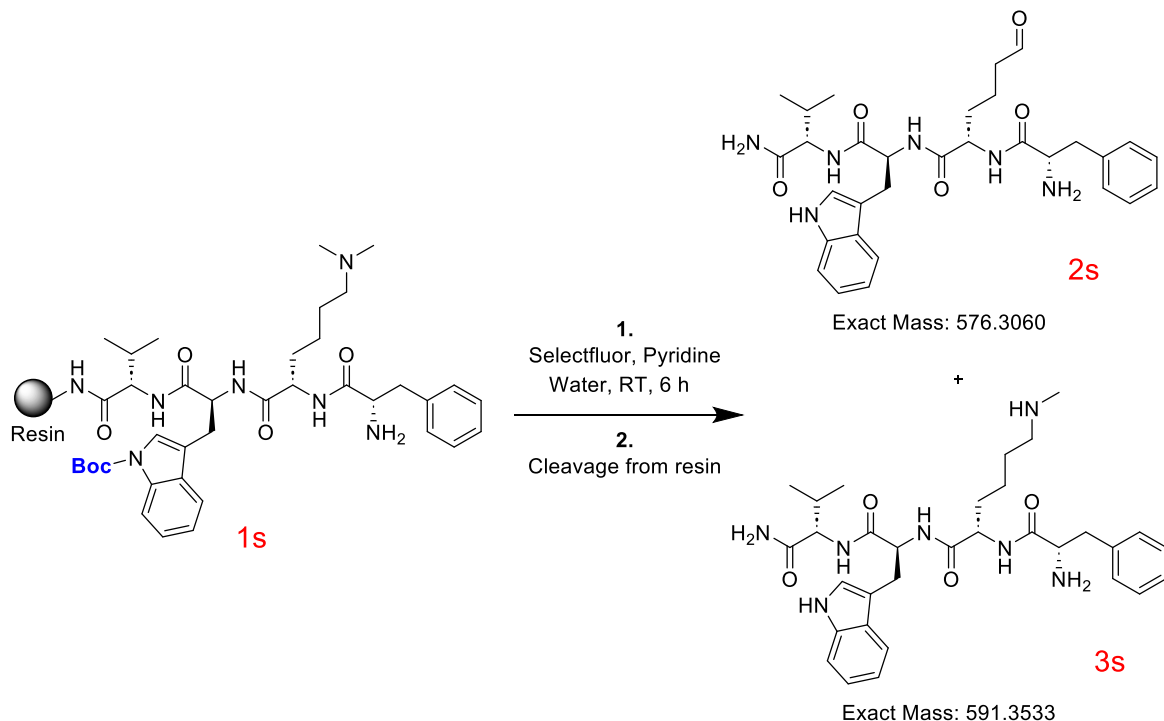## HPLC Trace of resin cleaved reaction after 6 h



## MS Trace of resin cleaved and deprotected 2r (peak 5.304)



## MS Trace of resin cleaved and deprotected 3r (peak 4.584)

**FKme₂WV:**



2s

Exact Mass: 576.3060

3s

Exact Mass: 591.3533

1.
Selectfluor, Pyridine
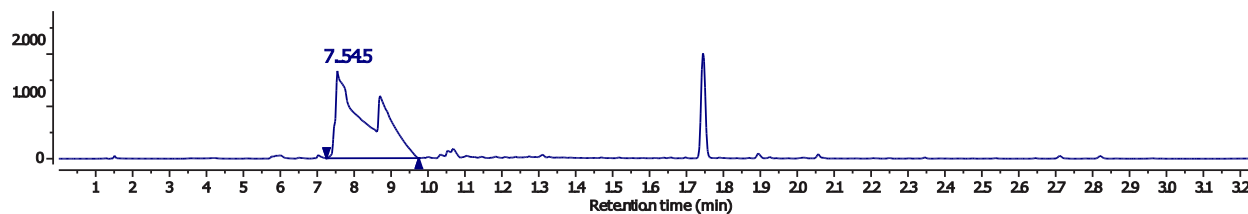Water, RT, 6 h

2.
Cleavage from resin

1s

Resin

Boc

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing tertiary amine peptide FKme₂W(Boc)V (1s) dissolved in 18 mL of water, was added 100 mg of selectfluor and 50 µL of pyridine. The reaction mixture was stirred for 6 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture. % conversion to peptide aldehyde was 84%.

**Resin cleaved FKme₂WV peptide 1s**. LCMS: m/z 606.3788 (calcd $[M+H]^+$ = 606.3762), m/z 628.3115 (calcd $[M+Na]^+$ = 628.3582), m/z 651.2158 (calcd $[M+2Na]^+$ = 651.3474), m/z 1233.6483 (calcd $[2M+Na]^+$ = 1233.7271). (HPLC analysis at 220 nm). Retention time in HPLC: 7.545

**FK(CHO)WV peptide 2s**. LCMS: m/z 577.3104 (calcd $[M+H]^+$ = 577.3133), m/z 599.3172 (calcd $[M+Na]^+$ = 599.2952). (HPLC analysis at 220 nm). Retention time in HPLC: 10.701

**FKme₁WV peptide 3s**. LCMS: m/z 592.3617 (calcd $[M+H]^+$ = 592.3606), m/z 1183.7066 (calcd $[2M+H]^+$ = 1183.7139). (HPLC analysis at 220 nm). Retention time in HPLC: 8.189
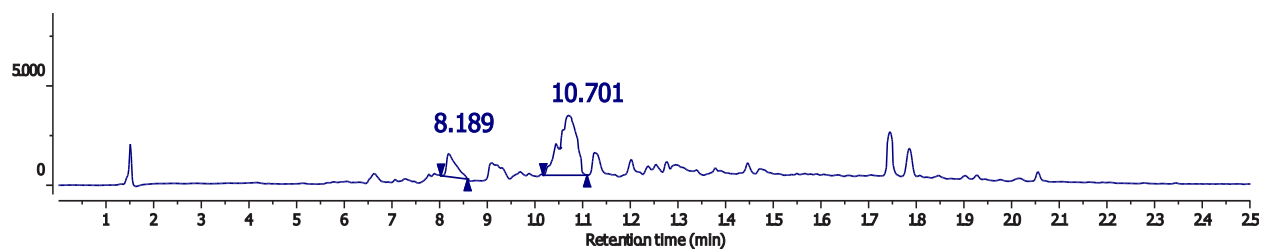
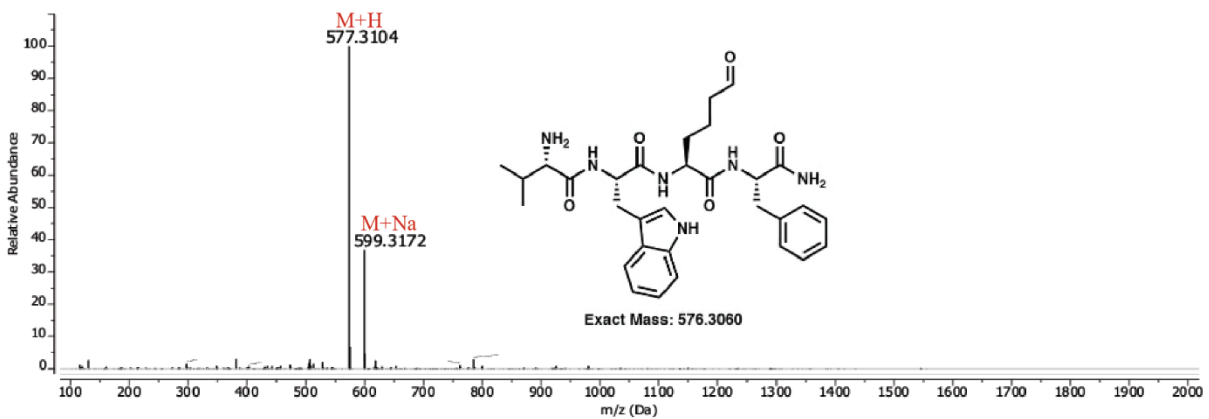## HPLC Trace of resin cleaved and deprotected 1s



