

Seton Hall University eRepository @ Seton Hall

Seton Hall University Dissertations and Theses
(ETDs)


Seton Hall University Dissertations and Theses

Spring 5-19-2018

Site-Selective Modification of Peptides and Proteins via Organocatalyzed Henry Reaction

Zilma Pereira Muneeswaran
zilma.pereira@student.shu.edu

Follow this and additional works at: <https://scholarship.shu.edu/dissertations>

 Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Enzymes and Coenzymes Commons](#), [Medicinal and Pharmaceutical Chemistry Commons](#), and the [Organic Chemicals Commons](#)

Recommended Citation

Muneeswaran, Zilma Pereira, "Site-Selective Modification of Peptides and Proteins via Organocatalyzed Henry Reaction" (2018). *Seton Hall University Dissertations and Theses (ETDs)*. 2533.
<https://scholarship.shu.edu/dissertations/2533>

SITE-SELECTIVE MODIFICATION OF PEPTIDES AND PROTEINS VIA ORGANOCATALYZED HENRY REACTION

Submitted by

Zilma Pereira Muneeswaran

MASTER'S THESIS

Submitted to the Department of Chemistry and Biochemistry at Seton Hall University in partial fulfillment of the requirements for the degree of Master of Science.

May 2018

© 2018 (Zilma Pereira Muneeswaran)

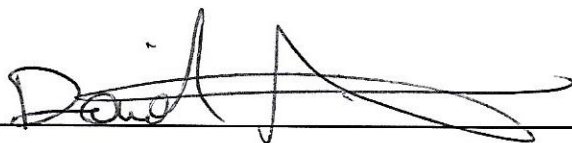
We certify that we have read this dissertation and that in our opinion it is adequate to scientific scope and qualify as a thesis for the degree of Master of Science.

APPROVED



Joseph J. Badillo, Ph. D

Research Mentor, Department of Chemistry and Biochemistry



David Sabatino, Ph. D

Member of the Dissertation Committee



Cecilia Marzabadi, Ph. D

Chair and Member of the Dissertation Committee, Department of Chemistry and Biochemistry

Acknowledgements

Throughout my graduate studies, I have received the opportunity to grow as a researcher and attain vast experience and knowledge in the process. This would not have been possible without the support of the Chemistry and Biochemistry Department, as well as, the knowledge and guidance of Dr. Monika Raj and her research group. I would like to thank my colleagues Tiauna Howard, for working with me in synthesizing the starting materials for this project, and Ryan Cohen, for his NMR and MS expertise. I would also like to acknowledge and thank Dr. Joseph Badillo for mentoring me in the final stages of my graduate studies and providing me the opportunity to learn new lab techniques. Lastly, I would like to give a special thanks to Seton Hall University and the Department of Chemistry and Biochemistry for the opportunity of achieving one of my goals as a graduate student.

Dedication

Dedicated to my mother Josefina Silva Pereira and my beloved husband Jegadheeshan Muneeswaran for encouraging me to pursue higher education. Also to my brother Victor Jose Pereira and my father Manoel Antonio Pereira for their unwavering support throughout my studies.

Abstract

In this research, peptides and protein containing serine on the N-terminus underwent site-selective modification following organocatalyzed bioconjugation that offered an additional functional group. It was shown that transforming the N-terminus serine to an aldehyde allowed site-specific bioconjugation to occur by utilizing the well-known Henry reaction. This method also grants a safer pathway for bioconjugation utilizing “green-chemistry” and biocompatible conditions. Amino acids and amino acid derived organocatalysts were utilized in the Henry reaction resulting in yields of up to 86 % conversion. Promising preliminary results were achieved in this research using peptides and myoglobin as the bioconjugation targets. Further investigation to be performed includes the analysis of the final product, as well as, applying this methodology to a number of other proteins.

Keywords

Bioconjugation, organocatalyst, Henry reaction, protein modification, site-selective.

Table of Contents

Acknowledgements.....	iv
Dedication.....	v
Abstract.....	vi
Keywords.....	vi
List of Figures and Tables.....	ix
List of Schemes.....	xii
List of Abbreviations.....	xiii
Chapter 1: Introduction.....	1
1.1 Methods for Bioconjugation.....	4
• Hydrazone and Oxime Ligation.....	4
• Pictet-Spengler.....	5
• Strain-Promoted Azide-Alkyne Cycloaddition.....	6
• Staudinger Ligation.....	7
• Native Chemical Ligation.....	8
1.2 Organocatalyst and the Henry Reaction.....	9
Chapter 2: Organocatalyzed Henry Reaction for Bioconjugation.....	10
2.1 Results and Discussion.....	10
2.2 Future Work.....	18
2.3 Instrumentation, Conditions, and Materials.....	19
• NMR.....	19
• Analytical HPLC.....	20
• HRMS.....	20
• LCMS.....	21
• LC-MS conditions.....	21
2.4 Experimental Procedures.....	23
• Synthesizing Phenylalanine Aldehyde.....	23
• Phenylalanine Aldehyde Henry Reaction.....	24
• Synthesizing Peptide Aldehydes.....	24
• Peptide Aldehyde Henry Reaction.....	24
• Fmoc Solid-Phase Peptide Synthesis.....	25
• Converting N-terminus Glycine to Aldehyde in Proteins.....	25

Chapter 3: Supporting Information.....	26
3.1 ^1H NMR, ^{13}C NMR, and MS.....	26
3.2 HPLC and LCMS Traces.....	30
References	36

List of Figures and Tables

Figure 1: Synthesis of oxime or hydrazone from carbonyl reacted with alkoxyamine or hydrazine	4
Figure 2: a) The traditional Pictet-Spengler ligation. b) The aminoxy Pictet-Spengler ligation ...	5
Figure 3: Two pathways for strain-promoted cycloaddition conjugation	6
Figure 4: Two pathways for the Staudinger Ligation	7
Figure 5: Native Chemical Ligation reaction	8
Figure 6: MS of Myoglobin aldehyde	15
Figure 7: MS of Myoglobin aldehyde	16
Figure 8: Myoglobin aldehyde reaction with nitromethane MS	17
Figure S1: ¹ H NMR of formylphenoxy NHS ester	26
Figure S2: ¹³ C NMR of formylphenoxy NHS ester	27
Figure S3: ¹ H NMR of phenylalanine aldehyde	28
Figure S4: ¹³ C NMR of phenylalanine aldehyde	29
Figure S5: a) HPLC and b) LCMS traces of phenylaldehyde aldehyde	30
Figure S6: a) HPLC of phenylalanine reaction with nitromethane using proline as catalyst; b) LCMS traces of crude phenylalanine reaction with nitromethane using proline as catalyst; c) LCMS of 17.7 min peak; d) LCMS of 20 min peak; 24h reaction	30

Figure S7: a) HPLC of phenylalanine reaction with nitromethane control reaction; b) LCMS traces of crude control reaction; 6h reaction	31
Figure S8: a) HPLC of phenylalanine reaction with nitromethane using DIEA as catalyst; b) LCMS of 17.7 min peak; c) LCMS of 20 min peak; 6h reaction	31
Figure S9: a) HPLC of phenylalanine reaction with nitromethane using N,N-Diphenylthiourea as catalyst; b) LCMS of 17.7 min peak; c) LCMS of 20 min peak; 6h reaction	32
Figure S10: a) HPLC of phenylalanine reaction with nitromethane using glycine methyl ester as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6h reaction	32
Figure S11: a) HPLC of phenylalanine reaction with nitromethane using threonine as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6h reaction	33
Figure S12: a) HPLC of phenylalanine reaction with nitromethane using glycine as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6h reaction	33
Figure S13: HPLC of phenylalanine reaction with nitromethane using glycine as catalyst under different pH conditions. a) reaction under acidic condition, pH 3.5 b) reaction under biological condition, pH 7.5; c) reaction under basic condition, pH 10.5	34
Figure S14: HPLC of SVF. a) SVF peptide b) SVF peptide aldehyde c) SVF product of nitromethane reaction using glycine as catalyst; 30 minutes	34
Figure S15: LCMS of SVF. a) SVF peptide; b) SVF peptide aldehyde; c) SVF product of nitromethane reaction using glycine as catalyst	35

Figure S16: LCMS of SAF. a) SAF peptide; b) SAF peptide aldehyde; c) SVF

product of nitromethane reaction using glycine as catalyst 35

Table 1: Percent conversion calculated for Michael addition product as determined by HPLC 12

List of Schemes

Scheme 1: General reaction for synthesizing the phenylalanine aldehyde moiety	10
Scheme 2: Proposed Henry reaction using organocatalysis	11
Scheme 3: a) Condensation and b) Michael addition mechanism	11
Scheme 4: Proposed reaction mechanism	12
Scheme 5: Modification of peptide to peptide aldehyde through sodium periodate oxidation	14
Scheme 6: N-terminus glycine modification of protein to form aldehyde	15
Scheme 7: Organocatalyzed Henry reaction of myoglobin aldehyde with nitromethane	16

List of Abbreviations

ACN	Acetonitrile
Ala	Alanine
CD	Circular dichroism
CID	Collision induced dissociation
DCE	Dichloroethane
DIC	N,N'-Diisopropylcarbodiimide
DIEA	N,N,-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	N,N-Dimethylformamide
DMSO	Dimethyl sulfoxide
DSC	Disuccinimidyl carbonate
ESI	Electrospray ionization
FA	Formic acid
HBTU	N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HOAt	1-Hydroxy-7-azabenzotriazole
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
LCMS	Liquid chromatography - mass spectrometry
MeOH	Methanol
MS	Mass spectrometry
NaIO ₄	Sodium periodate

NHS	N-Hydroxysuccinimide
NMR	Nuclear magnetic resonance
Phe	Phenylalanine
PLP	Pyridoxal 5'-phosphate
RPM	Rotation per minute
Ser	Serine
SPPS	Solid phase peptide synthesis
TES	Triethylsilane
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TOF/MS	Time-of-flight mass spectrometry
Val	Valine

Chapter 1: Introduction

Proteins are the backbone for many cellular functions as they take an array of roles including building our immune system¹, carrying out an array of reactions, sending messages from cell-to-cell and transporting materials in-and-out of cells^{2,3}, as well as build the structural components in cells⁴. When proteins begin to miscommunicate, or are damaged for any reason, diseases occur and many detrimental diseases are caused by unwanted protein-protein interactions. Researchers are constantly studying causes for diseases and developing new methods to deliver drugs are crucial for successful treatment.

Often, peptides are used to replicate a specific region of a protein that is known to be the cause of the undesired interaction. Like proteins, peptides are composed of amino acids; however, proteins are much larger molecules with larger complex structures. Researchers often look for the active site of a protein, the region which interacts with another protein or is responsible for the function of interest, and model peptide analog that may interact with the active site. The modeled peptide can then be modified in many ways to study the interactions. Frequently fluorescent tags are attached to the peptides in order to monitor interactions.

Bioconjugation is a very useful technology for labeling of biomolecules such as proteins, DNA, and carbohydrates with fluorescent probes, affinity tags, or isotope labels.⁵⁻¹⁰ The labeled biomolecules can then be used to study the function and to track the activity of a biomolecule. The synthesis of these constructs requires chemoselective and site-selective bioconjugation reactions that allow the formation of a covalent bond between two unique and chemically stable groups under mild and aqueous conditions.

Various techniques are known for bioconjugation but the most widely used methods are the Cu-catalyzed click reaction of azides and alkynes¹¹, the hydrazone and oxime ligations,¹² the Staudinger ligation,¹³ the Pictet-Spengler ligation,¹⁴ and native chemical ligation.¹⁵ Limitations of these methods include the presence of toxic reagents and relatively slow reaction rates. Most importantly, all of the chemoselective and site-selective bioconjugation methods that currently exist use a stoichiometric amount of the chemical reagents. The need of a stoichiometric reagents not only adds to expense to the method but also to the amount of chemical waste. In addition, existing methods can lead to various byproducts and side reactions that add to the complexity of the bioconjugation. A desired alternative in the field of chemical bioconjugation is a catalytic approach that eliminates the need for stoichiometric amounts of chemical reagents.

