

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced 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 or dissertation in whole or in part in all forms of media, now or hereafter known, 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 or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature

Kuei-Chien Tang

Date

Chemo and Site Selective Modification of Proteins and Chemical Tools for
Studying Mono-Methyl Lysine (Kme)

By

Kuei-Chien Tang

Doctor of Philosophy
Chemistry

Monika Raj, PhD
Advisor

Huw M. L. Davies, PhD
Committee Member

David Lynn, PhD
Committee Member
Accepted:

Kimberly J. Arriola, PhD
Dean of the James T. Laney School of Graduate Studies

Date

Site Selective Modification of Proteins and Chemical Tools for Studying Mono-Methyl Lysine
(Kme)

By

Kuei-Chien Tang

B.Sc., National Dong Hwa University, 2012

M.S., National Sun Yat-Sen University, 2014

Advisor: Monika Raj, PhD

An abstract of

A dissertation submitted to the Faculty of the James T. Laney of Graduate Studies of Emory
University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Chemistry

2022

Abstract

Site Selective Modification of Proteins and Chemical Tools for Studying Mono-Methyl Lysine

(Kme)

By: Kuei-Chien Tang

Site-selective modification of protein is a significant technique that can allow us to modify a single amino acid with biologically active small molecules or fluorophores to track and monitor the enzyme in the living system for biological study. The universal limitation for most of the current modification methods is the low detection sensitivity of the forming product, poor stability and low selectivity of probes for lysine modification, and lack of universal compatibility for the N-terminal modification. In order to circumvent the current limitations, I designed and synthesized various types of azole, azoline, and azolinium thioethers, which are highly selective and specific to our desired targets, and the reactivity and selectivity of those probes are highly tunable by the heteroaromatic ring, and methylation state. Moreover, those probes are able to act as a chemo-selective charge booster to enhance the mass detection sensitivity of the forming product. Our probes 1d-yne, 1c-yne, and N3-1o were applied to the proteomic profiling of lysine and cysteine in HEK293T cell lysate. The probe 1c-yne is highly selective and reactive to the lysine residues. Moreover, 1c-yne shows high permeability and hydrolytic stability and is able to label both cytoplasmic and nuclear proteins in live cells without necrosis those indicating its potential as a probe for rapid live cell labeling and covalent inhibitor. In tandem with the project above, I have developed two novel chemical tools for studying lysine methylation post-translation modification (PTM). Methylation of lysine regulates gene transcription and RNA, DNA binding, and any aberrant change of methylation state and site will cause various diseases such as cancer and diabetes. In the first method, I used triazene cyclization to target the mono-methyl lysine selectively. The forming product indazole is highly stable under basic and acid conditions and easy to modify with different affinity targets. Moreover, this method is the first chemical technique applied to a single-molecule protein sequence to read out unknown Kme1 sites in a single workflow by fluorosequencing. For the second method, I develop a strategy that enables site-selective modification of a high-frequency Lys residue in the mono-methyl lysine containing (Kme) protein for studying, monitoring, and tracking the PPI between the K-me protein and its reader and eraser proteins. The probes I have designed have a diazonium salt warhead with a mask group for selective labeling of mono-methyl lysine. The flexible linker connects with various electrophiles for intramolecular modification and the light-activated group for capturing the reader and eraser proteins. The probe AI showed high intramolecular labeling efficiency with a trace amount of intermolecular side product with different proline, K-me-containing peptides, and Histone-III truncated peptides. Those novel tools offer a chemical platform for identifying and studying the role of monomethyl lysine (Kme) in the whole proteome and a starting point for therapeutic interventions.