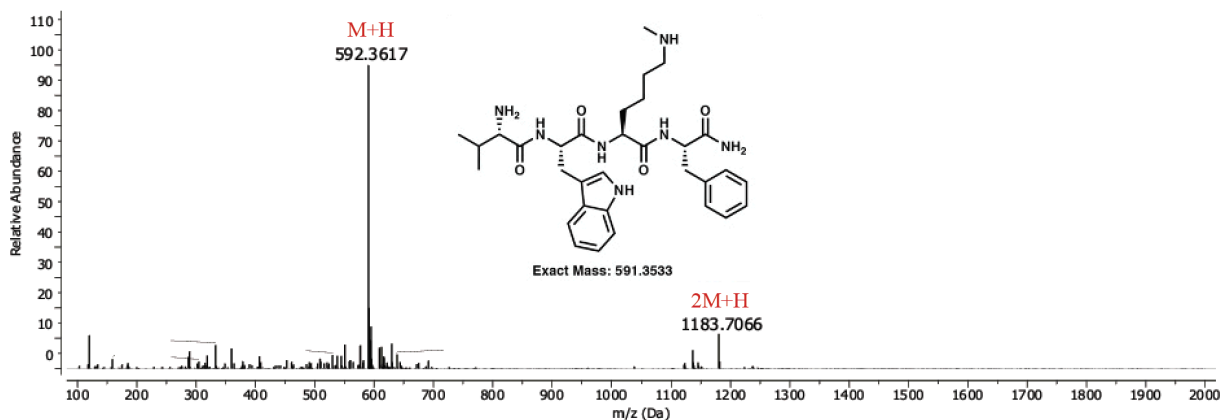## MS Trace of resin cleaved and deprotected 1s (peak 7.545)
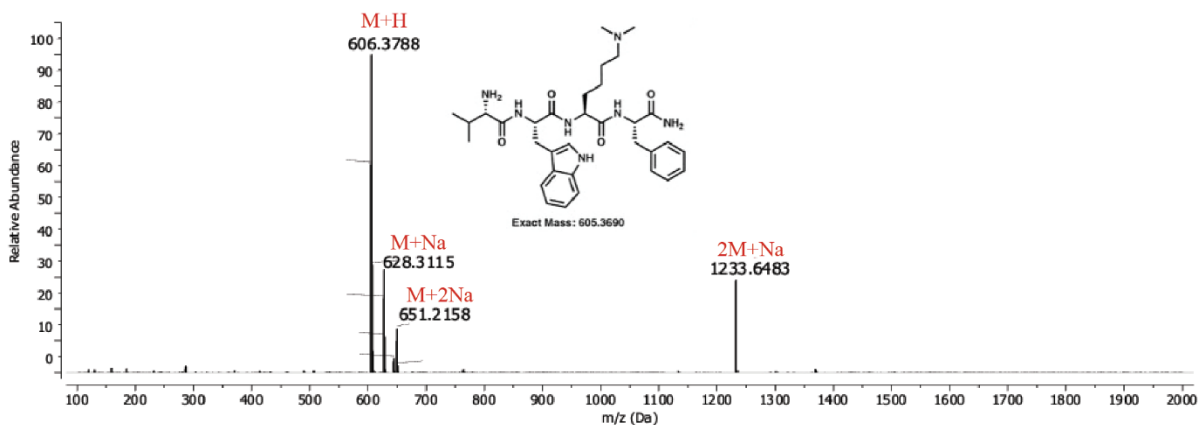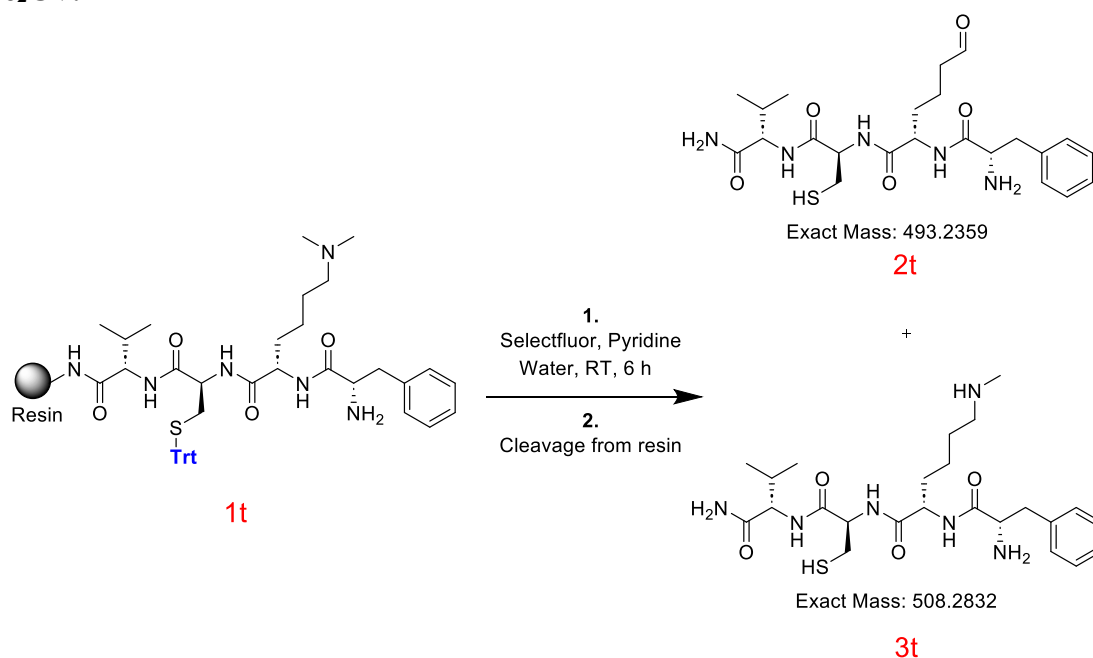


## HPLC Trace of resin cleaved reaction after 6 h



## MS Trace of resin cleaved and deprotected 2s (peak 10.701)

**MS Trace of resin cleaved and deprotected 3s (peak 8.189)**



Exact Mass: 591.3533

**FKme₂CV:**



Exact Mass: 493.2359
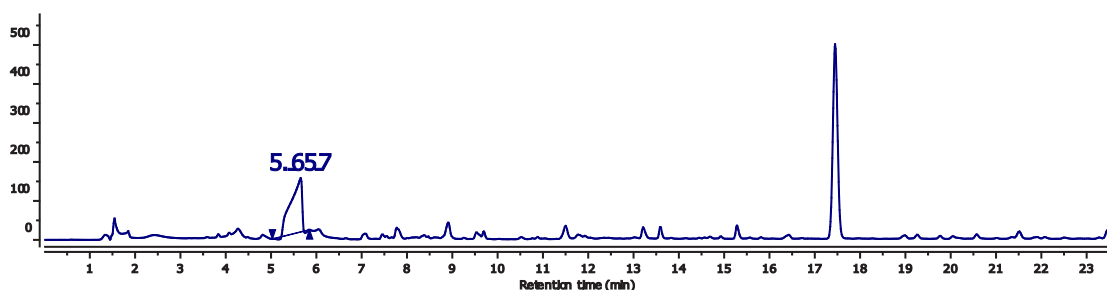**2t**

Exact Mass: 508.2832
**3t**

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing tertiary amine peptide FKme₂C(Trt)V (1t) dissolved in 18 mL of water, was added 100 mg of selectfluor and 50 µL of pyridine. The reaction mixture was stirred for 6 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into LC-MS to monitor the reaction. The reaction mixture was analyzed by the HPLC method reported in the analytical method. % conversion was determined by calculating the area under the HPLC peaks of reaction mixture.  % conversion to peptide aldehyde was 91%.

**Resin cleaved FKme₂CV peptide 1t**. LCMS: m/z 523.3788 (calcd $[M+H]^+$ = 523.3061), m/z 545.3115 (calcd $[M+Na]^+$ = 545.2880), m/z 1067.6862 (calcd $[2M+Na]^+$ = 1067.5869). (HPLC analysis at 220 nm). Retention time in HPLC: 5.657

**Resin cleaved FK(CHO)CV peptide 2t**. LCMS: m/z 494.2436 (calcd $[M+H]^+$ = 494.2432), m/z 512.2533 (calcd $[M+H_2O]^+$ = 512.2537), m/z 987.4796 (calcd $[2M+H]^+$ = 987.4791). (HPLC analysis at 220 nm). Retention time in HPLC: 8.259

**Resin cleaved FKme₁CV peptide 3t**. LCMS: m/z 509.2938 (calcd $[M+H]^+$ = 509.2905), m/z 527.3021 (calcd $[M+H_2O]^+$ = 527.3010), m/z 1017.5723 (calcd $[2M+H]^+$ = 1017.5736), m/z 1035.5864 (calcd $[2M+H_2O]^+$ = 1035.5842). (HPLC analysis at 220 nm). Retention time in HPLC: 5.653
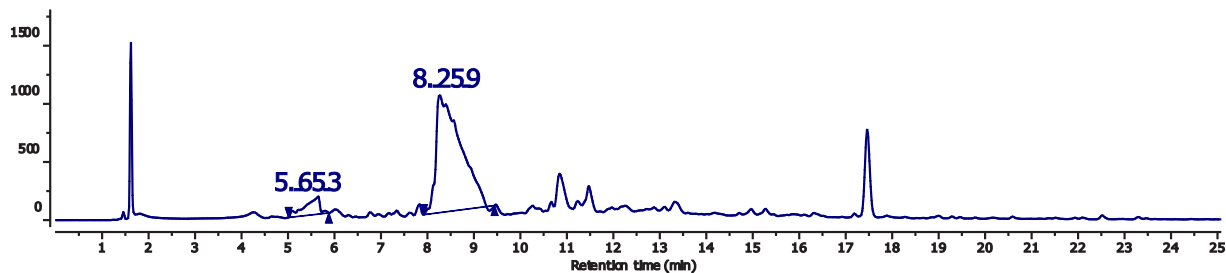
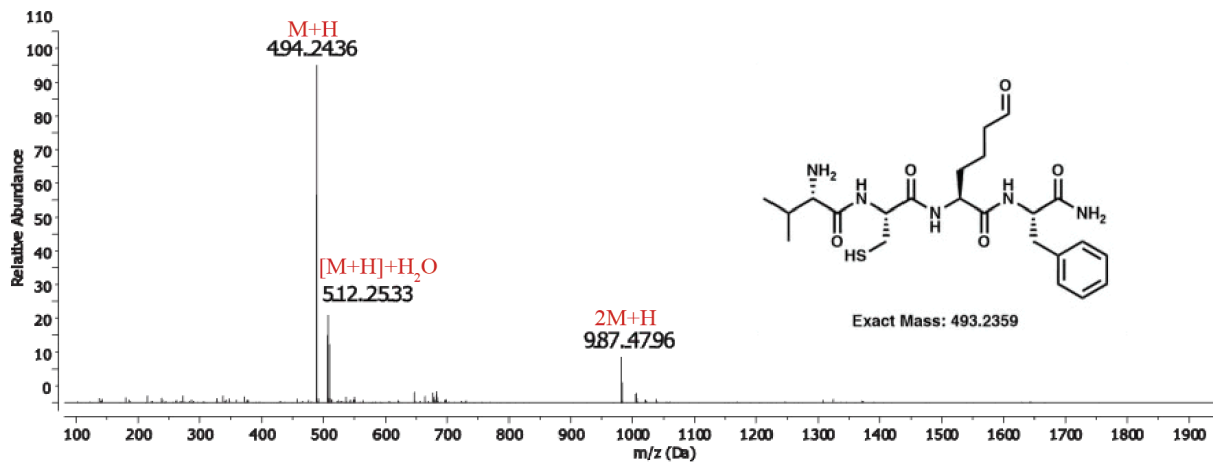**HPLC Trace of resin cleaved and deprotected 1t**