Here we describe a new approach for site-selective bioconjugation, which will require the use of a catalyst. Since biomolecules are involved, we propose to use non-toxic organocatalysts for carrying out the bioconjugation rather than toxic metal catalysts. Besides the nontoxic nature of organocatalysts there are also several other benefits over metal catalysts. The organocatalysts have the potential for savings in cost, time and energy, are operationally simple, and reduce chemical waste.¹⁶ Herein, we propose the use of organocatalysts for catalyzing a chemoselective and site-selective bioconjugation reaction by undergoing a Henry reaction. This reaction occurs between an aldehyde, or ketone, and a nitroalkane producing a nitroaldol compound. We chose to exploit this reaction as there has been many techniques developed to obtain a carbonyl functional group on biomolecules; however, there has yet to be any report on the Henry reaction being applied to bioconjugation. The nitroaldol product will provide a hydroxyl and nitro functional groups which can be exploited for conjugating tags, drugs, and probes to the biomolecule. This reaction will enable selective conjugations of large molecules using an organocatalyst, opening an avenue

to the synthesis of much larger, structurally defined molecules including multi-domain proteins, protein–polymer conjugates,¹⁷ and oligomeric biomolecules.¹⁸

1.1 Methods for Bioconjugation

- **Hydrazone and Oxime Ligation**

As previously mentioned, there are several pathways in which to achieve bioconjugation. Oximes and hydrazones have become of great interest for the achievement of bioconjugation as the reaction can be performed under physiological conditions with water as a byproduct when reacting a carbonyl containing moiety with an alkoxyamine or hydrazine (Figure 1). Due to their facile ligation hydrazones and oximes have been utilized to form new probes and sensors,^{19,20} provide isotopic labeling of biomolecules,²¹ and to synthesize heterocycles.²²⁻²⁴ In recent years, these reactions have also been fine-tuned to obtain faster reaction rates by incorporating catalysts and performing these reaction in aqueous media to better suit biological applications.²⁵

Although extensive research has improved the utility of these bioconjugates, the hydrolysis of hydrazones and oximes in certain biological medium has been reported, with the former also being hydrolyzed in cell cultures and plasma.^{26,27} The use of catalysts has proven to make reactions more efficient, but there still remains concern about the toxicity of some of the catalysts.²⁸ Therefore, there remains a need for the development of a method which can be performed under biological conditions while also being able to withstand degradation by hydrolysis.

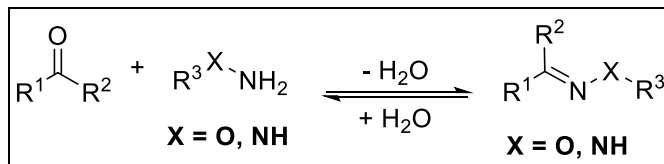


Figure 1: Synthesis of oxime or hydrazone from carbonyl reacted with alkoxyamine or hydrazine.

- **Pictet-Spengler**

In order to combat the shortcomings of the hydrazone and oxime ligation, researchers have developed new methodologies using modified Pictet-Spengler reactions.²⁹ Many natural products have been formed using the Pictet-Spengler reaction and researchers have chosen to exploit this methodology in order to facilitate bioconjugation under physiological conditions. When the traditional Pictet-Spengler reaction was performed the reaction proceeded slowly and required acidic conditions which are not suitable for biological studies. By adding an aminoxy to the tryptamine the reaction rate increased, except acidic conditions were still necessary (Figure 2b).³⁰ Researchers continue to explore this pathway for bioconjugation by applying this method to antibody drug conjugates (ADCs), to achieve target specific drug delivery, and an array of protein and peptides moieties.

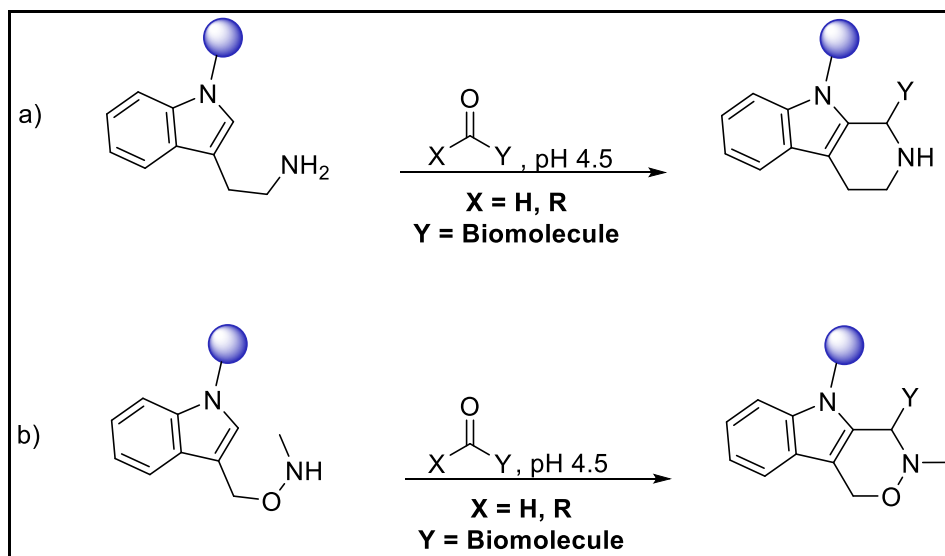


Figure 2: a) The traditional Pictet-Spengler ligation. b) The aminoxy Pictet-Spengler ligation

- **Strain-Promoted Azide-Alkyne Cycloaddition**

When designing a bioconjugation reaction method, researchers often think of means to reduce toxicity by eliminating the need for metal catalysts, or utilizing organocatalysts which produce more suitable for biorthogonal reactions. For these reasons, many biochemists have turned away from the canonical copper-catalyzed azide-alkyne cycloaddition reaction (CuAAC) and have moved towards strain-promoted cycloadditions. These include the strain-promoted azide-alkyne cycloaddition (SPAAC) and strain-promoted alkyne-nitrone cycloaddition reactions (SPANC) (Figure 3).³¹ The advantages of utilizing SPANC over SPAAC is that the former method allows for two R-functional groups in the nitrone comparison to the azide having only one R-group. Additionally, SPANC has shown faster reaction rates and biocompatibility than SPAAC.³² These two reactions allow for bioconjugation to occur without the need for a metal catalyst, thus reducing cytotoxicity. One drawback from these two methods is slow rate of reaction due to the poor reactivity of cyclooctyne,³³ and much effort is still being invested into improving the reactivity of these strained carbocycles. Additionally, the bulkiness of the cyclooctyne can cause perturbation on protein structures making the protein inactive.

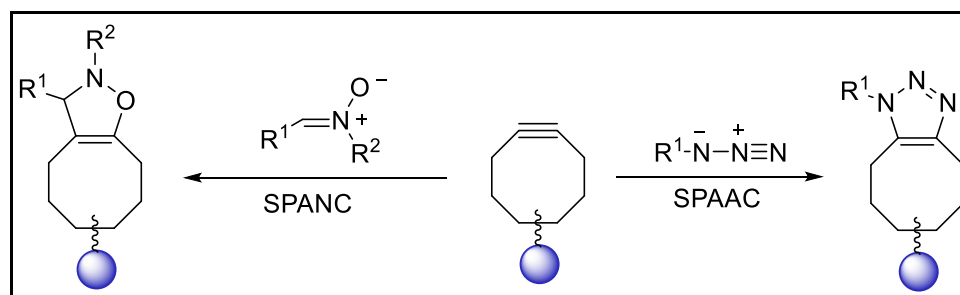


Figure 3: Two pathways for strain-promoted cycloaddition conjugation.

- **Staudinger Ligation**

The canonical Staudinger Ligation is an organic chemical reaction between an azide and a phosphine. Recently, Bertozzi and co-workers chose this reaction to perform bioconjugation due to the bio-orthogonal nature of both starting materials and it has since been applied to an array of biopolymers.³⁴ A modified approach to the Staudinger ligation has been developed by Bertozzi and co-workers which allows for a more direct approach to bioconjugation in which phosphine oxide is a byproduct; the reaction is termed “traceless” due to the amide bond having no residual atoms (Figure 4).³⁵ Overall, both methods are highly used in the synthesis of modified biomolecules in applications such as protein engineering and nucleic acid labeling.³⁶ Nonetheless, both of these approaches require the incorporation of an azide in the biomolecule, which is a synthetic challenge.

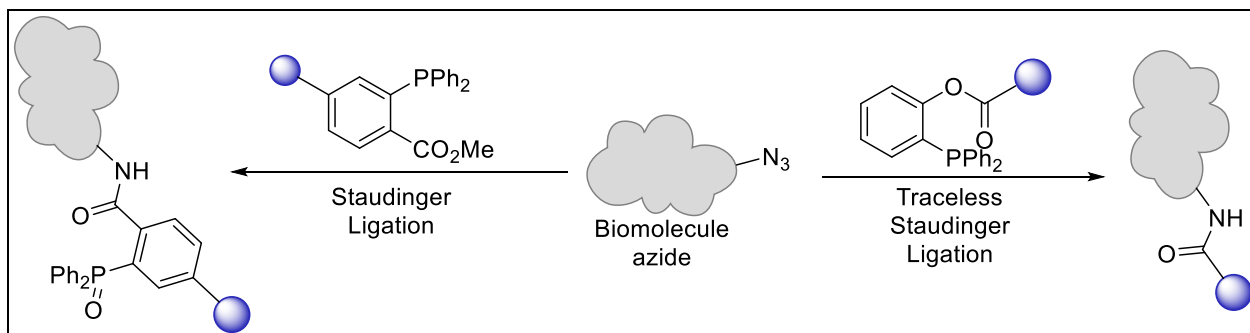


Figure 4: Two pathways for the Staudinger Ligation.

- **Native Chemical Ligation**

The requirements for a Native Chemical Ligation (NCL) are that the biomolecule contain a free N-terminus cysteine residue and the ligand contain a thioester. The reaction begins with a thioester exchange between the two moieties which results in thiol being removed. It then undergoes a transthioesterification step resulting in a cysteine residue being present at the ligation site (Figure 5).¹⁵ This method produces high yields, is irreversible, and occurs in aqueous media. The major drawback from this method is the requirement of a free N-terminus cysteine residue. Cysteine is usually found participating in disulfide bridges in proteins and is also one of the least abundant amino acids. Thus resulting in limited applicability with this method.

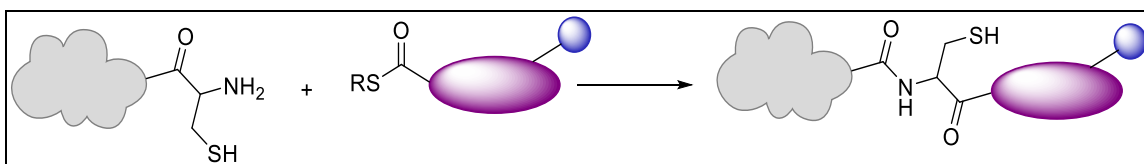


Figure 5: Native Chemical Ligation reaction.

1.2 Organocatalyst and the Henry Reaction

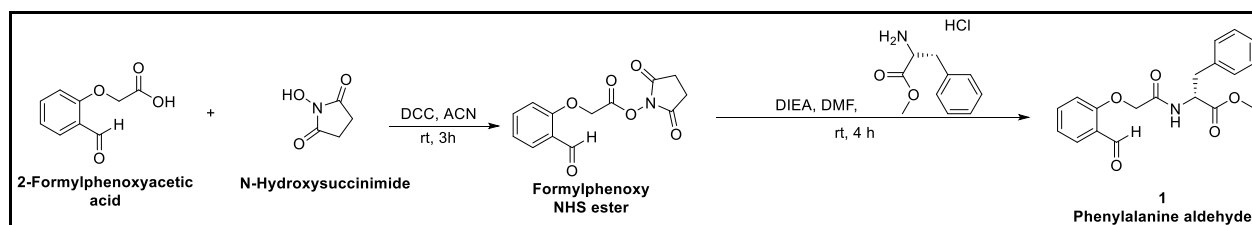
The concept of organocatalysis has quickly gained popularity in synthetic biology due to its “green-chemistry” potential as it replaces the need for using metal catalysts.³⁷ Organocatalyst are organic molecules that do not possess metals and are capable of increasing the rate of a chemical reaction. This reasoning also plays a role on its importance in biological reactions due to the toxic effects of some metals. One synthetic reaction that has successfully been accomplished through the use of organocatalysis is the Henry reaction. It has been reported that through the use of enantiomerically pure guanidines, reactions have produced products with up to 95 % of enantiomeric excess.³⁸ This report became interesting as this method has yet to be applied to biological reactions.

The Henry reaction requires a nitroalkane reacting with either an aldehyde or a ketone yielding a nitroaldol product. There have been many reports on how to convert the N-terminus residues of peptides or proteins into aldehydes or ketones in order to undergo bioconjugation; however, this method has yet to be applied for attaining bioconjugation through the Henry reaction. Thus, we felt the need to explore this pathway as its success would be of great use in the fields of medicine and pharmaceuticals. The nitroaldol product would result in the biomolecule containing two possible functional handles which can be reacted to conjugate with other biomolecules, drugs, tags, or peptides. Furthermore, if this method is applied for antibody drug conjugation (ADC), the antibody could possibly carry a drug while carrying a tag or marker which researchers can track to ensure target-specific delivery.