Site Selective Modification of Proteins and Chemical Tools for Studying Mono-Methyl Lysine
(Kme)

By

Kuei-Chien Tang

B.Sc., National Dong Hwa University, 2012

M.S., National Sun Yat-Sen University, 2014

Advisor: Monika Raj, PhD

A dissertation submitted to the Faculty of the James T. Laney of Graduate Studies of Emory
University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Chemistry

2022

Acknowledgements

First, I would like to thank the following people, Dr. Yang-Hsiang Chan, Dr. Kou-Cheng Peng, and Dr. Ming-Tsz Chen, for their assistance and advice. They supported me greatly during my graduate school application and made my Ph.D. study abroad possible. They have significantly influenced my life; I would like to share this thesis with them and wish them all the best in their life and research.

I would like to acknowledge and give my warmest thanks to my supervisor Dr. Monika Raj who made this work possible. Her guidance and advice carried me through all the stages of my research projects. I would also like to thank my committee members, Dr. Huw Davies and Dr. David Lynn, for letting my defense be an enjoyable moment and for your thoughtful comments and suggestions. Thanks to you.

I would also like to give special thanks to all my family and friends for their support and company for the past five years. Although we are far away from each other, we are still able to talk and share our daily life with each other by Line or Face book.

I would like to thank my cohorts Victor Adebomi and Ogonna Nwajiobi for their tremendous scientific and emotional support. I am fortunate to have witnessed their transformation from a single man to a husband and, eventually, a father. I wish all the best in their life and family. I also want to thank my lab mates Rachel Wills, Benjamin Emenike, Angele Bruce, Patricia Rodriguez, Lyndsey Prosser, Samrat Sahu, Rajendra Shirke, Patrick Czabala, Julia, and Daniel. They treat me like their own family and give me a lot of support during my Ph.D. I wish them all the best in their career and life.

I would like to thank all my collaborators, Dr. Keriann M. Backus from UCLA for the proteomic study, Dr. Jennifer Spangle and Jonathan Farhi from Emory for cellular imaging, Dr. Kendall Houk from UCLA for the computational study, Dr. Eric Anslyn from UT Austin for the single-molecule protein sequencing, Dr. Daniel K. Nomura from UC Berkeley for the proteomic study and Melissa from Auburn University for the MS analysis. Without their help, I would not achieve what I have achieved.

Lastly, I would like to thank my cats, Smokey and Takareni. They are the best cat in the world; for me, they are my children, friend, and family. I appreciate their company for the past few years, especially when I am sad or frustrated because of my work. I hope we will be together forever.

In the end, it is my pleasure to get my Ph.D. from Emory University, and It would not be possible to finish this thesis without the help of many kind people surrounding me.

Table of Contents

1.1 General Introduction to approaches to site-selective protein modification:	17
1.1.1 N terminus modification	17
1.1.2 Lysine residue modification:	22
1.1.3 Cysteine side chain modification	24
1.2 Results and Discussion	26
1.2.1 Screening of azoline,azole, and azolinum with standard peptide	26
1.2.2 Synthesis of Azoline, Azole, and Azolinum and their Analogs	27
1.2.2 One-Step Azolation Strategy for Site-and Chemo-Selective Labeling of Proteins with Mass-Sensitive Probes	32
1.2.2.2 Studying Universal Sequence Compatibility of N-terminus Azolztion	32
1.2.2.3 Protein modification-reaction optimization	34
1.2.2.4 Synthesis of functionalized Oxazoles for N-terminal modification	36
1.2.2.5 Protein Scope for N-terminal Oxazolation	36
1.2.2.6 Rate and Stability Studies of Oxazolation	37
1.2.2.7 Labeling of Proteolytic Fragments in the Complex Mixture	39
1.2.2.8 Oxazoline as Mass Sensitivity Booster	41
1.2.2.9 Unmodified Myoglobin and Modified Myoglobin Bioactivity Assay	45
1.2.2.10 Conclusion	46
1.2.3 Tunable heteroaromatic azoline thioethers (HATs) for cysteine profiling	48