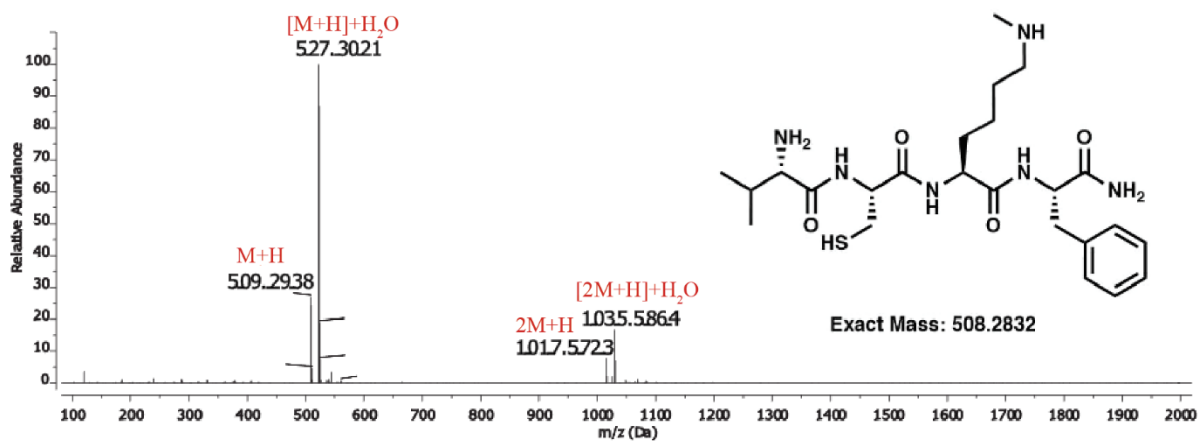**MS Trace of resin cleaved and deprotected 1t (peak 5.657)**
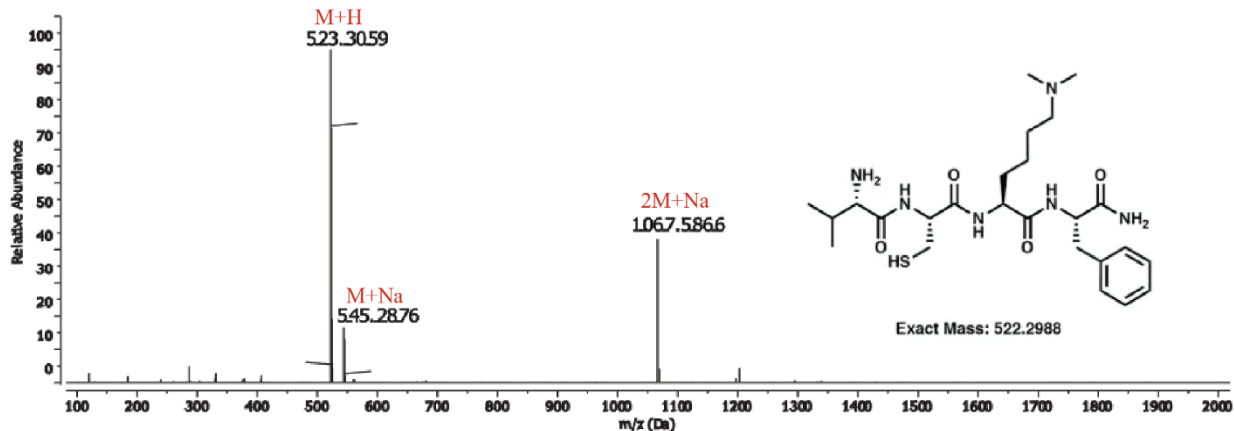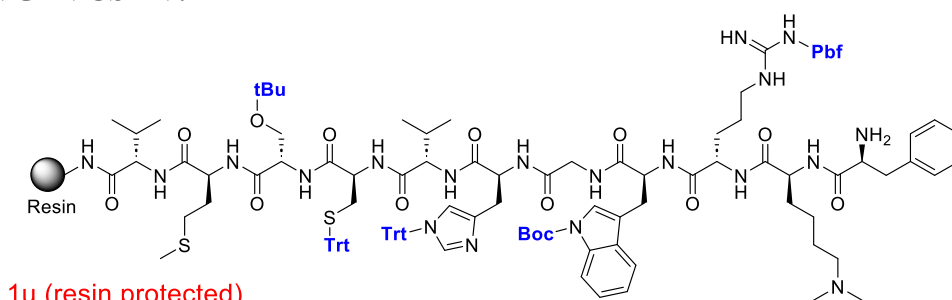


**HPLC Trace of resin cleaved reaction after 6 h**
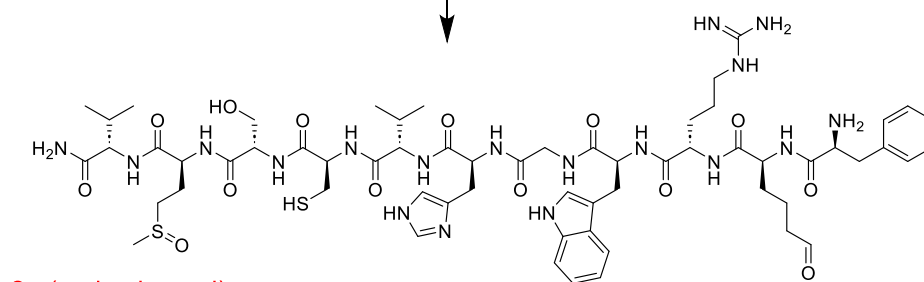
**MS Trace of resin cleaved and deprotected 2t (peak 8.259)**



**MS Trace of resin cleaved and deprotected 3t (peak 5.653)**

**FKme₂RWGHVCSMV:**



1u (resin protected)

*Exact Mass of deprotected peptide: 1376.7078*

**1.**
Selectfluor, Pyridine
Water, RT, 6 h

**2.**
Cleavage from resin

2u (resin cleaved)

Exact Mass: 1362.6325

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing dimethyllysine peptide FKme₂R(Pbf)W(Boc)GH(Trt)VC(Trt)S(tBu)MV dissolved in 18 mL of water, was added 100 mg of selectfluor and 50 µL of pyridine. The reaction mixture was stirred for 6 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into MS to monitor the reaction. No dimethyllysine peptide (1u) was observed in the reaction. Instead, we observed the aldehyde peptide product (2u).
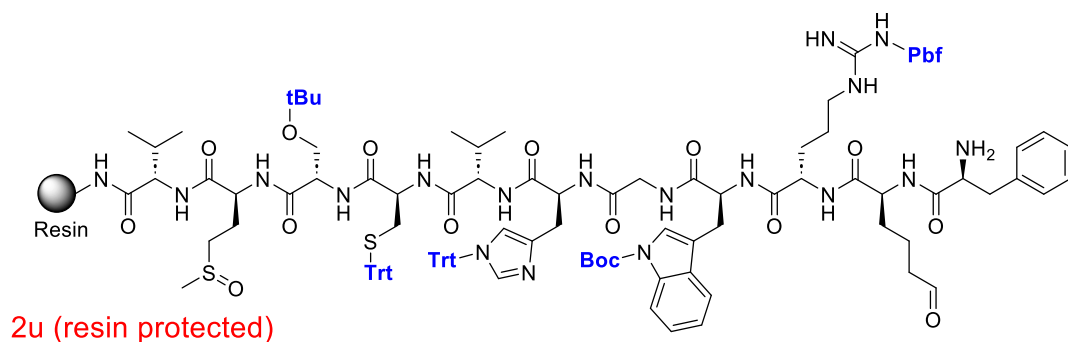
**Resin cleaved FKme₂RW(Boc)GH(Trt)VC(Trt)SMV peptide 1u**. LCMS: m/z 1376.6647 (calcd [M+H]$^+$ = 1376.7078), m/z 688.8541 (calcd [M+2/2]$^+$ = 688.8539), m/z 459.5699 (calcd [M+3/3]$^+$ = 459.5692).