Chapter 2: Organocatalyzed Henry Reaction for Bioconjugation

2.1 Results and Discussion

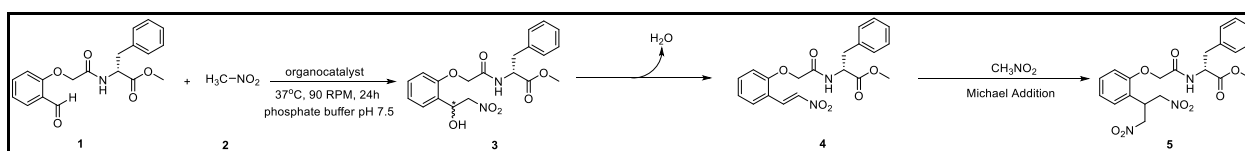
To begin this investigation, we designed an amino acid aldehyde as the model aldehyde for optimizing the Henry reaction. We chose to synthesize a phenylalanine aldehyde **1** to undergo preliminary reactions and monitored the reaction using HPLC (Scheme 1). The procedure for synthesizing phenylalanine aldehyde can be found in Section 2.3.



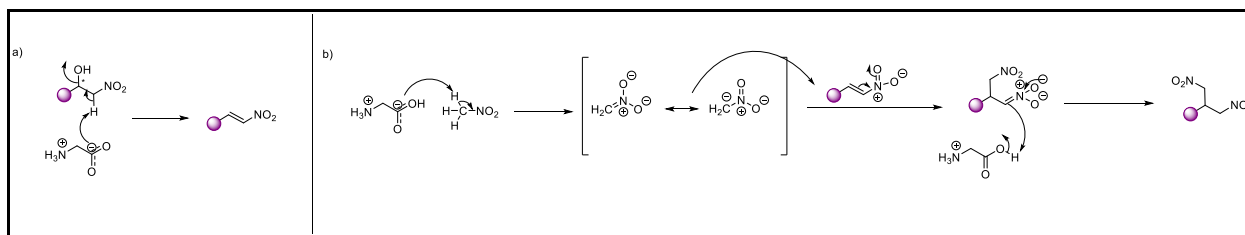
Scheme 1: General reaction for synthesizing the phenylalanine aldehyde moiety.

The first reaction was set using the phenylalanine aldehyde and nitromethane with proline as the catalyst using 100mM phosphate buffer pH 7.5 as solvent (Scheme 2). Proline was chosen as the initial organocatalyst for this reaction due to its wide application as an organocatalyst in the literature along with its availability in our lab. Nitromethane was chosen as the model nitroalkane for the Henry reaction due to its simple structure and because it would provide sufficient information as to whether this methodology is feasible. The reaction was set for 24 h and was analyzed using HPLC and LCMS. Although the reaction yielded product, the reaction was not complete as there still remained unreacted phenylalanine aldehyde in the reaction mixture. Upon analyzing the fractions by LCMS, it was noted that product, **3**, underwent condensation to **4**

followed by a Michael addition of a second nitromethane molecule **5** resulting in 52% conversion of the Michael addition (Scheme 3). It was difficult to account for how much starting material, **1**, Henry reaction product, **3**, and condensation product, **4**, remained in the reaction as these three coeluded on the HPLC (Supporting Information Figure S6). Therefore, percent conversions were calculated and reported only for the Michael addition product.



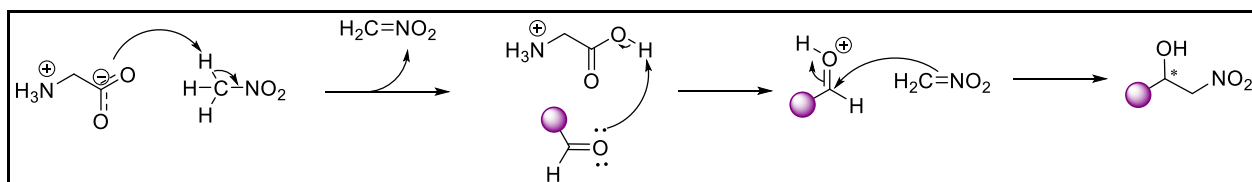
Scheme 2: Proposed Henry reaction using organocatalysis.



Scheme 3: a) Condensation and b) Michael addition mechanism.

Upon attaining these results different catalysts readily available in our lab were screened to achieve higher yields (Table 1). All reactions were completed with 100 mM phosphate buffer, and were incubated at 37 °C and with 90 RPM. Reactions were monitored after 6 h by HPLC and LCMS (see Supporting Information). As a control, two reactions were set, one without the use of any catalyst and after 6 h only 5 % conversion of product was detected. For the second control reaction we chose to use DIEA (N,N-diisopropylethylamine), in place of a catalyst since the Henry reaction is a base driven reaction, however only 32 % conversion to product, **5**, was observed as well as another product which was unidentified (Supporting Figure S8). Although the Henry

reaction is a base driven reaction, under our reaction conditions poor conversion was acquired when using a base. We then chose to use N,N-Diphenylthiourea as an organocatalyst after seeing organic synthesis reports of thiourea being utilized to catalyze the Henry reaction.³⁸ Nonetheless, we only observed a 10 % conversion. Next, we sought to utilize amino acids and amino acid derivatives as organocatalysts due to their availability in our lab and compare them to the results previously observed with proline. Glycine methyl ester, proline, and threonine resulted in low to moderate conversions (Table 1). Although threonine showed promising results, 80 % conversion, it also produced undesired products which were not identifiable. When glycine was used we observed 86 % conversion with the least amount of side product and was thus chosen as the catalyst for the remainder of the research project (Scheme 4).



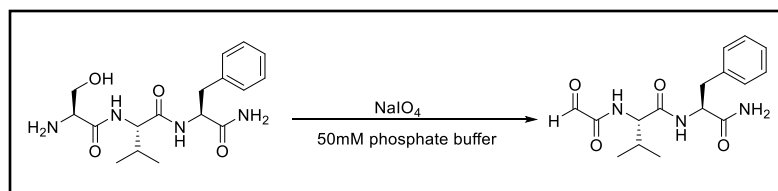
Scheme 4: Proposed reaction mechanism.

Table 1: Percent conversion calculated for Michael addition product as determined by HPLC.

Entry	Catalyst	% Conversion
1	None	5
2	Proline	44
3	Glycine	86
4	Threonine	80
5	DIEA	32
6	Glycine methyl ester	22
7	N,N,-Diphenylthiourea	10

Following the catalytic studies we determined that in order to deem this reaction suitable for biological systems it was crucial to perform pH studies on the reaction system. The reaction was subsequently repeated using glycine as the catalyst and 100 mM phosphate buffers at pH 4.5, 7.5, and 10.5. We found that the reaction occurred as previously observed under acidic and neutral conditions; however, when the reaction was set under basic conditions the reaction produced unforeseen side products (Supporting Figure S13). With these results we were satisfied to proceed with peptide studies and eventually protein models using glycine as the catalyst (7.5×10^{-4} mmol) using phosphate buffer pH 7.5 as the solvent and nitromethane as the nitroalkane for the Henry reaction.

After determining the catalyst and reaction pH, two tripeptides, Ser-Ala-Phe and Ser-Val-Phe, were synthesized through the well-known solid-phase-peptide-synthesis method (SPPS).³⁹ Upon synthesizing, cleaving, and purifying the peptides (see Section 2.3 for procedure) the N-terminus serine was modified to yield an aldehyde through sodium periodate oxidation on each of the peptides (Scheme 5).⁴⁰ Peptide aldehydes were purified by HPLC and analyzed through MS (see Supporting Information). Peptide Ser- Ala- Phe and Ser-Val-Phe aldehyde were subsequently reacted with nitromethane using glycine as the organocatalyst for the reactions. Surprisingly, the peptide aldehydes were consumed in 30 min with > 99 % conversion yield to a mixture of Henry product, condensation product, and Michal addition product eluding in the same peak on the HPLC (Supporting Figure S14). After 6 h the reaction still contained a mixture of products. We attribute the fast consumption of aldehyde in peptides as compared to the phenylalanine aldehyde adduct used previously to steric. The peptides were linear and the aldehyde was able to freely react; whereas in the phenylalanine aldehyde adduct, the aldehyde had the bulk of the molecule in ortho-position in the phenyl ring.



Scheme 5: Modification of peptide to peptide aldehyde through sodium periodate oxidation.

There have been reports of proteins with the N-terminus glycine being converted into aldehydes through the use of pyridoxal 5'-phosphate (PLP).⁴¹ Myoglobin is a commercially available protein which contains an N-terminus glycine and we had previously successfully converted the N-terminal glycine to an aldehyde; therefore, we chose to apply our preliminary findings to myoglobin. The MS of myoglobin was taken prior to experimental procedures (Figure 6) and myoglobin was subsequently modified to contain an N-terminus aldehyde (Scheme 6, Figure 7) and reacted with nitromethane using a glycine catalyst in 100mM phosphate buffer pH 7.5 for 3 h at 37 °C stirring at 90 RPM (Scheme 7). The reaction was then analyzed by MS (Figure 8). Although the MS of the myoglobin Henry reaction show background interference product was still observed. Further work and analysis are still required for the completion of this work; however, the preliminary work reported here demonstrates the possibility of applying this methodology for bioconjugation in the future.

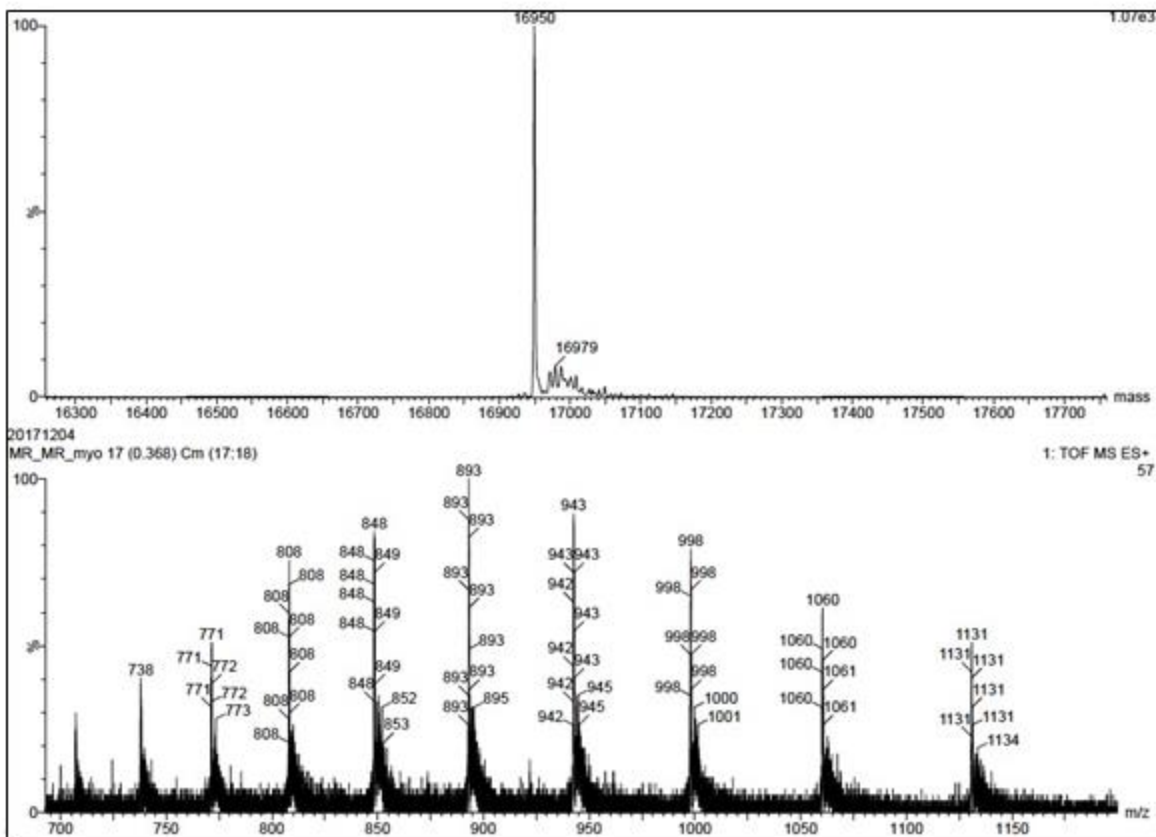
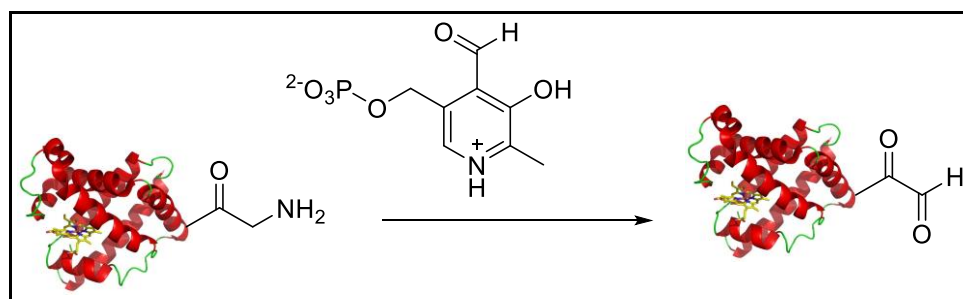


Figure 6: MS of unmodified myoglobin.