1.2.3.1 Rate and Stability Study of HAT Probe.....	48
1.2.3.2 Protein Modification with HAT	50
1.2.3.3 Selective tagging of Cys in a complex mixture	55
1.2.3.4 Reversibility and stability of the HAT–cysteine conjugation.....	57
1.2.3.5 Reactivity inversion by HAT probes	60
1.2.3.6 HAT as mass sensitivity booster	63
1.2.3.7 HAT probes for gel-based ABPP studies	68
1.2.3.8 Conclusion	70
1.2.4 Tunable amine-reactive electrophiles for selective profiling of lysine	71
1.2.4.1 Chemoselectivity Reaction of TARE Probes with Proteins.....	71
1.2.4.2 Rate, Stability and Reversibility of TARE Probes.....	73
1.2.4.3 Chemo-proteomic studies of TAREs	80
1.2.4.4 Live cell label ligand amino-acid selectivity in proteome by probe 1c-yne.....	86
1.2.4.5 Computational study of probe 1c-yne	89
1.2.4.6 Conclusion	90
Reference:	92
Chapter 2: Chemical Tool for Tagging Mono-Methyl Lysine and K-me-Directed Modification ...	101
2.1 General Introduction of Mono-Methyl Lysine	101
2.2 Coarctate Cyclization for Selective labeling of Monomethyl Lysine Posttranslational Modifications	103
2.2.1 Development of Triazeneation Coarctate Cyclization (TCC).....	104
2.2.2 Stability Studies.....	107

2.2.3 Chemoselectivity studies for the formation of 2H-indazole-3-carbaldehyde.....	108
2.2.4 Tyrosine Vs Kme modification	109
2.2.5 Substrate scope with varying 2-ethyne phenyldiazonium ions.	112
2.2.6 Pan-specificity of TCC: Further Diversification.....	115
2.2.7 Selective labeling of Kme1 peptides in a complex cell lysate mixture by TCC	117
2.2.8 Single-molecule sequencing for identification of Kme1 sites by TCC.....	118
3.1 Design of Chemical Probes for Kme-Directed Modification (Kme-DM).....	125
3.2 Design and synthesis of various probes for Kme-DM.....	126
3.3 Optimization of intramolecular reaction under UV lamp with different proline containing peptides and chemo selectivity study.....	137
3.4 Chemoselectivity and control study of AI-I probe	140
3.5 Screening of AI-III probe with K-2, K-3 and K-4 peptides under UV lamp	140
3.6 Site-specific study with multiple lysins containing peptide	141
3.7 Intramolecular reaction study of various AI analogs with different affinity tags	142
3.8 Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe.....	143
3.9 Kme-Directed modification of truncated histone peptide:	146
Supporting Information for Chapter One:.....	148
Supplementary Figure 1 : MS/MS of modified peptide.....	148
Supplementary Figure 2: General method for the verification of the chemo- and site-selective nature of oxazoline probe with peptides.	153
Supplementary Figure 3. Azolation vs 2-PCA method.....	161
Supplementary Figure 4. Optimization of the oxazolation on protein Myoglobin Mb.	163

Supplementary Figure 5. Modification of myoglobin Mb by different oxazoline derivatives.	165
Supplementary Figure 6. Mass intensity enhancement of digested myoglobin by oxazolation.	175
Supplementary Figure 7. Mass intensity enhancement of intact Myoglobin by oxazolation as compared to unmodified myoglobin.....	177
Supplementary Figure 8: Rate study of HAT bioconjugation.....	178
Supplementary Figure 9. Modification of myoglobin Mb with different heteroaromatic azoline compounds 1a-1c, 1j, and 1n.....	184
Supplementary Figure 10. Selective cysteine bioconjugation of Insulin a chain and b chain with compound 1o.....	192
Supplementary Figure 11. Modification of reduced insulin with 1m.....	192
Supplementary Figure 12. Modification of bovine serum albumin (BSA) and lysozyme with HAT probe 1o	193
Supplementary Figure 13. Enrichment of cysteine containing peptides with 1o in mixture of proteolytic fragments.....	195
Supplementary Figure 14. Stability study of Ac-GCF-1o under different pH conditions.....	199
Supplementary Figure 15. Reversible study of Cys-HAT biconjugate with NaBH₄.....	202
Supplementary Figure 15. Reversible study of probe 1o.....	205
Supplementary Figure 16. Dehydroalanine synthesis from cysteine 2c using 1o at high pH.....	207
Supplementary Figure 17. Aza-Michael addition and thiol-ene reaction of dehydroalanine	209
Supplementary Figure 18. Aza-Michael addition of dehydroalanine lysozyme.....	212
Supplementary Figure 20. Mass sensitivity of 1o-Ac-GCF (low concentration)	214