**Resin cleaved FK(CHO)RWGHVCSMV peptide 2u**. LCMS: m/z 1363.6442 (calcd [M+H]$^+$ = 1363.6398), m/z 682.3213 (calcd [M+2/2]$^+$ = 682.3199), m/z 455.2160 (calcd [M+3/3]$^+$ = 455.2132).
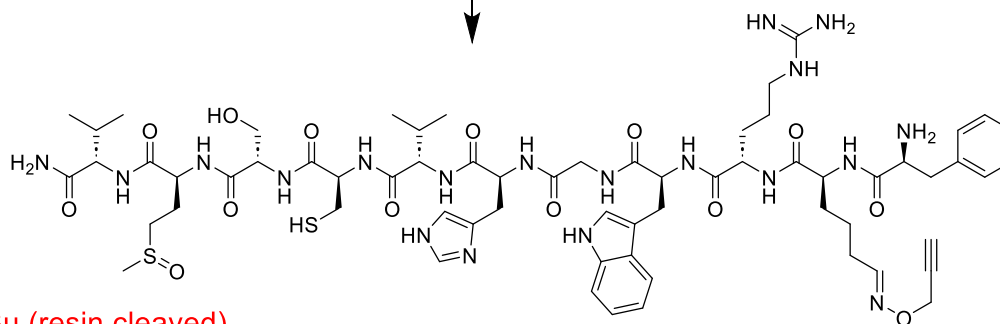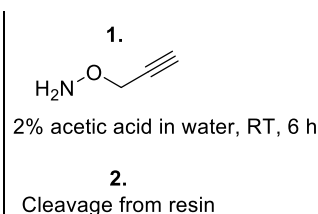
**MS Trace of resin deprotected and cleave peptide (1u) FKme₂RWGHVCSMV**
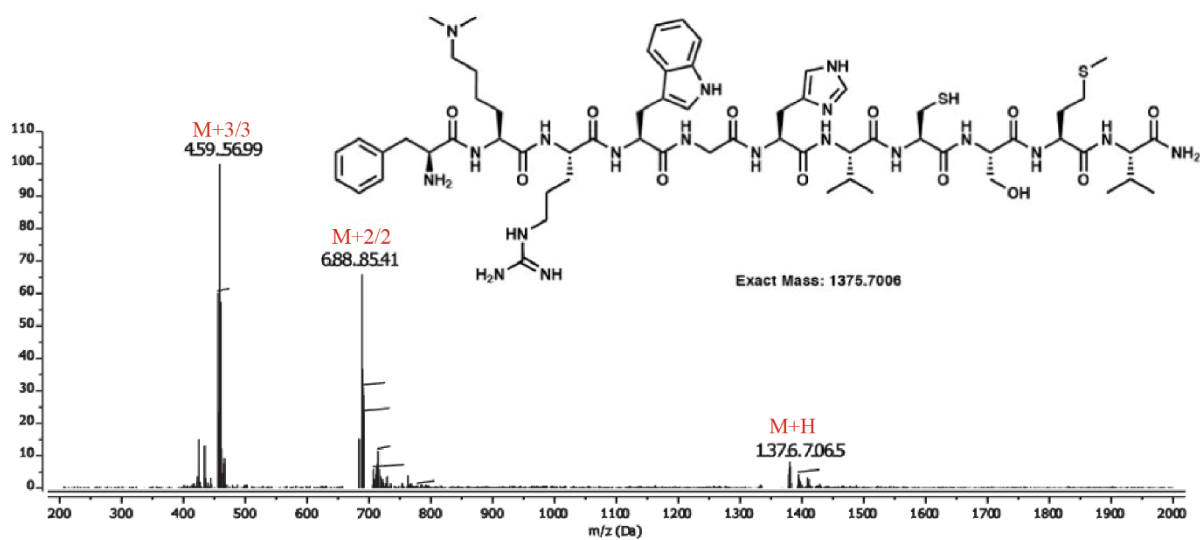


**MS Trace of resin deprotected and cleave peptide (2u) FK(CHO)RWGHVCSMV**
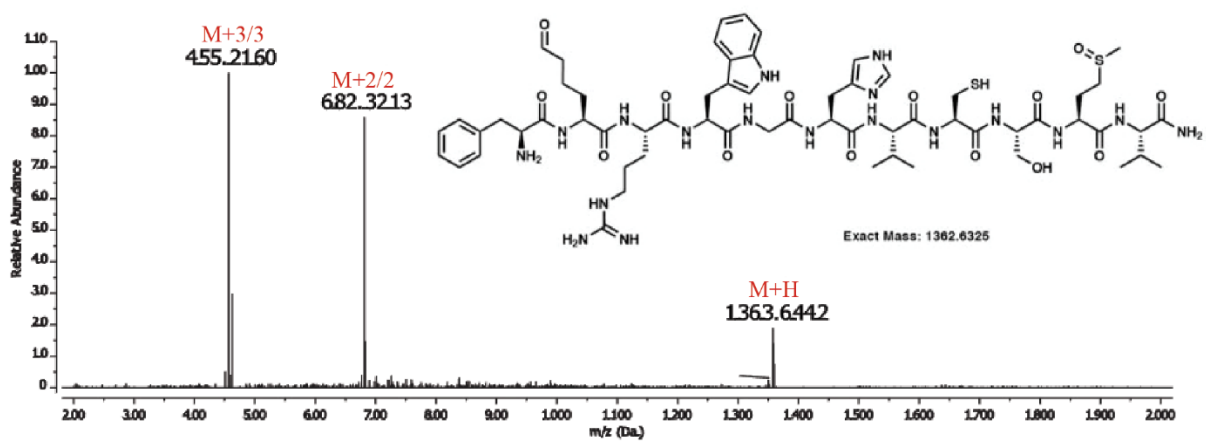
## XVII. Supplementary Figure 9: Solid-support mediated diversification of allysine peptides

**FK(CHO)RWGHCSMV reaction with *o*-2-propynylhydroxylamine hydrochloride:**



2u (resin protected)

*Exact Mass of deprotected peptide: 1363.6398*

1.

H₂N—O—propargyl

2% acetic acid in water, RT, 6 h

2.
Cleavage from resin

3u (resin cleaved)

Exact Mass: 1415.6591

To 100 mg of PEGA-amide (0.20 - 0.50 mmol/g) resin containing dimethyllysine peptide FK(CHO)RW(Boc)GH(Trt)VC(Trt)SMV dissolved in 18 mL of 2% acetic acid in water, was added 10 mg of *o*-2-propynylhydroxylamine hydrochloride. The reaction mixture was stirred for 6 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into MS to monitor the reaction. No aldehyde containing peptide (2u) was observed in the reaction. Instead, we observed the oxime containing peptide product (3u).

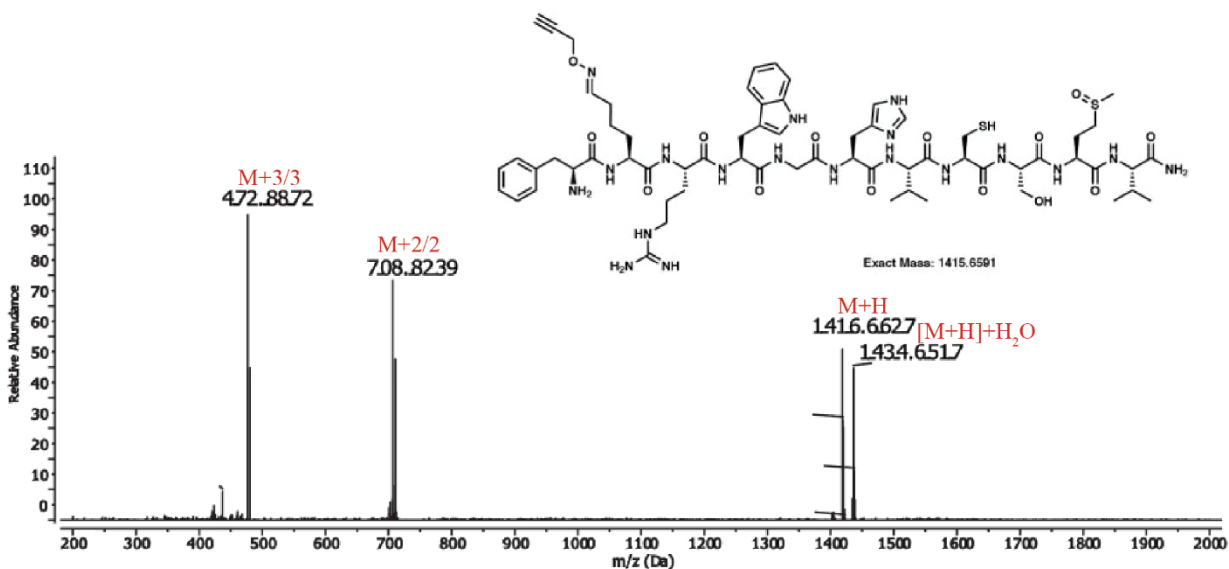**Resin cleaved FK(CHO)RWGHVCSMV peptide 2u**. LCMS: m/z 1363.6442 (calcd $[M+H]^+$ = 1363.6398), m/z 682.3213 (calcd $[M+2/2]^+$ = 682.3199), m/z 455.2160 (calcd $[M+3/3]^+$ = 455.2132).
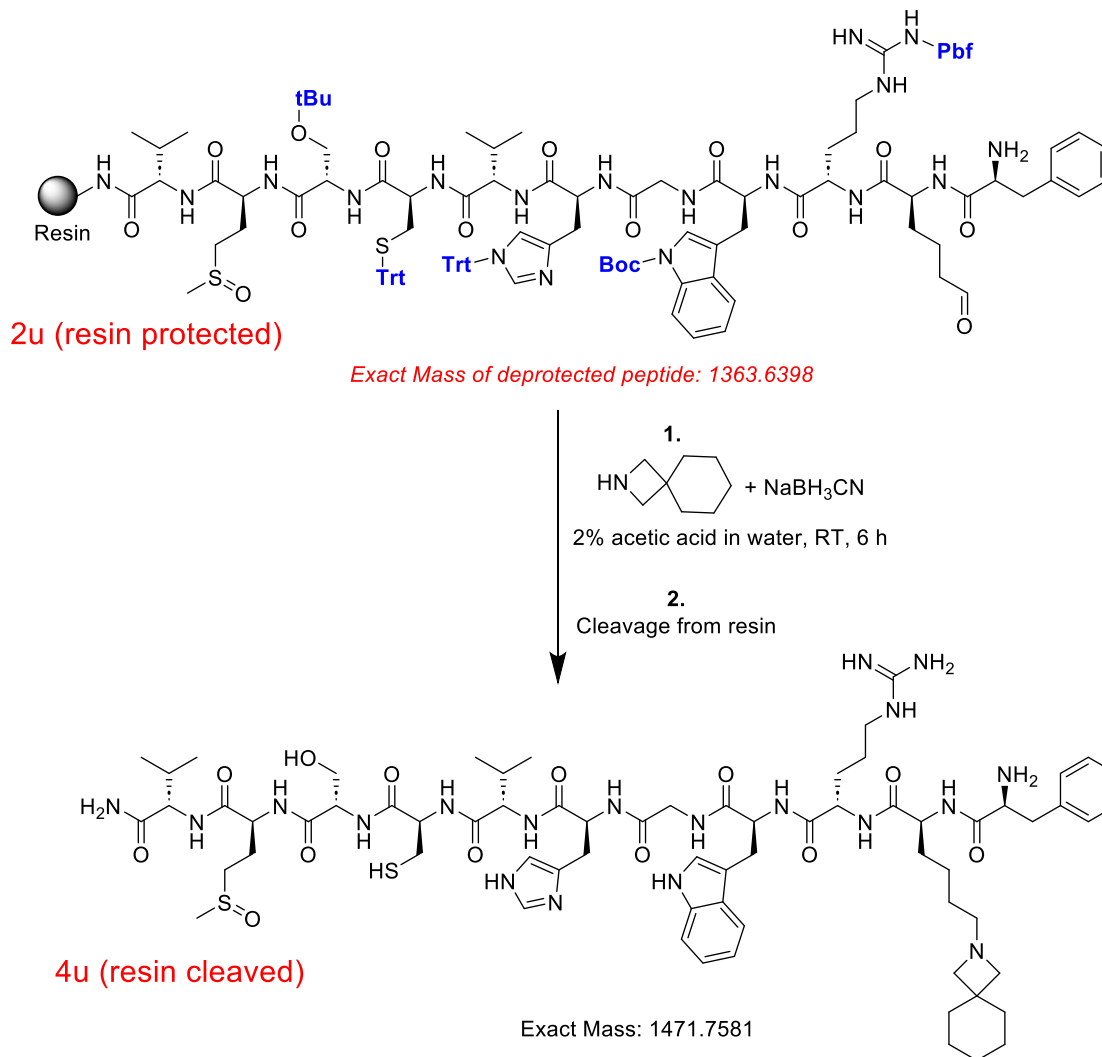
**Resin cleaved FK(CHO)RWGHVCSMV oxime peptide 3u**. LCMS: m/z 1416.6627 (calcd $[M+H]^+$ = 1416.6664), m/z 1434.6517 (calcd $[M+H_2O+1]^+$ = 1434.6769), m/z 708.8239 (calcd $[M+2/2]^+$ = 708.8332), m/z 472.8872 (calcd $[M+3/3]^+$ = 472.8888).