Scheme 6: N-terminus glycine modification of protein to form aldehyde.

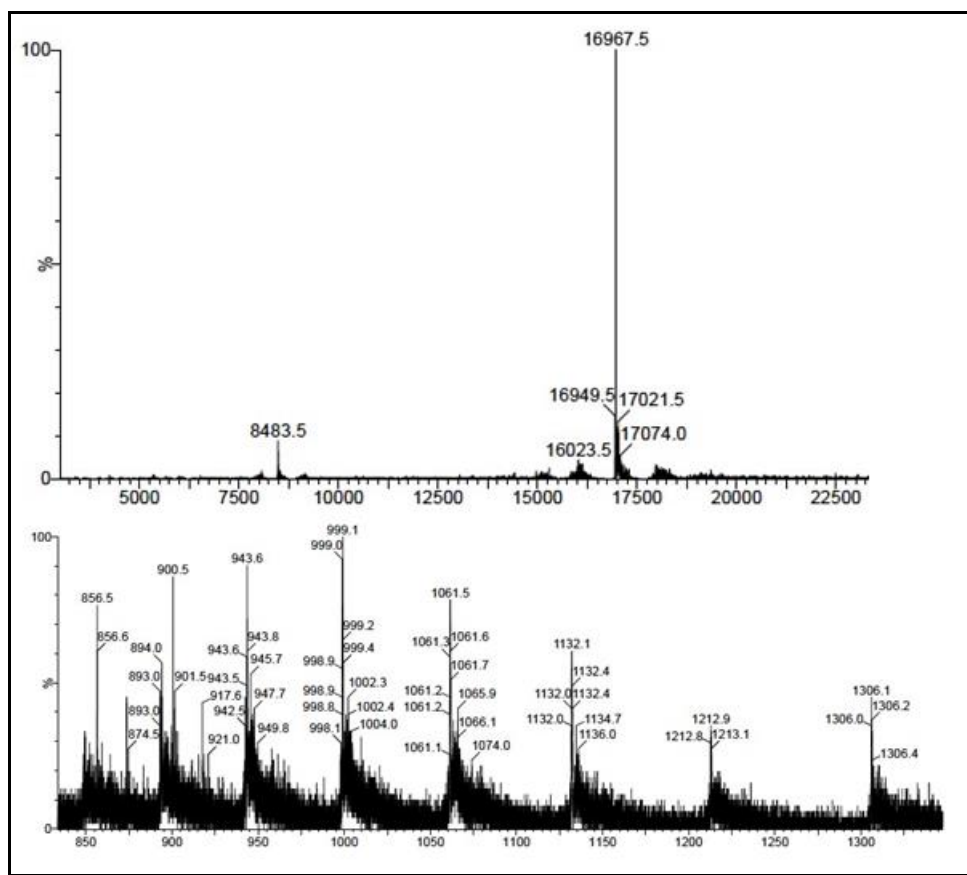
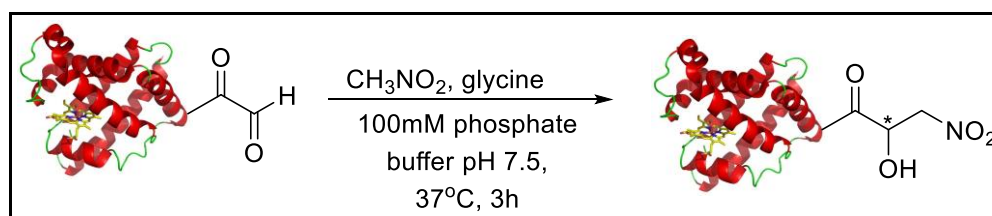


Figure 7: MS of myoglobin aldehyde.



Scheme 7: Organocatalyzed Henry reaction of myoglobin aldehyde with nitromethane.

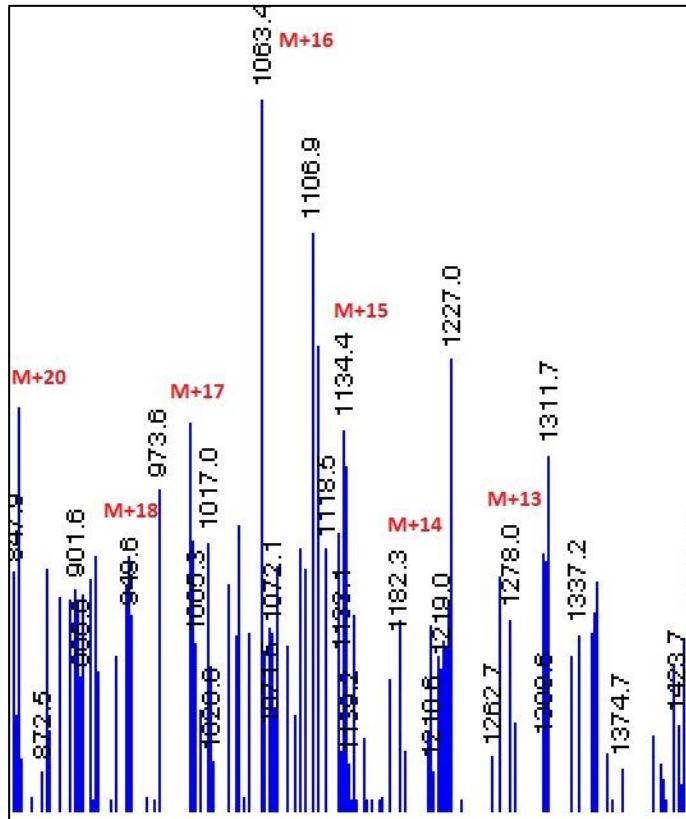


Figure 8: Myoglobin aldehyde reaction with nitromethane MS.

2.2 Future Work

Overall, this work highlights the potential of applying the well-known Henry reaction to bioconjugation using simple amino acids as catalyst. This method resulted in yields of up to 86 % conversion in mild conditions which is suitable for biological systems. Additionally, by using the Henry reaction the biomolecule is equipped with new functional handles which can be exploited for conjugating a drug or tag.

Further investigation is needed to characterize the final products and to determine the stereo-selectivity of the organocatalyst by using ^1H and ^{13}C NMR for the phenylalanine aldehyde Henry product **3**, **4**, and **5**, as well as the peptides moieties. Myoglobin, myoglobin aldehyde, and the final myoglobin Henry product needs to be characterized for its secondary structure through the use of CD-spectrometry to demonstrate that the protein will still remain active after modification. Lastly, we hope to be able to apply this methodology to antibody-drug-conjugates (ADCs) in order to facilitate target drug deliveries.

2.3 Instrumentation, Conditions, and Materials.

Fmoc-amino acids were obtained from Nova Biochem (EMD Millipore Corporation) (Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide was obtained from ChemPep Inc. (Wellington, Florida). N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU) was obtained from CreoSalus (Louisville, Kentucky). N,N'-Disuccinimidyl carbonate (DSC) was obtained from Nova Biochem, under (EMD Millipore Corporation) (Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP): Merck KGaA (Darmstadt, Germany). N,N-Dimethylformamide (DMF): Macron Fine Chemicals (Center Valley, Pennsylvania). Dichloroethane (DCE), acetonitrile, N,N-Diisopropylethylamine (DIEA), N,N'-diisopropylcarbodiimide (DIC), 1-Hydroxy-7-azabenzotriazole (HOAt), were purchased from (EMD Millipore Corporation)(Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR 100 Matsonford Road Radnor, PA. Diethyl Ether: Sigma Aldrich (St. Louis, Missouri). Water was purified using a Millipore Milli-Q water purification system.

All commercial materials (Aldrich, Fluka, Nova) were used without further purification. All solvents were reagent grade or HPLC grade (Fisher). Anhydrous THF, diethyl ether, dichloromethane, and DMF were obtained from Aldrich and Fluka.

- **NMR.**

Proton NMR spectra were acquired at 25 °C in DMSO-*d*₆ using an Agilent DD2 (600 MHz) spectrometer and carbon NMR spectra on a 151 MHz equipped with a 3 mm He HCN cryoprobe from Merck. All NMR chemical shifts (δ) are referenced in ppm relative to residual solvent or internal tetramethylsilane. ¹H NMR chemical shifts referenced to residual DMSO-*d*₅ at 2.50 ppm, and ¹³C NMR chemical shifts referenced to

DMSO- d_6 at 39.52 ppm. Carbon NMR spectra are proton decoupled. NMR spectral data are reported as chemical shift (multiplicity, coupling constants (J), integration). Multiplicity is reported as follows: singlet (s), broad singlet (bs), doublet (d), doublet of doubles (dd), doublet of triplet (td), triplet (t) and multiplet (m). Coupling constant (J) in Hertz (Hz).

- **Analytical HPLC.**

Analytical HPLC chromatography (HPLC) was performed on an Agilent 1100 series HPLC equipped with a 4.6 mm C-18 reversed-phase column. All separations involved mobile phase of 0.1 % FA (v/v) in water (solvent A) and 0.1 % FA (v/v) in acetonitrile (solvent B). Peptide compositions were evaluated by analytical reverse phase HPLC using a gradient of 0.1 % FA in acetonitrile versus 0.1 % FA in water. Analytical HPLC method using a linear gradient of 0 – 80 % 0.1 % FA (v/v) acetonitrile in 0.1 % aqueous FA over 30 min at room temperature with a flow rate of 1.0 mL min⁻¹. The eluent was monitored by absorbance at 220 nm unless otherwise noted.

- **HRMS**

HRMS data were acquired on a Q-ToF mass spectrometer using positive polarity electrospray ionization (+ESI). Tandem MS experiments were performed using collision induced dissociation (CID) with N₂ as the collision gas.

- **LCMS**

Mass spectrometry was performed using ultra high performance liquid chromatography-mass spectrometry using the Agilent 1100 Series LCMSD VL MS Spectrometer.

Instrument: Agilent 1290 UHPLC with 6520 Q-ToF | Mass Range: m/z 150 to 3200 |

Ionization: (+) ESI | Voltages: capillary = 3.5 kV, fragmentor = 175 V, skimmer = 65 V |

Drying Gas: N₂ at 300 °C | Nebulizer: N₂ at 35 L/min | Acquisition Rate: 1 spectra/s |

Flow Rate: 0.1 mL/min | Injection Volume: 15 µL | Mobile Phase: A = 0.1% formic acid

in water; B = acetonitrile | Gradient: 15 % B to 50 % B in 55 minutes, followed by 5 min

re-equilibration time | Column: Acquity UPLC Peptide BEH C18, 300 Å, 1.7 mm, 1.0 X

150 mm | Column Temp.: 50 °C | Deconvolution using maximum entropy algorithm in

Masshunter software.

- **LC-MS conditions**

Analysis was performed on an ultra-performance 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). 5 µL of the sample was injected onto a C4 column (Phenomenex Aeris™ 3.6 µm WIDEPOR C4 200 Å, LC Column 50 x 2.1 mm) with a 200 µL/min flow rate of mobile phase of 95 % H₂O, 5 % acetonitrile and 0.1 % formic acid (solution A) and 95 % acetonitrile, 5 % H₂O, and 0.1 % formic acid (solution B) beginning gradient time - 0 min 10 % B; 5 min 28 % B; 20 min 38 % B; 22 min 100 % B. Capillary voltage was set at 3.0 kV, the sample cone voltage was 50 V, and the extraction cone was 4.0 V. The source and desolvation temperature were maintained at 95 and 300 °C, respectively, with the

desolvation gas flow at 400 L/h. The Time of Flight Mass Spectrometry (TOF/MS) scan was 1 s long from 100 to 1100 m/z with a 0.02 s inter-scan delay using extended dynamic range acquisition and centroid data format. The lock mass was used to correct instrument accuracy with a 0.5 μ M solution of HP 1221 (Agilent part number G196985003). Ion source parameters such as the source temperature (gas and sample cone), mobile phase flow rate, and cone voltage were fixed throughout the study.