Supplementary Figure 21. Mass intensity enhancement of N3-1o -reduced insulin bioconjugate products.	216
Supplementary Figure 24. Rate study of 1c-yne and 1d-yne with peptide Ac-GKF 2c.....	225
Supplementary Figure 25. Stability study of 1c-yne, 1d-yne, NHS ester and STPyne	228
Supplementary Figure 26. Reversibility study of 1e-Cys-conjugate (VCF-1e) and 1d-thio-conjugate with lysine methylester.	229
Supplementary Figure 27. Chemoproteomic analysis of residue selectivity.....	234
Supplementary Figure 28. TAREs for gel-based ABPP	237
Supplementary Figure 29. Stability study of GCF-1c-yne conjugate in TCEP, and 1d-thio-conjugate under sodium phosphate buffer.....	243
Supplementary Figure 30. Live cell labeling of proteins using different probes.....	244
Supplementary Figure 31. Cell viability studies with 1cyne	247
Supplementary Figure 32	247
Supporting Information of Chapter Two:	257
Supplementary Figure 1: Synthesis of 2H-indazole-3-carbaldehyde and benzotriazole	257
Supplementary Figure 2: General procedure for the modification of FAP with probes 2a and 3a.	263
Supplementary Figure 3: Stability study of 2H-indazole-3-carbaldehyde group in TFA and pyridine.	267
Supplementary Figure 4: Chemoselectivity studies for the TCC reaction	269
Supplementary Figure 5: Restriction of TCC side reaction with tyrosine	270
Supplementary Figure 6: Selective blocking of tyrosine by 1,3 diphenyl propynone.....	272
General procedure for tyrosine labeling by 1,3 diphenyl propynone	272

Supplementary Figure 7: Optimization of triazination coarctate cyclization with different diazonium salts analogs.	274
Supplementary Figure 8: Pan-specificity of triazination coarctate cyclization with truncated histone peptides.	291
Supplementary Figure 9: Enrichment of modified peptides by cysteine condensation and oxime chemistry	294
Supplementary Figure 10: Enrichment of modified truncated peptides by cysteine condensation and oxime chemistry.	296
Supplementary Figure 11: Selective enrichment of Kme1 containing peptides in a complex cell lysate mixture by TCC	300
Supplementary Figure 12: Single-molecule sequencing for identification of Kme1 sites by TCC.	302
Supporting Information for Chapter Three:	303
Supplementary Figure 1. Synthesis of different K-me-DM probes	303
Supplementary Figure 2. Optimization of intramolecular reaction under UV lamp with different proline containing peptides and chemo selectivity study	339
Supplementary Figure 3. Chemoselectivity and control study of AI-I probe	342
Supplementary Figure 4. Screening of AI-III probe with K-2, K-3 and K-4 peptides under UV lamp	343
Supplementary Figure 5. Site-specific study with multiple lysins containing peptide	344
Supplementary Figure 6. Digestion of modified peptide.	346
Supplementary Figure 7. Intramolecular reaction study of various AI analogs with different affinity tags	346