**MS Trace of resin deprotected and cleave peptide (2u) FK(CHO)RWGHVCSMV**
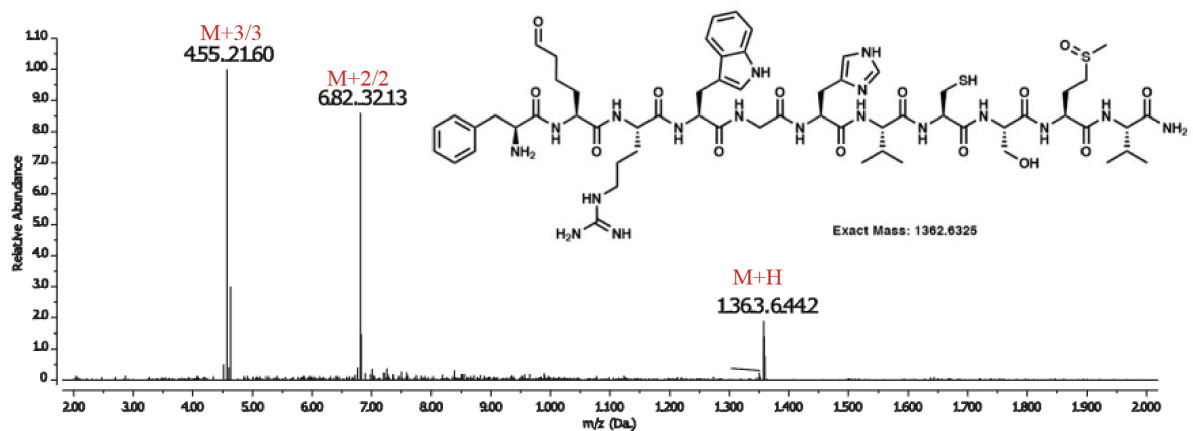


**MS Trace of Resin cleaved FK(CHO)RWGHVCSMV oxime peptide 3u**

**FK(CHO)RWGHCSMV reaction with 2-Azaspiro[3,5]nonane Hydrochloride:**



2u (resin protected)

*Exact Mass of deprotected peptide: 1363.6398*

1.

+ NaBH$_3$CN

2% acetic acid in water, RT, 6 h

2.
Cleavage from resin

4u (resin cleaved)
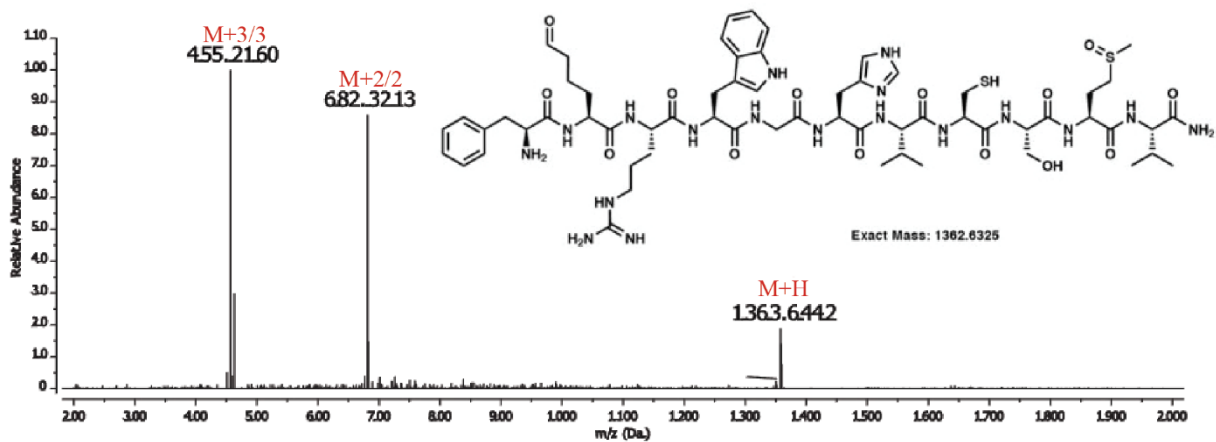
Exact Mass: 1471.7581

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing dimethyllysine peptide FK(CHO)RWGHVCSMV dissolved in 18 mL of 2% acetic acid in water, was added 10 mg of 2-Azaspiro[3,5]nonane Hydrochloride. The reaction mixture was stirred for 1 h, followed by the addition of 20 mg of sodium cyanoborohydride. The reaction mixture was stirred for 5 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into MS to monitor the reaction. No aldehyde containing peptide (2u) was observed in the reaction. Instead, we observed the tertiary amine peptide product (4u).

**Resin cleaved FK(CHO)RWGHVCSMV peptide 2u**. LCMS: m/z 1363.6442 (calcd [M+H]$^+$ = 1363.6398), m/z 682.3213 (calcd [M+2/2]$^+$ = 682.3199), m/z 455.2160 (calcd [M+3/3]$^+$ = 455.2132).

**Resin cleaved FK(CHO)RWGHVCSMV Azaspiro[3,5]nonane peptide 4u**. LCMS: m/z 1472.7651 (calcd [M+H]$^+$ = 1472.7653), m/z 1490.7789 (calcd [M+H$_2$O+H]$^+$ = 1490.7759), m/z

736.8818 (calcd [M+2/2]$^+$ = 736.8826), m/z 754.8861 (calcd [(M+2/2)+H$_2$O]$^+$ = 754.8826), m/z 491.5881 (calcd [M+3/3]$^+$ = 491.5884).

**MS Trace of resin deprotected and cleave peptide (2u) FK(CHO)RWGHVCSMV**



**MS Trace of resin cleaved FK(CHO)RWGHVCSMV Azaspiro[3,5]nonane peptide 4u**

**FK(CHO)RWGHCSMV reaction with methylhydrazine:**



2u (resin protected)

*Exact Mass of deprotected peptide: 1363.6398*

**1.**

2% acetic acid in water, RT, 6 h

**2.**
Cleavage from resin

5u (resin cleaved)

Exact Mass: 1390.6751

To 100 mg of PEGA-amide resin (0.20 - 0.50 mmol/g) containing dimethyllysine peptide FK(CHO)RWGHVCSMV dissolved in 18 mL of 2% acetic acid in water, was added 50 µL of methylhydrazine. The reaction mixture was stirred for 5 h. Resin was washed with DMF, MeOH, and DCM before cleavage in 95% TFA. Following cleavage, peptide was dried under air and subsequently injected into MS to monitor the reaction. No aldehyde containing peptide (2u) was observed in the reaction. Instead, we observed the hydrazone peptide product (5u).