2.4 Experimental Procedures

- **Synthesizing Phenylalanine Aldehyde**

To prepare phenylalanine aldehyde, formylphenoxy acetic acid (2.1 mmol) was dissolved in 3 mL of acetonitrile (3 mL) along with DCC (2.4 mmol) and N-hydroxy succinimide (2.3 mmol) in a round bottom flask with a stir bar and left stirring for 3 h. Reaction was filtered with vacuum to remove urea formed within the reaction. The solvent was removed and reaction was dried until formyl phenoxy NHS ester solid was recovered. The crude solid was recrystallized using THF; upon crystallization the crystals were then dried overnight under vacuum. ^1H and ^{13}C NMR analysis confirmed formation of formyl phenoxy NHS ester. ^1H NMR (600 MHz, DMSO- d_6) δ 10.42 (s, 1H), 7.74 (dd, $J = 7.7$, 1.8 Hz, 1H), 7.69 (ddd, $J = 8.9$, 7.3, 1.9 Hz, 1H), 7.27 (d, $J = 8.4$ Hz, 1H), 7.18 (t, $J = 7.5$ Hz, 1H), 5.55 (s, 2H), 2.83 (s, 4H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 188.8, 169.9, 165.2, 159.3, 136.2, 127.9, 124.8, 122.2, 114.1, 63.7, 25.5.

Formyl phenoxy NHS ester (0.65 mmol) was then dissolved by anhydrous DMF (2 mL) and DIEA (280 μL) in round bottom flask along with L-Phenylalanine methyl ester and stirred at room temperature for 4 h. Once again solvent was removed by vacuum. Column chromatography was utilized to purify phenylalanine aldehyde product and was analyzed by MS and ^1H and ^{13}C NMR. ^1H NMR (600 MHz, DMSO- d_6) δ 10.40 (s, 1H), 8.59 (d, $J = 8.1$ Hz, 1H), 7.74 (dd, $J = 7.7$, 1.9 Hz, 1H), 7.59 (ddd, $J = 8.8$, 7.3, 1.9 Hz, 1H), 7.32 – 7.23 (m, 2H), 7.23 – 7.18 (om, 3H), 7.12 (t, $J = 7.5$ Hz, 1H), 6.98 (d, $J = 8.4$ Hz, 1H), 4.70 (d, $J = 14.9$ Hz, 1H), 4.63 (d, $J = 14.9$ Hz, 1H), 4.63 – 4.53 (om, 1H), 3.64 (s, 3H), 3.13 (dd, $J = 13.8$, 5.1 Hz, 1H), 3.00 (dd, $J = 13.8$, 9.8 Hz, 1H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 189.7, 171.6, 167.4, 159.7, 137.1, 136.1, 129.1, 128.2, 128.1, 126.6, 124.4, 121.4, 113.6, 67.0, 53.2, 52.0, 36.3. HRMS (+ESI) $\text{C}_{19}\text{H}_{20}\text{NO}_5^+$: 342.1336, found 341.9.

- **Phenylalanine Aldehyde Henry Reaction**

For the organocatalyzed Henry reaction, phenylalanine aldehyde (0.003 mmol) was dissolved with 100 μ L of 50:50 mixture of DMSO and 100 mM phosphate buffer (pH 7.5) in a dry and clean vial. After dissolving the phenylalanine aldehyde, an additional 400 μ L of 100 mM phosphate buffer (pH 7.5) was added to the solution. A stock solution of the catalyst was prepared to accurately deliver the catalyst (7.5×10^{-4} mmol). The catalyst was pipetted to the dissolved phenylalanine aldehyde solution. Lastly, nitromethane (0.036 mmol) was added to the reaction mixture and the vial was placed in an incubator at 37 $^{\circ}$ C and set to 90 RPM.

- **Synthesizing Peptide Aldehydes**

Peptides with serine at the N-terminus (0.014 mmol) were dissolved with 1200 μ L of 50 mM phosphate buffer (pH 7.5). Sodium periodate (0.071 mmol) was added to the solution and stirred at room temperature for 3 h. Unreacted sodium periodate was filtered out of the solution and peptide aldehydes were purified by HPLC and mass confirmed through LCMS (Figure S14 and S15).

- **Peptide Aldehyde Henry Reaction**

Peptide aldehyde (0.003 mmol) was dissolved in 500 μ L of 100 mM phosphate buffer (pH 7.5). A stock solution of the catalyst was prepared to accurately deliver the catalyst (7.5×10^{-4} mmol). The catalyst was pipetted to the dissolved peptide aldehyde solution. Lastly, nitromethane (0.036 mmol) was added to the reaction mixture and the vial was placed in an incubator at 37 $^{\circ}$ C and set to 90 RPM. Reaction was purified by HPLC and mass confirmed through LCMS (Figure S14 and S15).

- **Fmoc Solid-Phase Peptide Synthesis**

Peptides were synthesized manually on a 0.25 mmol scale using Rink amide resin. Fmoc-group was deprotected with 20% piperidine-DMF for 20 min to obtain a deprotected peptide-resin. Fmoc-protected amino acids (1.25 mmol) were coupled on the resin using a HBTU (1.25 mmol) and DIEA (1.25 mmol) for 2 h at room temperature. Peptides were synthesized using standard protocols and were cleaved from the resin using a cocktail of trifluoroacetic acid: triisopropyl silane: water (95: 2.5: 2.5, respectively) for 2 h. The resin was removed by filtration and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was dissolved in acetonitrile/water mixture and purified by HPLC.

- **Converting N-terminus Glycine to Aldehyde in Proteins**

Myoglobin (3.0×10^{-4} mmol) was dissolved in sodium phosphate buffer (25 mM, pH 6.5). A stock solution of PLP was prepared by dissolving PLP (0.021 mmol) in 1 mL of sodium phosphate buffer (25 mM, pH 6.5) and adding 24 μ L of NaOH solution (1 M) and vortexing. A microelectrode was used to ensure the pH of the stock solution remained at pH 6.5. The dissolved protein and PLP stock solution were combined in a 1:1 ratio by volume and incubated at 37°C for 4 h. The protein and protein aldehyde were analyzed by ultra-performance LC coupled with a quadrupole time-of-flight mass spectrometer (Supporting Information).

Chapter 3: Supporting Information

3.1 ^1H NMR, ^{13}C NMR, and MS

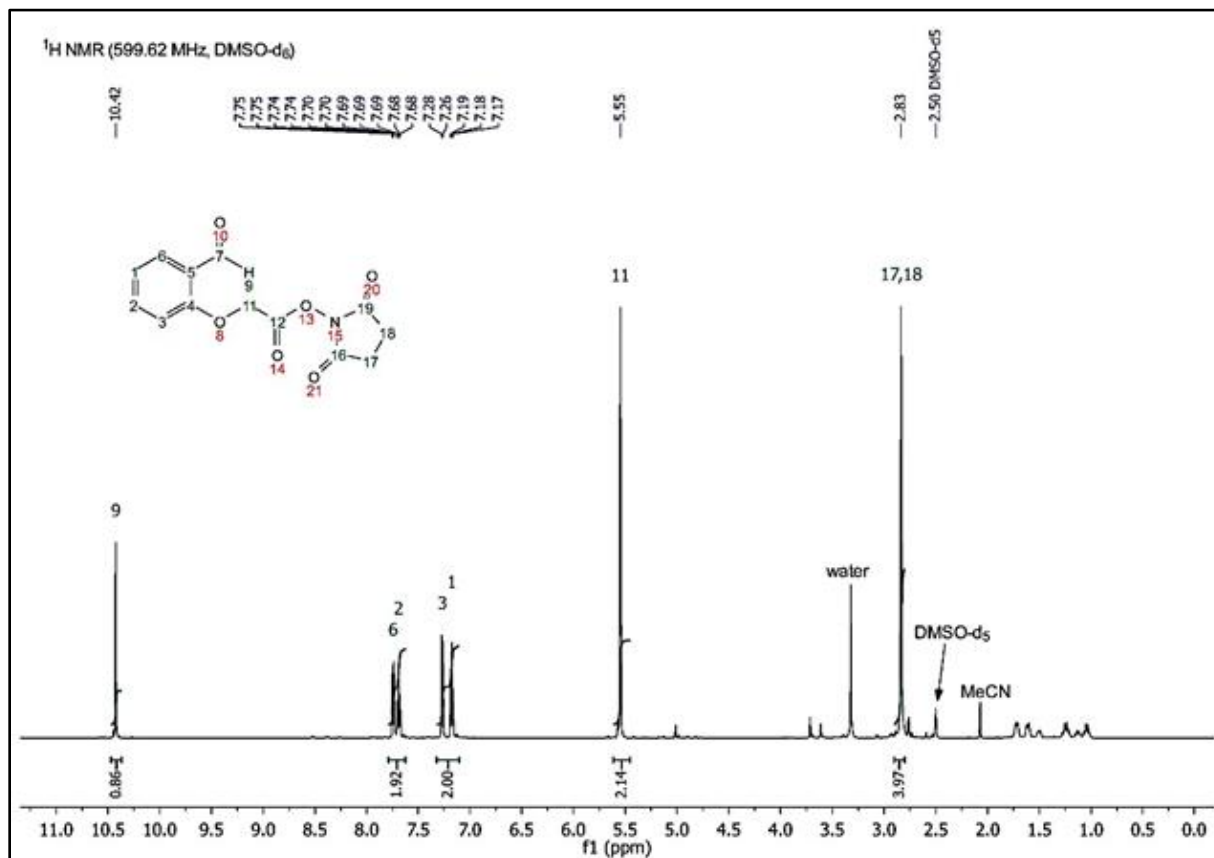


Figure S1: ^1H NMR of formylphenoxy NHS ester.

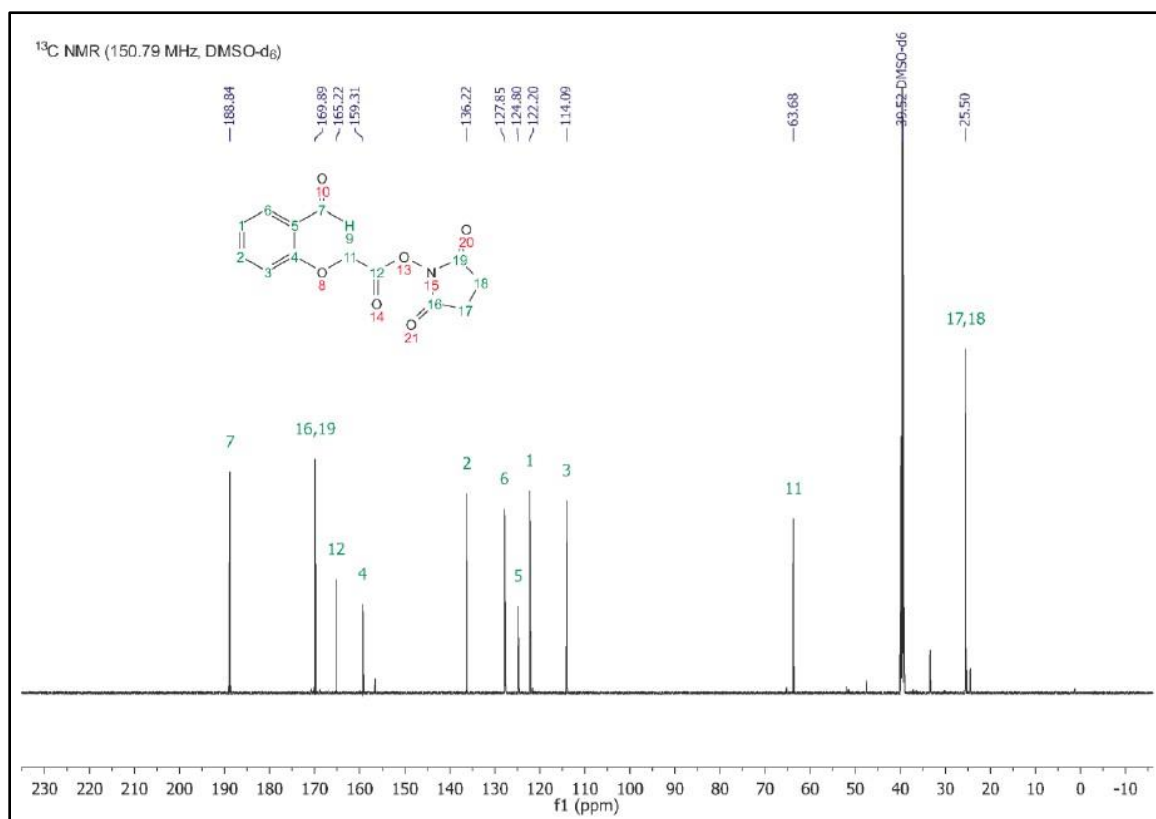


Figure S2: ¹³C NMR of formylphenoxy NHS ester.