Supplementary Figure 8. Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe	348
Supplementary Figure 9. Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe with K-me peptide.....	353
Supplementary Figure 10. Chemo-selective study of unmasked AI probe.....	355
Supplementary Figure 11. Kme-Directed modification of truncated histone peptide.....	356
NMR for Chapter One.....	360

List of Abbreviations

¹³C NMR: Carbon-13 Nuclear Magnetic Resonance Spectroscopy

¹⁹F NMR: Fluorine Nuclear Magnetic Resonance Spectroscopy

¹H NMR: Proton Nuclear Magnetic Resonance Spectroscopy

ACN: Acetonitrile

ADCs: Antibody-Drug Conjugates

Ala: Alanine

Arg: Arginine

Asn: Asparagine

Asp: Aspartate

Bn: Benzyl

Boc: tert-Butyloxycarbonyl

COESY: Correlation Spectroscopy

CREST: Conformer-Rotamer Ensemble Sampling Tool

CTP: Cyclic Tetrapeptide

Cu(OTf)₃: Copper(II) trifluoromethanesulfonate

Cys: Cysteine

DARs: Drug-to-Antibody ratio

DCM: Dichloromethane
DEPBT: 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one
DFT: Density Functional Theory
DIEA: N,N-Diisopropylethylamine
DMAP: 4-dimethylaminopyridine
DMF: dimethylformamide
DMSO: Dimethyl sulfoxide
DSC: N,N'-Disuccinimidyl carbonate
Et: ethyl
Et₃N: Triethylamine
EtOAc: ethylacetate
Gln: Glutamine
Glu: Glutamate
Gly: Glycine
H₂O: Water
HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate)
HMBC: Heteronuclear Multiple Bond Correlation
HOBT: hydroxybenzotriazole
HPLC: High Performance Liquid Chromatography
HPLC: high-performance liquid chromatography
HRMS: High-Resolution Mass Spectrometry
HSQC: Heteronuclear Single Quantum Coherence
iBu: isobutyl
Ile: Isoleucine
IMHB: Intramolecular Hydrogen Bonding
K₂CO₃: Potassium carbonate
LC-MS and LC-MS/MS: Liquid Chromatography with tandem mass spectrometry
Leu: Leucine

Lys: Lysine
Me: methyl
Me: Methyl
MeOH: methanol
Met: Methionine
MS/MS: Tandem mass spectrometry
MS: Mass Spectrometry
NMR: Nuclear Magnetic Resonance Spectroscopy
OMe: Methoxy
Pd-PEPPSI: Palladium - pyridine-enhanced precatalyst preparation stabilization and initiation
PEG: Polyethylene glycol
Ph: Phenyl
pH: potential of Hydrogen
Phe: Phenylalanine
PPIs: Protein-Protein Interaction
Pro: Proline
PTM: Post-Translational Modification
PyBOP: benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
ROESY: Rotating-frame nuclear overhauser effect correlation spectroscopy
Ser: Serine
TFA: Trifluoroacetic acid
TfOH: Trifluoromethanesulfonic acid
THF: Tetrahydrofuran
thio-CDI: Thiocarbonyldiimidazole
Thr: Threonine
TMS(OTf)₃: Trimethylsilyl trifluoromethanesulfonate
TOCSY: Total Correlation Spectroscopy
Trp: Tryptophan

Ts: Tosyl

TS: Transition State

Tyr: Tyrosine

UV spectroscopy: Ultraviolet-visible spectroscopy

Val: Valine

Chapter 1: Site-Selective Modification of Proteins

1.1 General Introduction to approaches to site-selective protein modification:

Site-specific chemical modification of protein with fluorescent or biologically active small molecules has a crucial role in the development of novel biologically active protein conjugates for applications in biology and medicine, the discovery of the tools for biochemical and biophysical studies, monitoring, modulating, and tracking proteins in living systems for biological research^{1,2}. The chemo-selectivity for certain of the natural 20 amino acids and N-terminus and mildness of the selected approaches are essential to precisely install modifications at pre-determined sites without disturbing the structure, function, and activity of the protein³.