**FK(CHO)RWGHVCSMV peptide 2u**. LCMS: m/z 1363.6442 (calcd [M+H]$^+$ = 1363.6398), m/z 682.3213 (calcd [M+2/2]$^+$ = 682.3199), m/z 455.2160 (calcd [M+3/3]$^+$ = 455.2132).
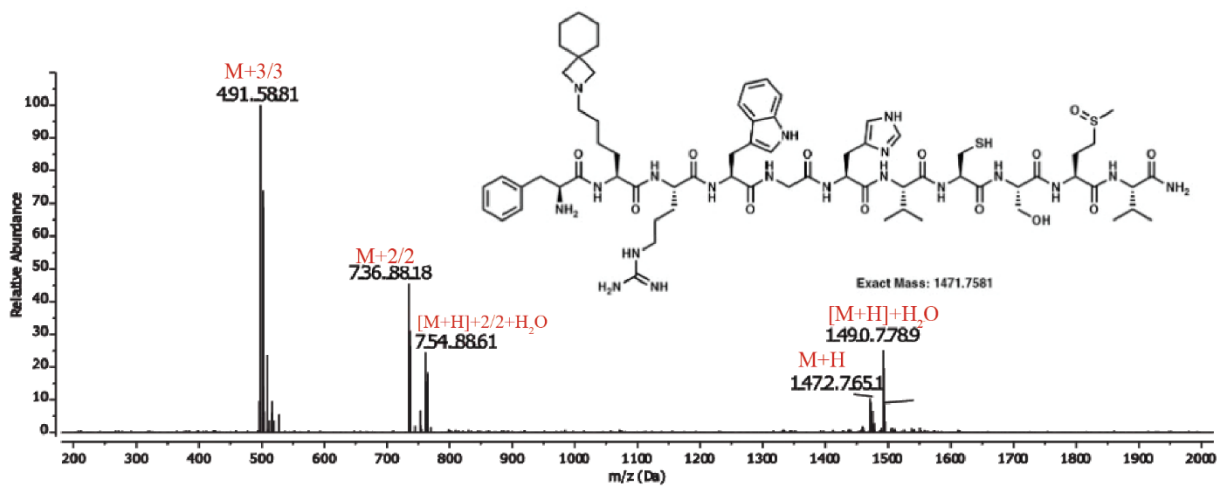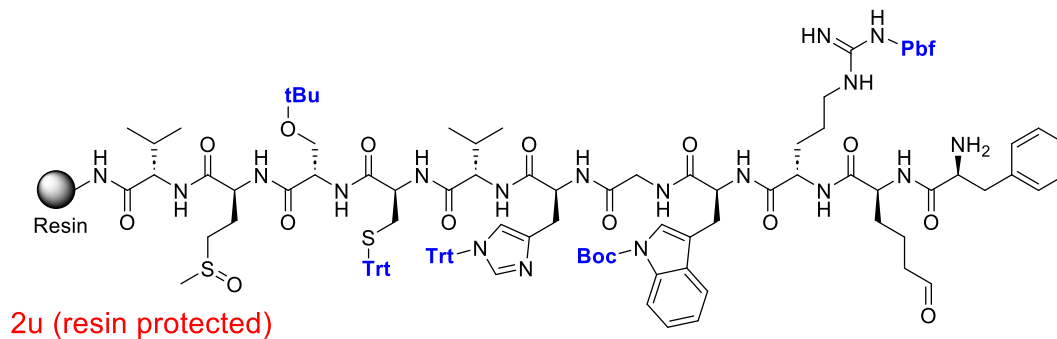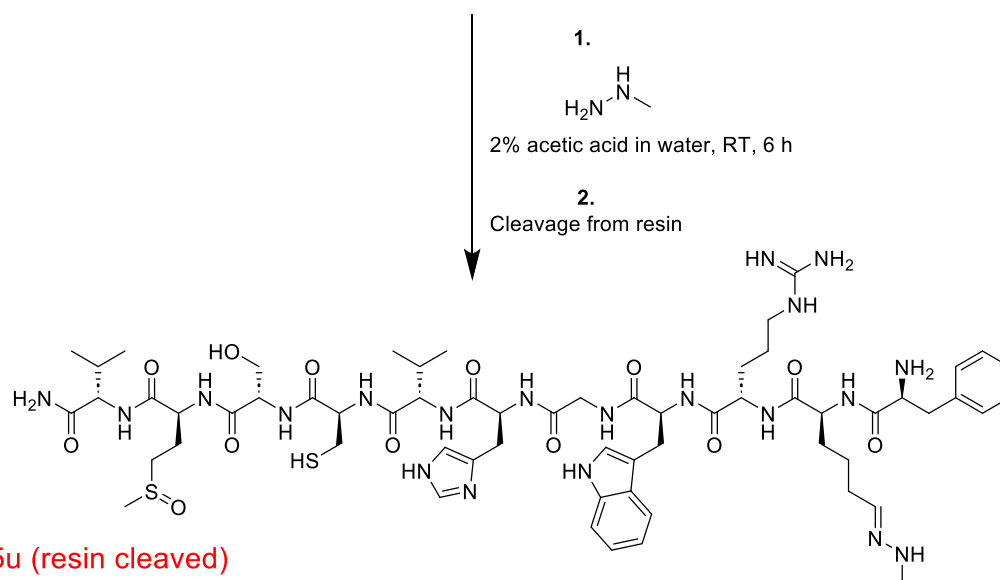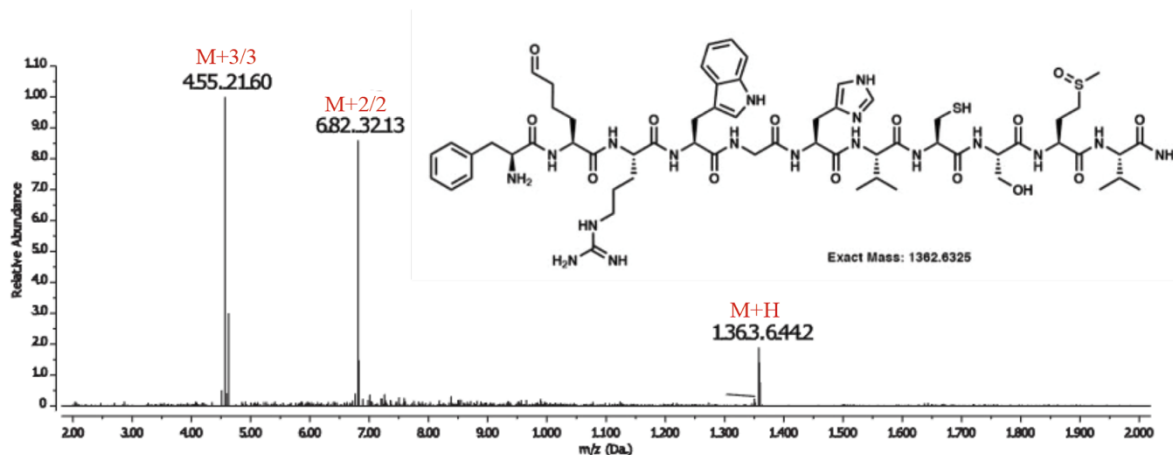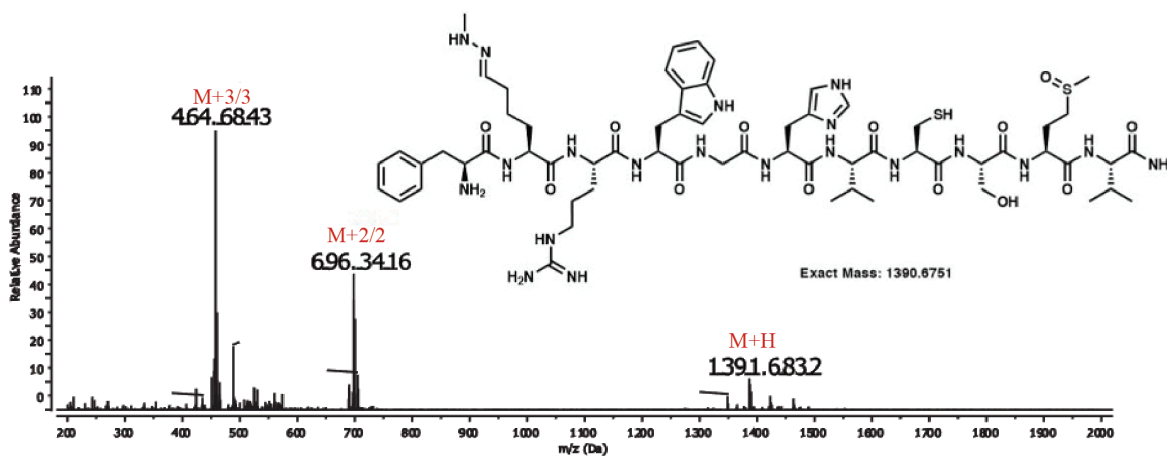
**FK(CHO)RWGHVCSMV hydrazine peptide 5u**. LCMS: m/z 1391.6832 (calcd [M+H]$^+$ = 1391.6823), m/z 696.3416 (calcd [M+2/2]$^+$ = 696.3411), m/z 464.6843 (calcd [M+3/3]$^+$ = 464.5607).

**MS trace of FK(CHO)RWGHVCSMV hydrazine peptide 5u**



## XVIII. Supplementary Figure 10. Confocal microscopy imaging of peptide-fluorophore conjugates.

**Synthesis of peptide fluorophore conjugate:**

To 0.5 mg of allysine containing peptide LILK(CHO)PF (2b) dissolved in 400 µL of water, was added 0.5 eq of Alexafluor-647 dye. The reaction mixture was stirred for 6 h. Samples were taken from the reaction mixture and injected into LC-MS to monitor the generation of peptide-fluorophore product.

**LILKme$_2$PF peptide 2b**. LCMS: m/z 758.5179 (calcd [M+H]$^+$ = 758.5175), m/z 379.7628 (calcd [M+2/2]$^+$ = 379.4545). (HPLC analysis at 220 nm). Retention time in HPLC: 10.793.