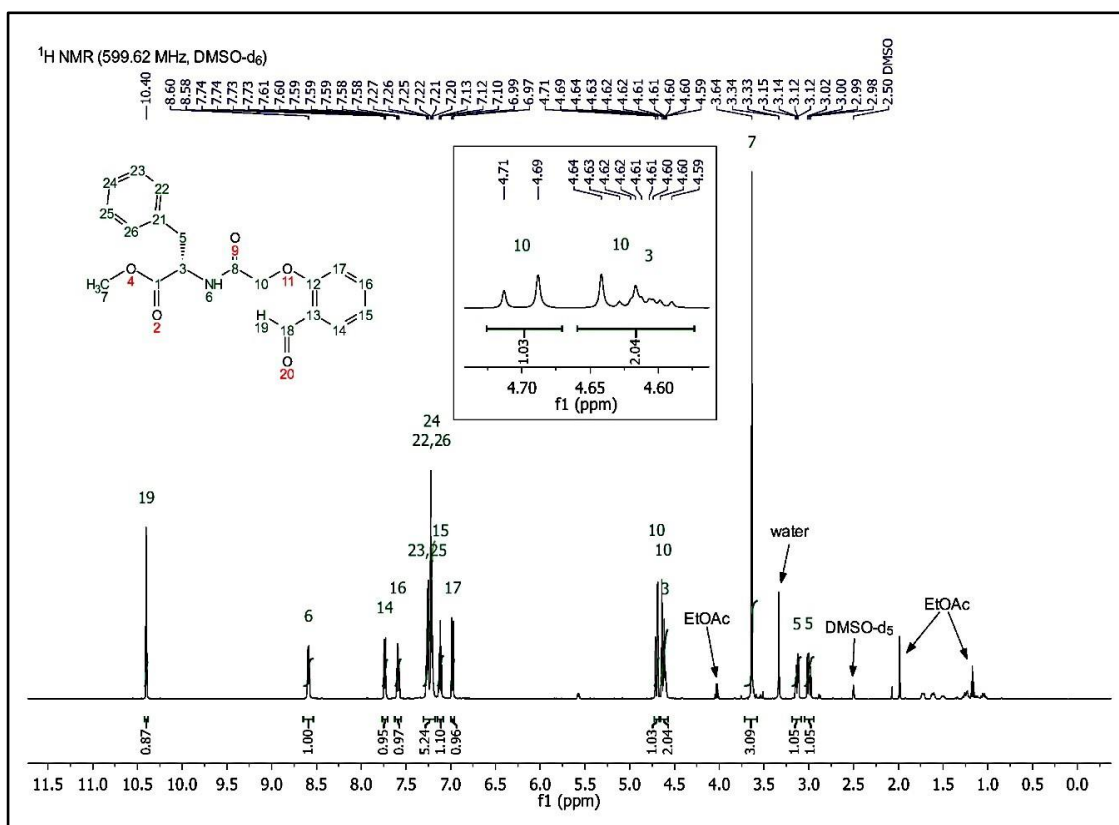


Figure S3: ¹H NMR of phenylalanine aldehyde.

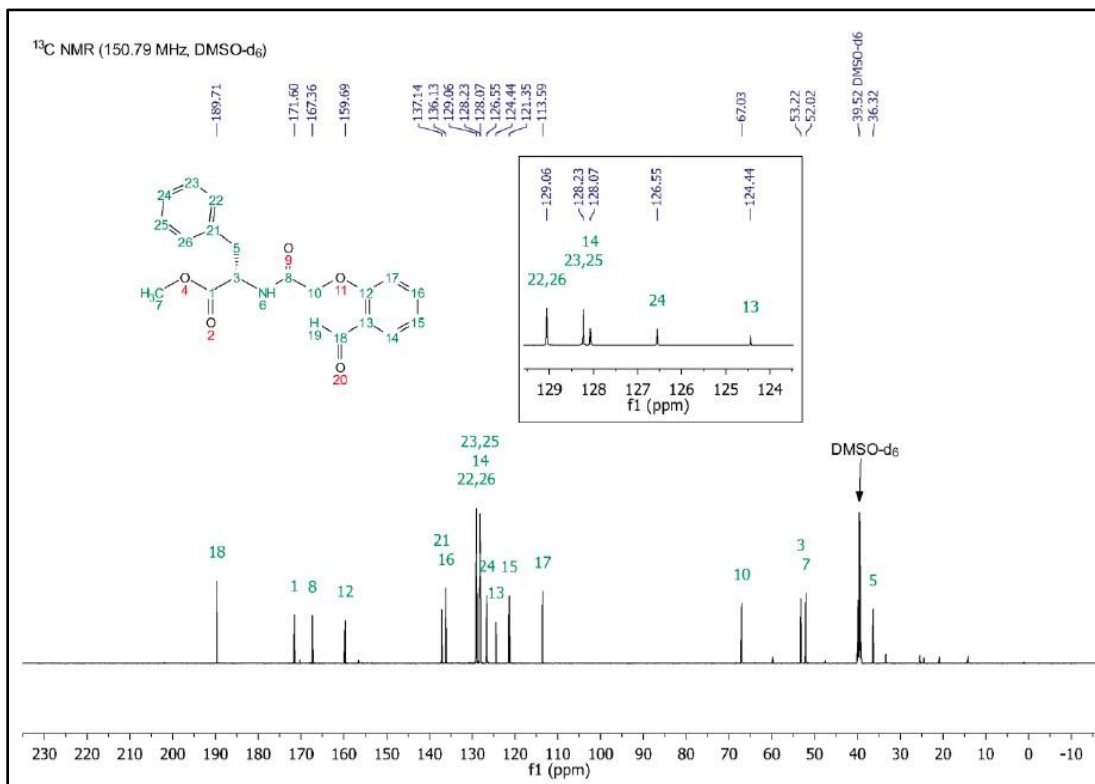


Figure S4: ¹³C NMR of phenylalanine aldehyde.

3.2 HPLC and LCMS Traces

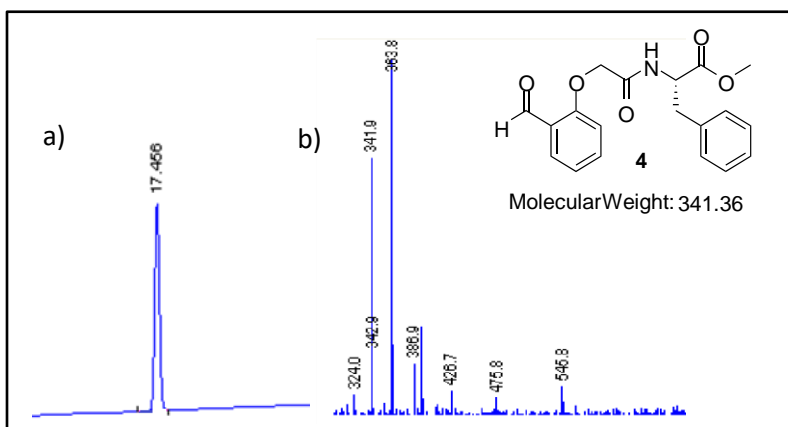


Figure S5: a) HPLC and b) LCMS traces of phenylaldehyde aldehyde.

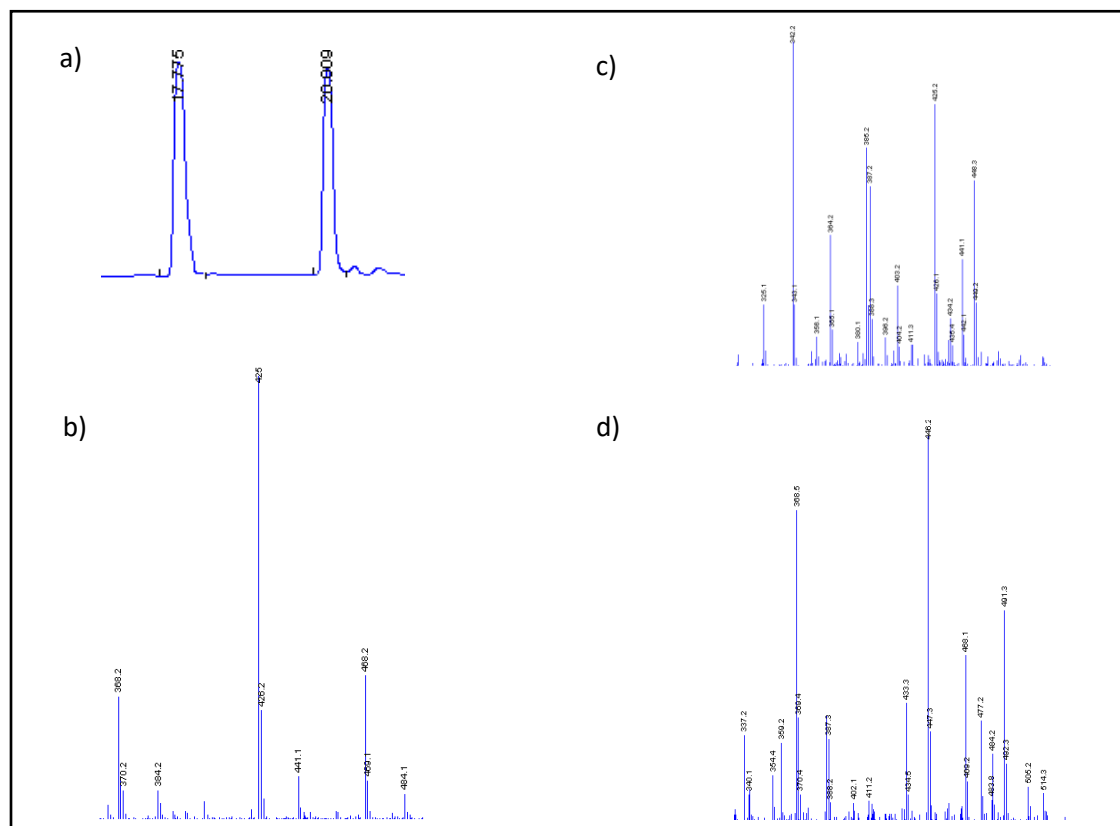


Figure S6: a) HPLC of phenylalanine reaction with nitromethane using proline as catalyst; b) LCMS traces of crude phenylalanine reaction with nitromethane using proline as catalyst; c) LCMS of 17.7 min peak; d) LCMS of 20 min peak; 24 h reaction.

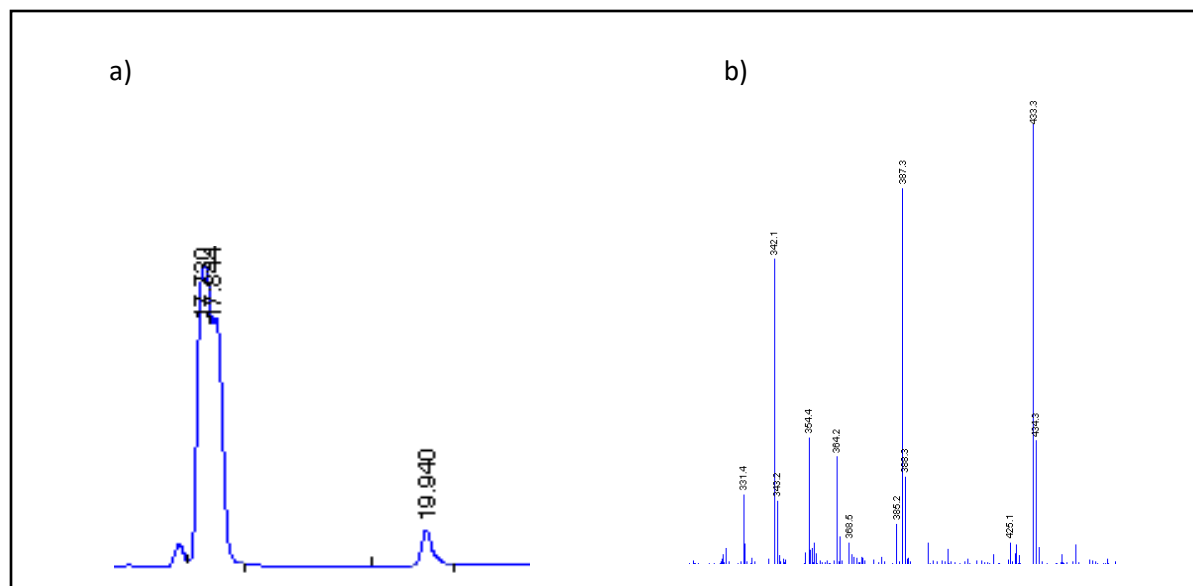


Figure S7: a) HPLC of phenylalanine reaction with nitromethane control reaction; b) LCMS traces of crude control reaction; 6 h reaction.