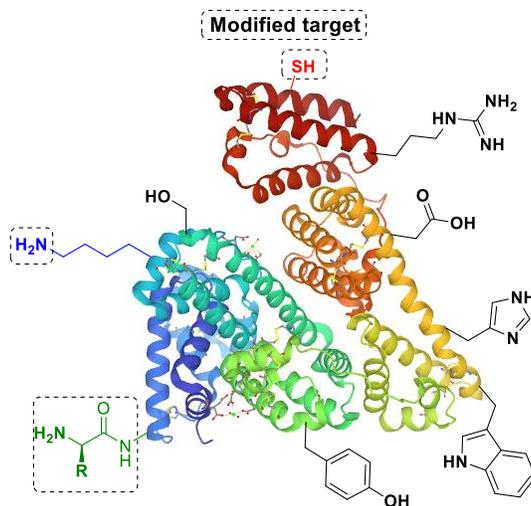


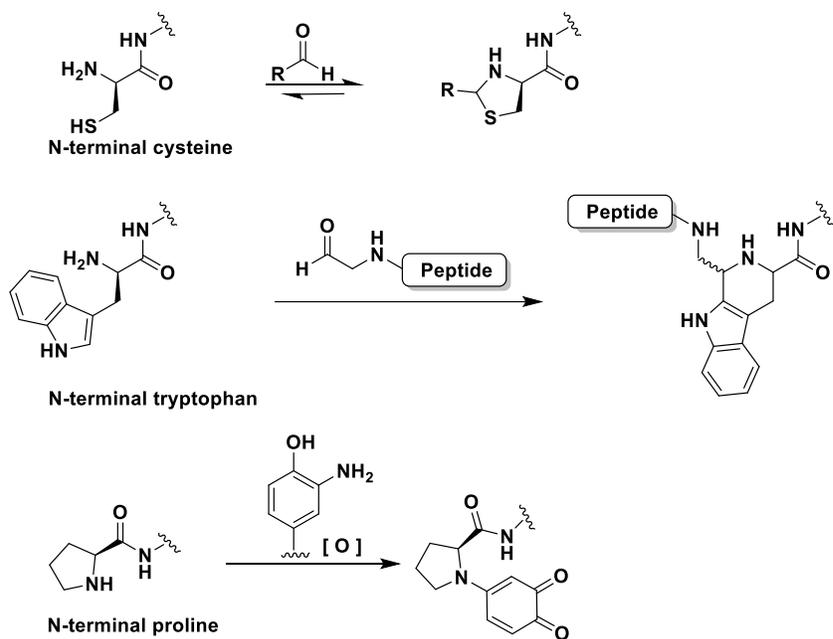
Figure 1. Potential site for chemo-selective modification