**Peptide-2b-fluorophore**. LCMS: m/z 956.51541 (calcd [M+2/2]$^+$ = 956.5958), m/z 637.66616 (calcd [M+3/3]$^+$ = 637.6855). Retention time in HPLC: 19.142

**HPLC Trace of peptide 1b**
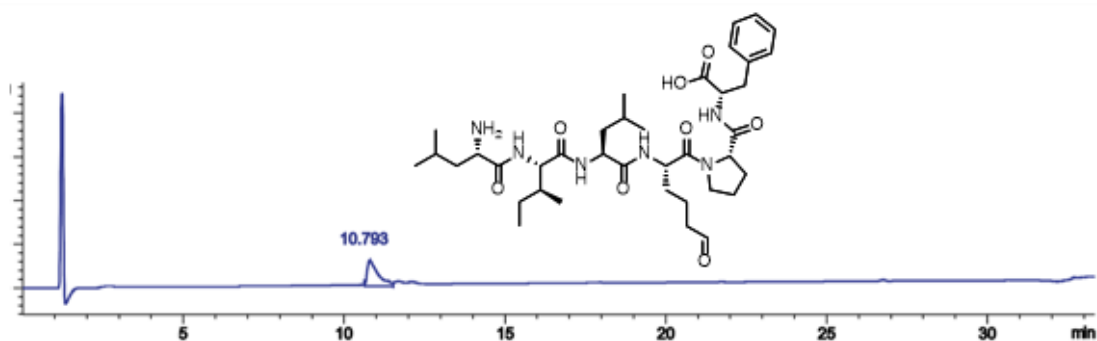


**MS-Trace of peptide 2b**


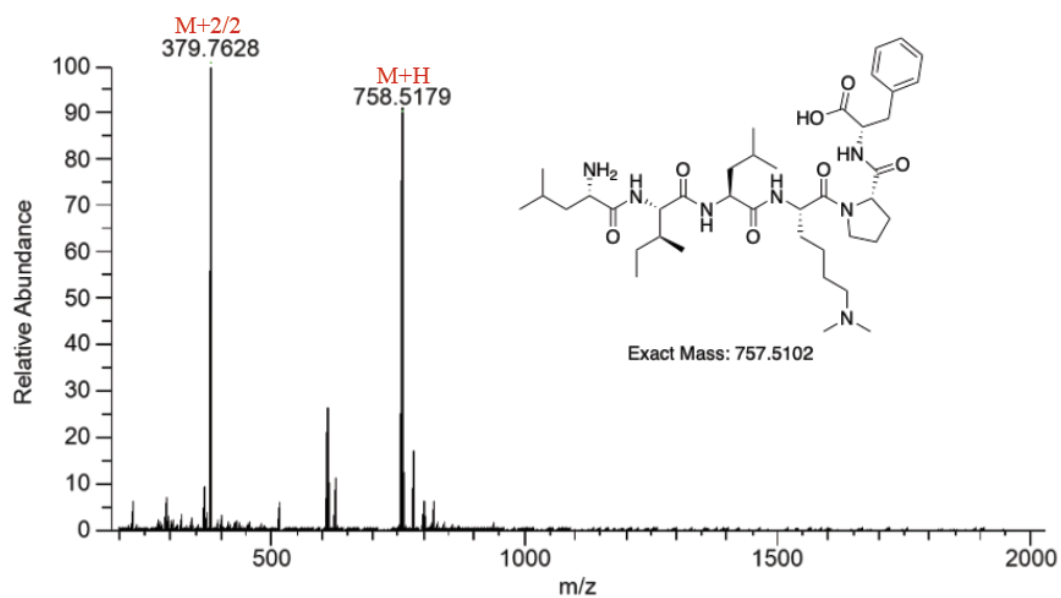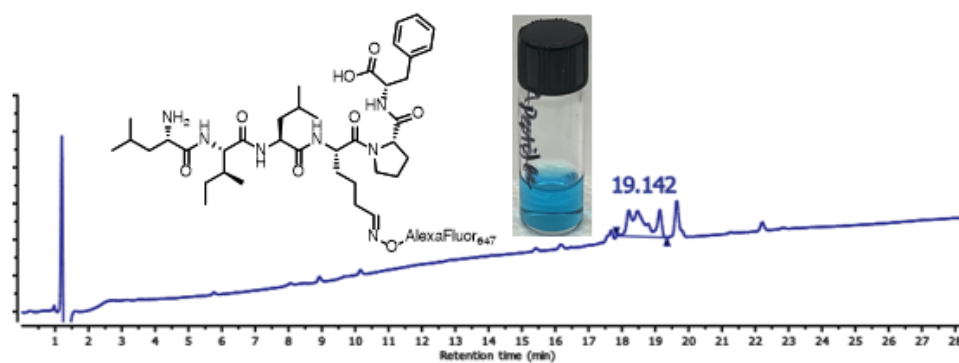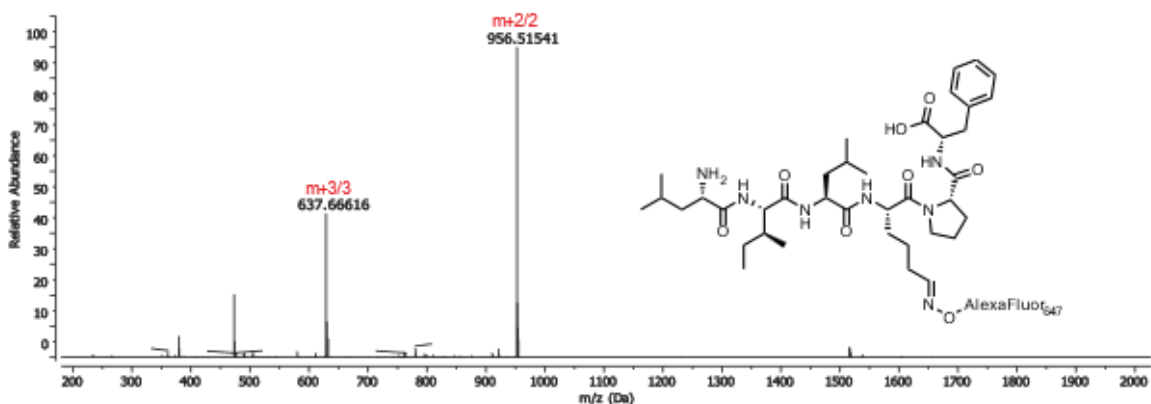
**HPLC Trace of reaction**