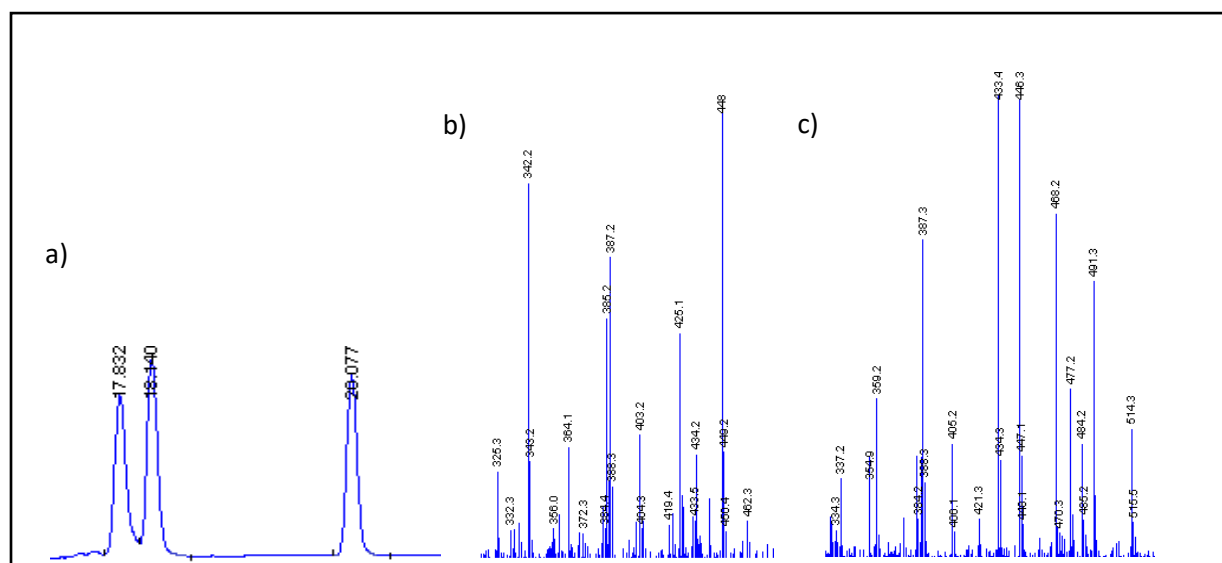


Figure S8: a) HPLC of phenylalanine reaction with nitromethane using DIEA as catalyst; b) LCMS of 17.7 min peak; c) LCMS of 20 min peak; 6 h reaction.

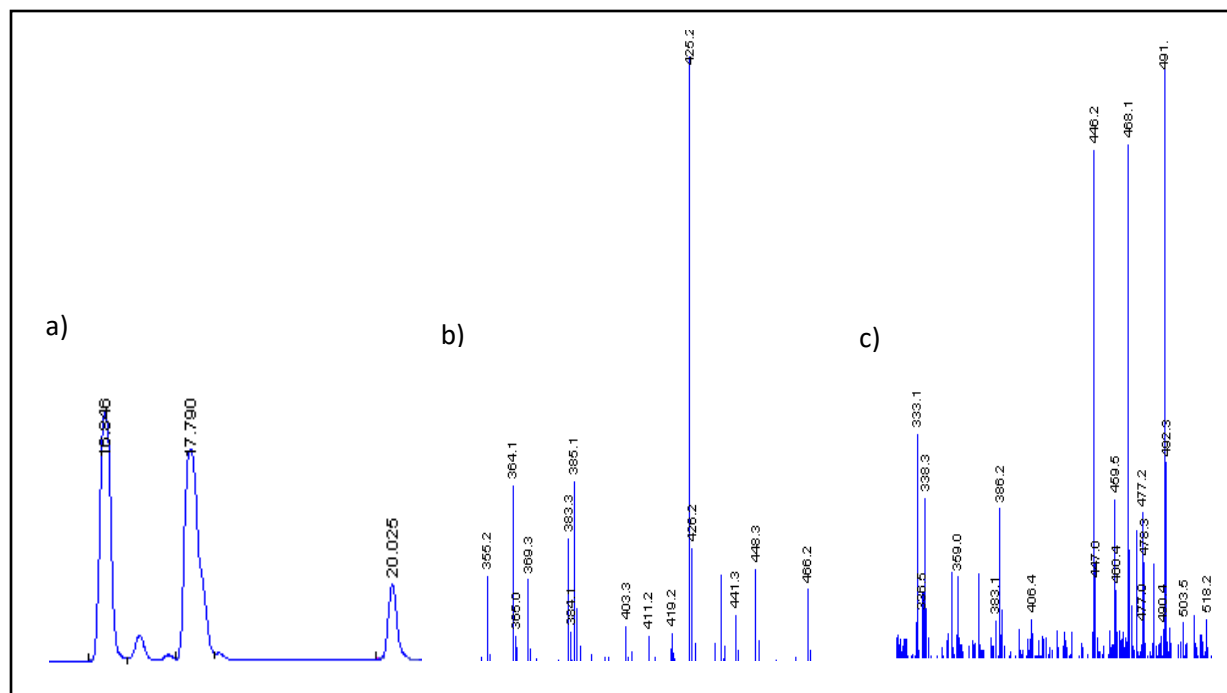


Figure S9: a) HPLC of phenylalanine reaction with nitromethane using N,N-Diphenylthiourea as catalyst; b) LCMS of 17.7 min peak; c) LCMS of 20 min peak; 6 h reaction.

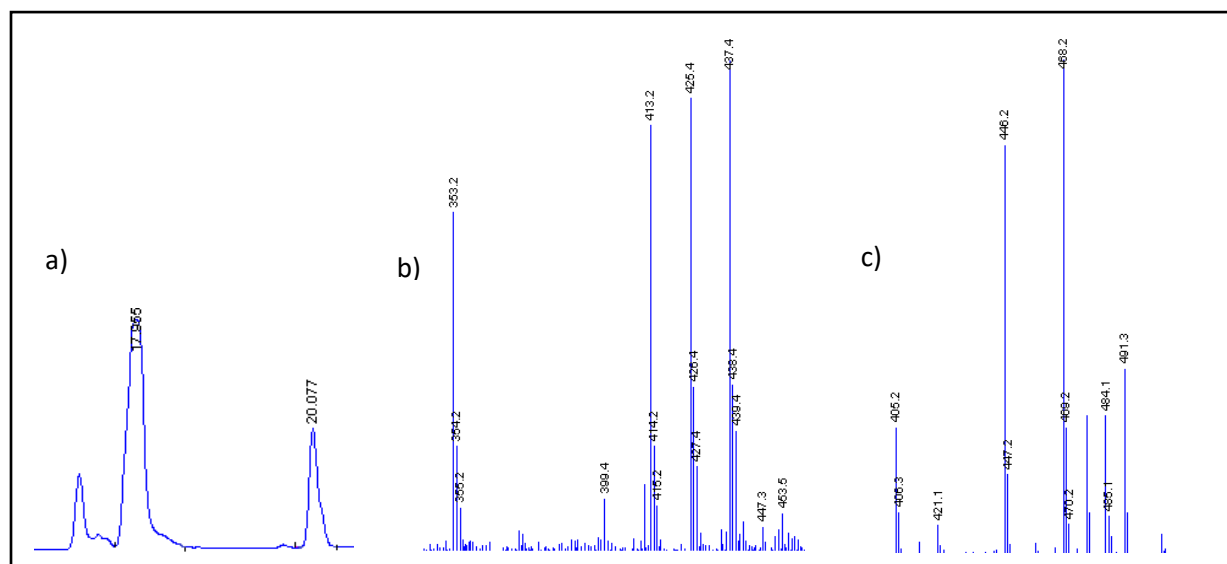


Figure S10: a) HPLC of phenylalanine reaction with nitromethane using glycine methyl ester as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6 h reaction.

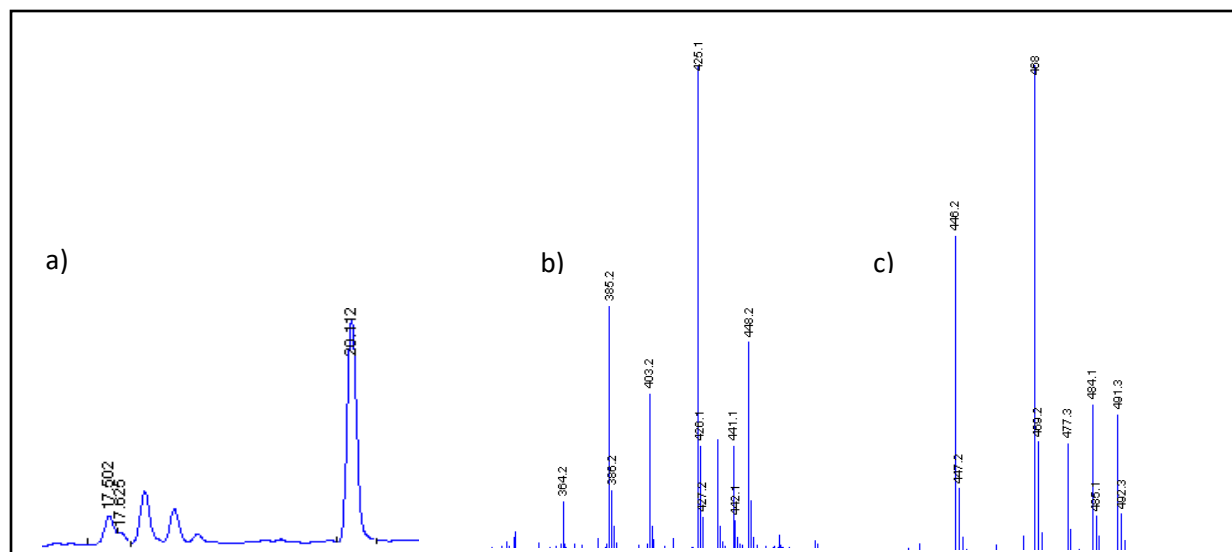


Figure S11: a) HPLC of phenylalanine reaction with nitromethane using threonine as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6 h reaction.

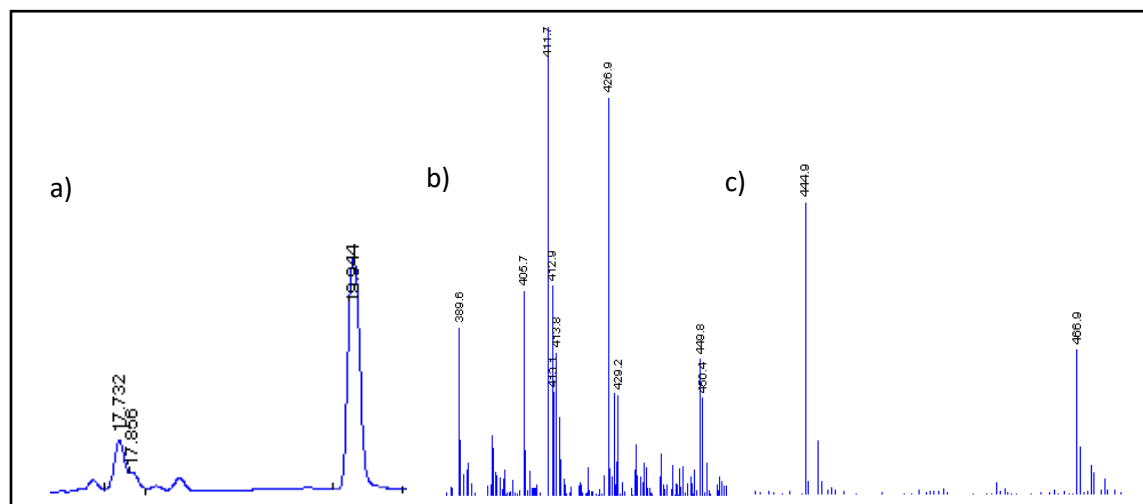


Figure S12: a) HPLC of phenylalanine reaction with nitromethane using glycine as catalyst; b) LCMS of 17.9 min peak; c) LCMS of 20 min peak; 6 h reaction.

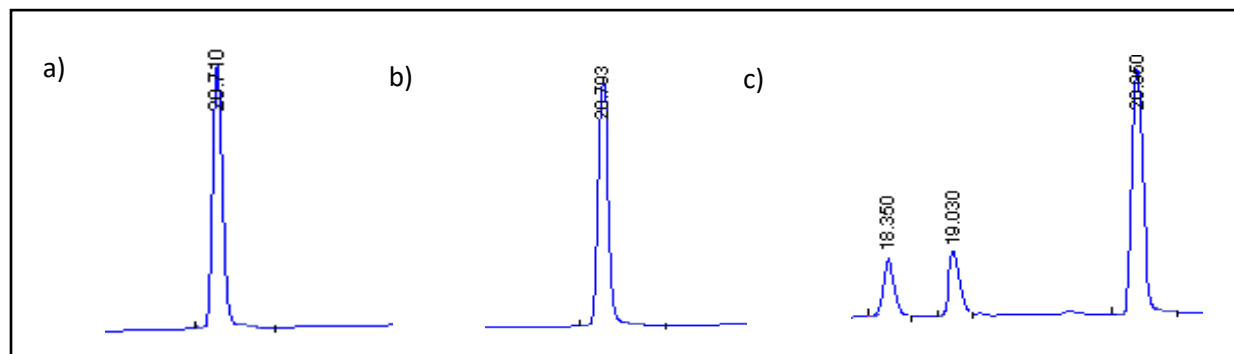


Figure S13: HPLC of phenylalanine reaction with nitromethane using glycine as catalyst under different pH conditions. a) reaction under acidic condition, pH 3.5 b) reaction under biological condition, pH 7.5; c) reaction under basic condition, pH 10.5.

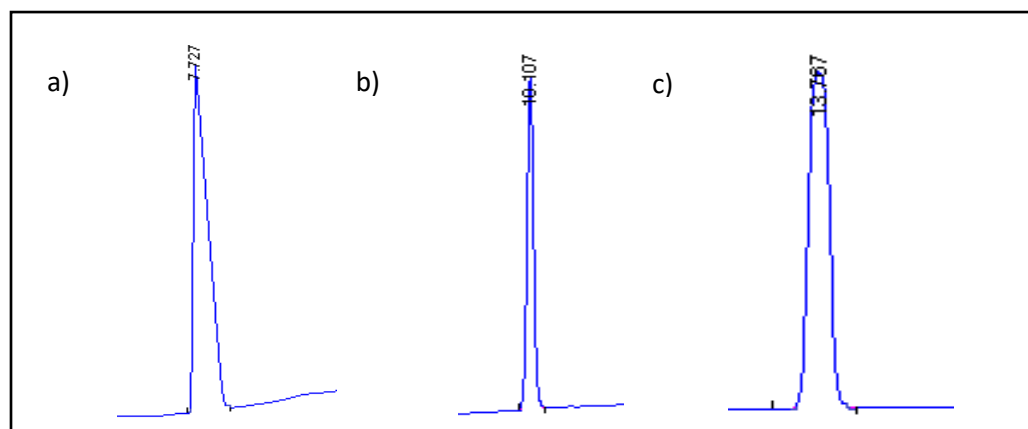


Figure S14: HPLC of SVF. a) SVF peptide b) SVF peptide aldehyde c) SVF product of nitromethane reaction using glycine as catalyst; 30 min.

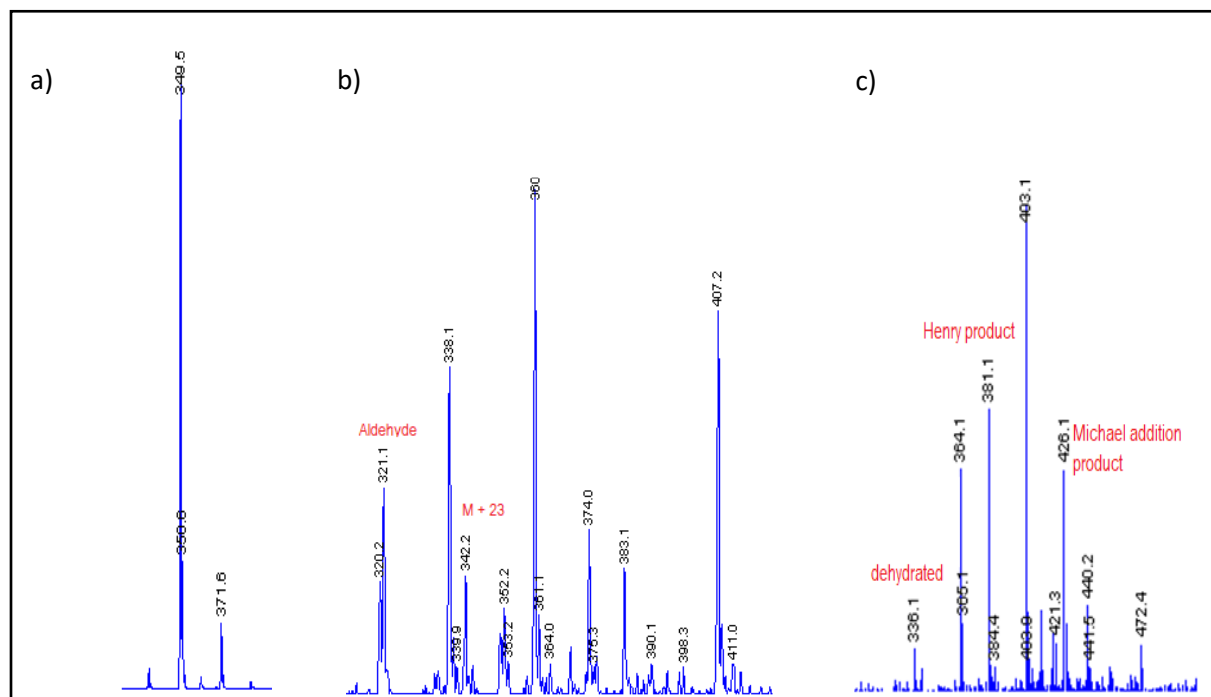


Figure S15: LCMS of SVF. a) SVF peptide; b) SVF peptide aldehyde; c) SVF product of nitromethane reaction using glycine as catalyst.

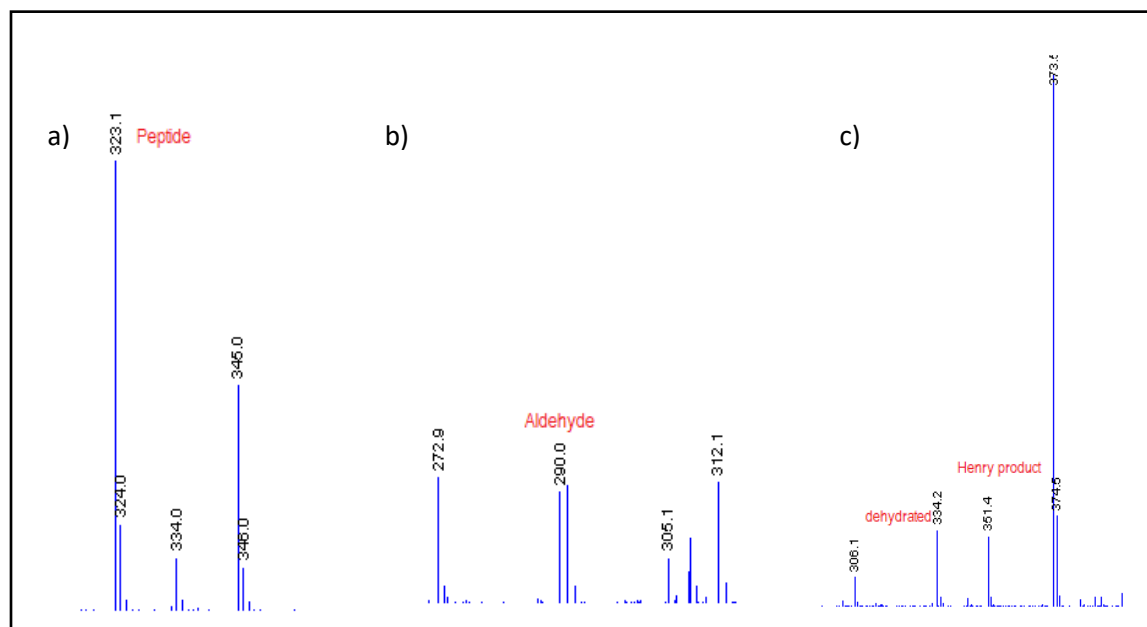


Figure S16: LCMS of SAF. a) SAF peptide; b) SAF peptide aldehyde; c) SAF product of nitromethane reaction using glycine as catalyst.

References

1. Chaplin, D. D. *J. Allergy Clin. Immunol.* **2010**, *125*, S2.
2. Wheeler-Jones, C. P. D. *Heart.* **2005**, *91* (10), 1366.
3. Yang, N. J.; Hinner, M. J. *Methods Mol. Biol.* **2015**, *1255*, 29.
4. Schwartz, D. S.; Blower, M. D. *Cell Mol. Life Sci.* **2016**, *73*, 79.
5. Berkel, S. S.; Eldijk, M. B.; Hest, J. C. M. *Angew. Chem. Int. Ed.* **2011**, *50*, 8806.
6. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, **48**, 6974.
7. Hackenberger, C. P. R. ; Schwarzer, D. *Angew. Chem.* **2008**, *120*, 10182.
8. Zalipsky, S. *Bioconjugate Chem.* **1995**, *6*, 150.
9. Roberts, M. J.; Bentley, M. D.; Harris, J. M. *Adv. Drug. Deliv. Rev* **2002**, *54*, 459.
10. Chapman, A. P. *Adv. Drug. Deliv. Rev* **2002**, *54*, 531.
11. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.
12. Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24* (3), 333.
13. Leffler, J. E.; Temple, R. D. *J. Am. Chem. Soc.* **1967**, *89* (20), 5235.
14. Cox, E. D.; Cook, J. M. *Chem. Rev.* **1995**, *95*, 1797.
15. Seayad, J.; List, B. *Org. Biomol. Chem.* **2005**, *3*, 719.
16. Broyer, R. M.; Grover, G. N.; Maynard, H. D. *Chem. Commun.* **2011**, *47*, 2212.
17. Hudak, J. E.; Barfield, R. M.; de Hart, G. W.; Grob, P.; Nogales, E.; Bertozzi, C. R.; Rabuka, D. *Angew. Chem. Int. Ed.* **2012**, *51*, 4161.
18. Spears, R.J.; Fascione, M. A. *Org. Biomol. Chem.* **2016**, *14*, 7622.
19. Kolmel, D. K.; Kool, E. T. *Chem. Rev.* **2017**, *117*, 10358.

20. Su, X.; Aprahamian, I. *Chem. Soc. Rev.* **2014**, *43*, 1963.
21. Tatum, L. A.; Su, X.; Aprahamian, I. *Acc. Chem. Res.* **2014**, *47*, 2141.
22. Li, X.-G.; Haaparanta, M.; Solin, O. *J. Fluorine Chem.* **2012**, *143*, 49.
23. Walton, J. C. *Acc. Chem. Res.* **2014**, *47*, 1406.
24. Johnson, E. C.; Kent, S. B. *J. Am. Chem. Soc.* **2006**, *128* (20), 6640.
25. Dang, T. T.; Dang, T. T.; Langer, P. *Synlett.* **2011**, *2011*, 2633.
26. Abele, E.; Abele, R. *Curr. Org. Synth.* **2014**, *11*, 403.
27. Kovarikova, P.; Mrkvickova, Z.; Klimes, J. *J. Pharm. Biomed. Anal.* **2008**, *47*, 360.
28. Buss, J. L.; Ponka, P. *Biochem. Biophys. Acta, Gen. Subj.* **2003**, *1619*, 177.
29. Agarwal, P.; Kudirka, R.; Albers, A. E.; Barfield, R. M.; de Hart, G. W.; Drake, P. M.; Jones, L. C.; Rabuka, D. *Bioconjugate Chem.* **2013**, *24*, 846.
30. MacKenzie, D. A.; Sherratt, A. R.; Chigrinova, A.; Cheung, L. L. W.; Pezacki, J. P. *Curr Opin. Chem. Biol.* **2014**, *21*, 81.
31. McKay, C. S.; Chigrinova, M.; Blake, J. A.; Pezacki, J. P. *Org. Biomol. Chem.* **2012**, *10*, 3066.
32. McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21* (9), 1075.
33. Van Berkel, S. S.; Van Eldijk, M. B.; Van Hest, J. C. *Angew. Chem. Int. Ed. Eng.* **2011**, *50* (38), 8806.
34. Agarwal, P.; Weijden, J. V.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci.* **2013**, *110* (1), 46.
35. Lin, F. L.; Hoyt, H. M.; Van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 2686.
36. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2* (13), 1939.

37. Li, M.; Berritt, S.; Wang, C.; Yang, X.; Liu, Y.; Sha, S.; Wang, B.; Wang, R.; Gao, X.; Li, Z.; Fan, X.; Tao, Y.; Walsh, P. J. *Nat. Commun.* **2018**, *9*, 1754.
38. Alvarez-Casao, Y.; Marques-Lopez, E.; Herrera, R. P. *Symmetry*. **2011**, *3*, 220.
39. Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: a Practical Approach*; Oxford University Press: New York, **2000**.
40. Tsuda, S.; Yoshiya, T.; Mochizuki, M.; Nishiuchi, Y. *Org. Lett.* **2015**, *17*, 1806.
41. Shaikh, I. R. *J. Catal.* **2014**, *402860*, 1.