**MS-Trace of peptide-2b-fluorophore**



**Peptide fluorophore delivery into T47D cells:**

Live T47D cells were plated on microscope slides in a 6 cm petri dish supplemented with RPMI media containing 10% (v/v) fetal bovine serum (FBS), 1% (v/v) penicillin/streptomycin (100 µg/mL), and amphotericin B (1 mL), and incubated at 37 °C and 5% $CO_2$ for 24 h. T47D cells were incubated with 150 µM of peptide-2b-fluorophore conjugate for 12 h in the dark. Negative control was generated by incubating cells with 150 µM peptide 2b without any fluorophore for 12 h. Cells were washed 3 times with cold PBS and fixed in 4% formaldehyde solution for 10 min, and subsequently washed 3 times with PBS. Another negative control includes untreated cells. Nuclear staining of cells (experiment and controls) was done with 2 drops of Fluoroshield-DAPI mounting media on each microscope slide and subsequently imaged on a Leica SP8 confocal microscope. The images were processed and analyzed using ImageJ software. Experiments were performed in duplicates.
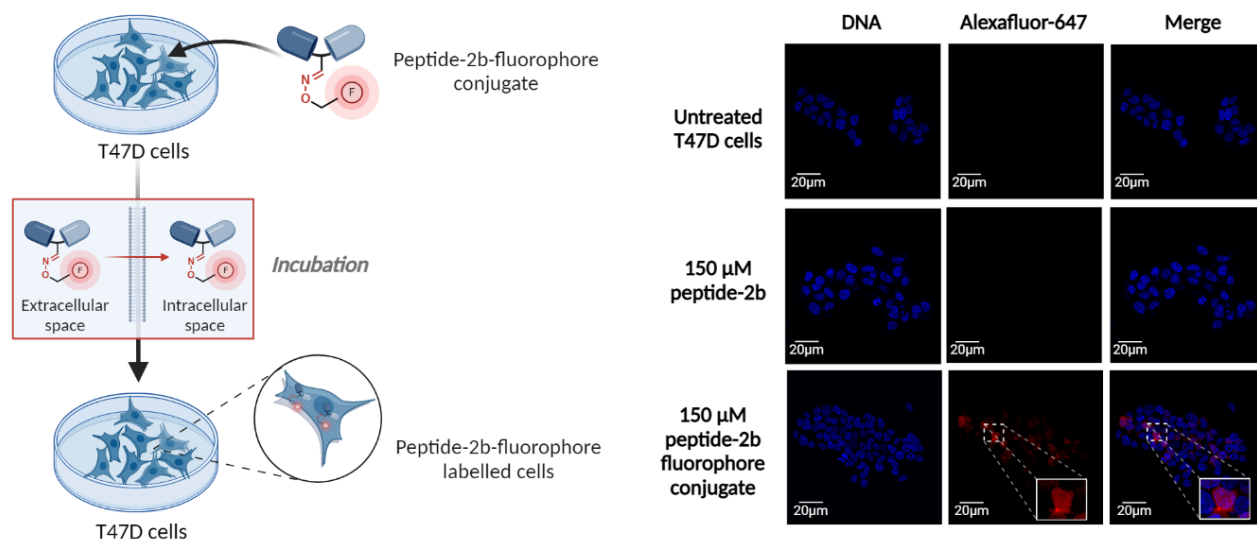
**XIX. Supplementary Figure 11. Confocal microscopy imaging of allysine labeled cells.**
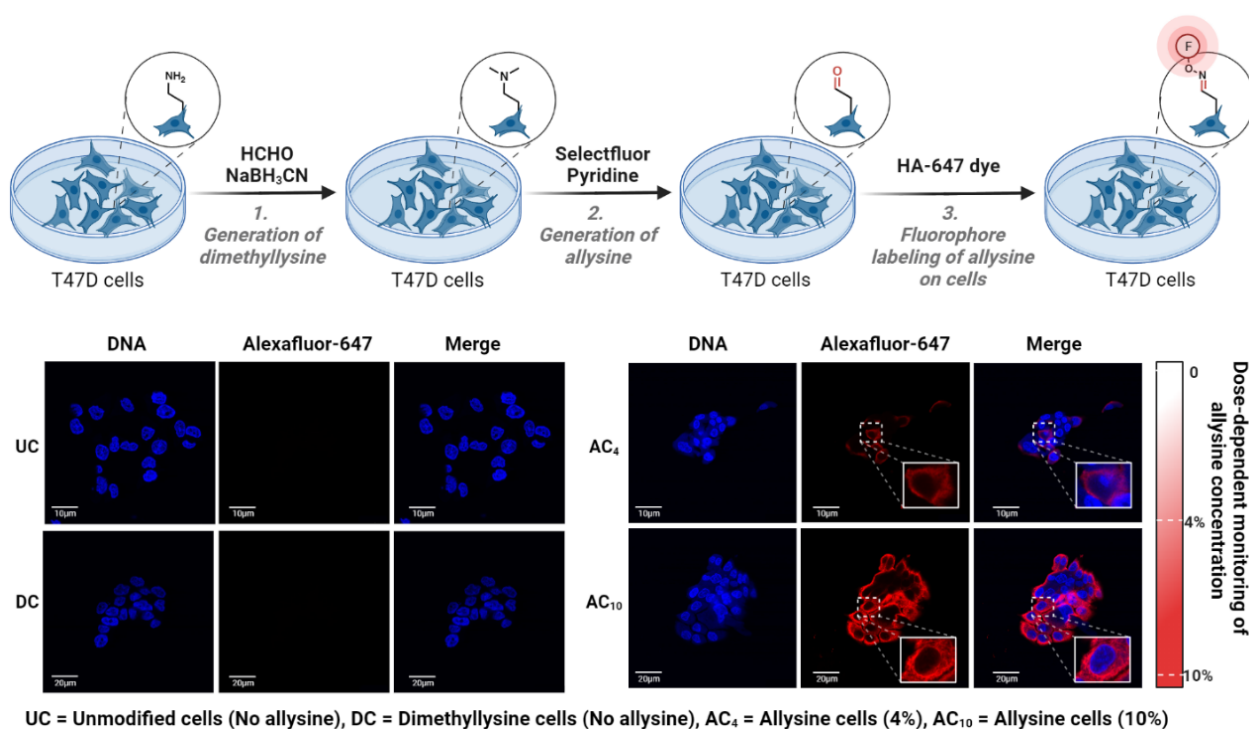
Live T47D cells were plated on microscope slides in a 6 cm petri dish supplemented with RPMI media containing 10% (v/v) fetal bovine serum (FBS), 1% (v/v) penicillin/streptomycin (100 µg/mL), and amphotericin B (1 mL), and incubated at 37 °C and 5% CO2 for 24 h. Cells were washed 3 times with cold PBS and fixed in 4% formaldehyde solution for 10 min. Cells were subsequently washed 3 times with PBS (5 min) and permeabilized with freshly prepared 0.1% Triton-X solution in PBS for 10 min. Cells were incubated with 4% and 10% formaldehyde solution for 10 min, followed by addition of 100 µL of 600 mM sodium cyanoborohydride in water. Cells were incubated for 20 min and washed 3 times with cold PBS. Microscope slides with dimethyllysine containing cells submerged in 5 mL of water, were incubated with 50 mg of selectfluor and 10 µL pyridine for 1 h and washed 3 times with PBS. For fluorescent labeling of allysine containing cells, cells were incubated with 50 µM of alexafluor-647 hydroxylamine dye for 2 h and washed 3 times with PBS. Negative control cells (unmodified) were treated with 50 µM of alexafluor-647 hydroxylamine dye for 2 h and washed 3 times with PBS. Negative control cells with only dimethyllysine residues were treated with 50 µM of alexafluor-647 hydroxylamine dye for 2 h and washed 3 times with PBS. Nuclear staining of cells (experiment and controls) was done with 2 drops of Fluoroshield-DAPI mounting media on each microscope slide and subsequently imaged on a Leica SP8 confocal microscope. The images were processed and analyzed using ImageJ software. Experiments were performed in duplicates.
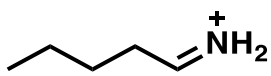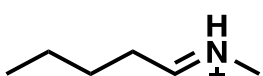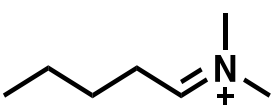


UC = Unmodified cells (No allysine), DC = Dimethyllysine cells (No allysine), AC₄ = Allysine cells (4%), AC₁₀ = Allysine cells (10%)

## XX. Cartesian coordinates

Cartesian coordinates, Thermal correction to Gibbs Free Energies (TCG), and single-point electronic energies (E) of reactants and intermediates.

**Primary amine iminium**

| C | -0.77113 | -0.76412 | 0.06871 |
|---|---|---|---|
| C | 0.49491 | 0.10895 | -0.01186 |
| H | -1.63368 | -0.13547 | 0.14439 |
| H | -0.84533 | -1.36694 | -0.8122 |
| H | -0.71405 | -1.39654 | 0.92991 |
| C | 1.73632 | -0.79583 | -0.12079 |
| H | 0.43782 | 0.74137 | -0.87307 |
| H | 0.56911 | 0.71177 | 0.86904 |
| C | 3.00237 | 0.07723 | -0.20136 |
| H | 1.79341 | -1.42826 | 0.74042 |
| H | 1.66213 | -1.39866 | -1.00169 |
| H | 2.94528 | 0.70966 | -1.06257 |
| H | 3.07657 | 0.68006 | 0.67954 |
| C | 4.24378 | -0.82755 | -0.31029 |
| N | 5.37674 | -0.43212 | 0.17287 |
| H | 5.44645 | 0.46108 | 0.6171 |
| H | 6.18286 | -1.01964 | 0.10214 |
| H | 4.1692 | -1.78327 | -0.78561 |

TCG = 0.135899
E = -252.312738

**Secondary amine iminium**

| C | -0.74756 | -0.48212 | 0. |
|---|---|---|---|
| C | 0.52473 | 0.36557 | -0.06021 |
| H | -1.6458 | 0.14074 | -0.02892 |
| H | -0.79562 | -1.17756 | -0.84388 |
| H | -0.78094 | -1.0745 | 0.91984 |
| C | 1.80445 | -0.47368 | -0.02407 |
| H | 0.51881 | 0.97253 | -0.9733 |
| H | 0.53484 | 1.07228 | 0.77802 |
| C | 3.08121 | 0.36731 | -0.08473 |
| H | 1.81044 | -1.0817 | 0.88922 |
| H | 1.79513 | -1.18041 | -0.86349 |
| H | 3.08023 | 0.97636 | -0.9957 |
| H | 3.09418 | 1.06746 | 0.7605 |
| C | 4.35129 | -0.47715 | -0.05333 |
| N | 5.55647 | 0.35265 | -0.11046 |
| H | 5.53065 | 1.01521 | 0.6593 |
| H | 4.36685 | -1.15135 | -0.91757 |
| C | 6.7868 | -0.43232 | -0.01949 |
| H | 7.6502 | 0.23577 | 0.00637 |
| H | 6.82535 | -1.08426 | 0.86859 |
| H | 6.87958 | -1.06938 | -0.90376 |

TCG = 0.161578
E = -291.613658

**Tertiary amine iminium**

| C | -0.25972 | -0.39909 | -0.35675 |
|---|---|---|---|
| C | 1.00669 | 0.4517 | -0.14707 |
| H | -1.10693 | 0.24446 | -0.47064 |
| H | -0.14569 | -0.99765 | -1.23632 |
| H | -0.40644 | -1.03523 | 0.49101 |
| C | 2.22604 | -0.47454 | 0.01685 |
| H | 1.15341 | 1.08784 | -0.99483 |
| H | 0.89265 | 1.05025 | 0.7325 |
| C | 3.49244 | 0.37626 | 0.22654 |
| H | 2.07931 | -1.11067 | 0.86461 |
| H | 2.34007 | -1.07309 | -0.86271 |
| H | 3.63917 | 1.01239 | -0.62122 |
| H | 3.37841 | 0.97481 | 1.1061 |
| C | 4.7118 | -0.54998 | 0.39046 |

```
N          5.88215 -0.15212  0.00917
H          4.59095 -1.52261  0.81973
C          7.04608 -1.03626  0.16563
H          7.75761 -0.82915 -0.60621
H          7.49686 -0.86674  1.12112
H          6.73098 -2.05643  0.09586
C          6.04817  1.18411 -0.58058
H          6.86     1.16807 -1.27742
H          5.14854  1.46455 -1.08745
H          6.25683  1.89236  0.19386
```
**TCG = 0.186653**
**E = -330.912242**

## XXI. References

1. W. C. Chan, P. D. White, Fmoc solid phase peptide synthesis: A practical approach (Oxford Univ. Press, New York, **2000**).
2. M. J. Frisch *et. al.* Gaussian 16, Revision C.01, Gaussian, Inc., Wallingford CT, **2019**.
3. A. D. Becke, *Phys. Rev. A.* **1998**, 38, 3098-3100.
4. C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B.* **1988**, 37, 785-789.
5. A. D. Becke, *J. Chem. Phys.* **1993**, 98, 1372-1377.
6. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, 132, 154104-154122.
7. A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, 123, 154101-154106.
8. A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, 122, 154104-154109.
9. E. R. Johnson, A. D. Becke, *J. Chem. Phys.* **2006**, 124, 174104-174112.
10. A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B.* **2009**, 113, 6378-6396.