1.1.1 N terminus modification:

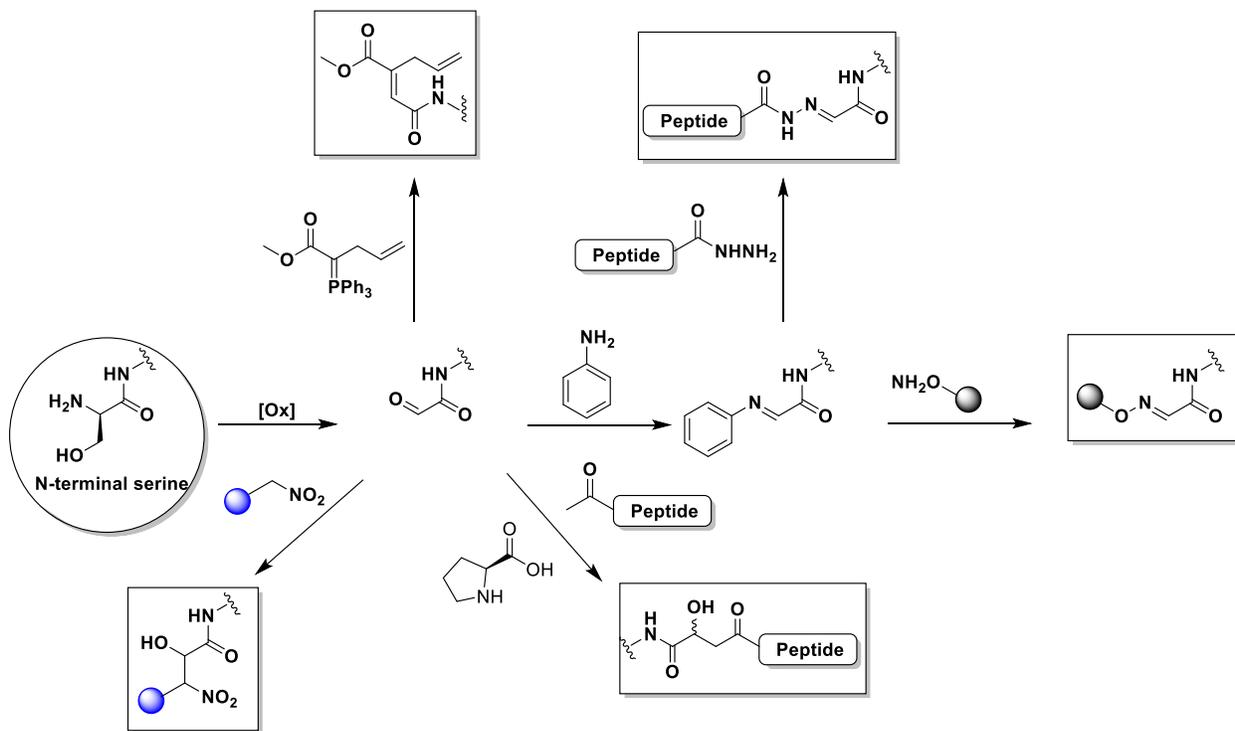
Modification of proteins at a single site with varying synthetic molecules is of high significance in the field of chemical biology, for efficient drug delivery, material science and synthesis of bio-hybrid materials. Recently, there has been a significant increase in the number of methods for site-selective labeling of N-terminus but is limited by the requirement of particular amino acids at the N-terminus. Such as N-terminal cysteines leading to the formation of thiazolidines with

aldehydes⁴ (Figure 2a). Periodate oxidation of N-terminal serine and residues for oxime⁵, hydrazone⁶, Wittig, Aldol⁷ and Henry bioconjugation⁸ (Figure 2b). N-terminal tryptophans can be modified via Pictet-Spengler reactions⁹ (Figure 2a). N-terminal proline can be selectively labeled by the oxidative coupling with o-aminophenols¹⁰ (Figure 2a). Some of the methods also required multiple steps¹¹ (Figure 2c). A few strategies that work with most N-terminal amino acids have been developed^{12,13} (Figure 2d). A recent 2 pyridinecarboxaldehyde (2PCA) method provides one-step approach for the N-terminal modification of proteins¹⁴ but is unable to tag proteins with proline amino acid in the second position, and doubly modify the peptide with glycine amino acid at the N-terminus, thus lack universal sequence compatibility (Figure 2d). Another major limitation with the current methods of N-terminal labeling of proteins is the low mass detection sensitivity of the resulting protein bioconjugates thus making it difficult to characterize. The low mass sensitivity of the N-terminal labeled bioconjugates is due to the blockage of the free N-terminus by hydrophobic groups. This is a major issue while analyzing the conjugated proteins in a proteomic mixture. Consequently, there is a great need to develop new synthetic methodologies that can circumvent the aforementioned limitations and provide an efficient strategy for site-selective labeling of proteins with mass sensitive probes and exhibit universal sequence compatibility.

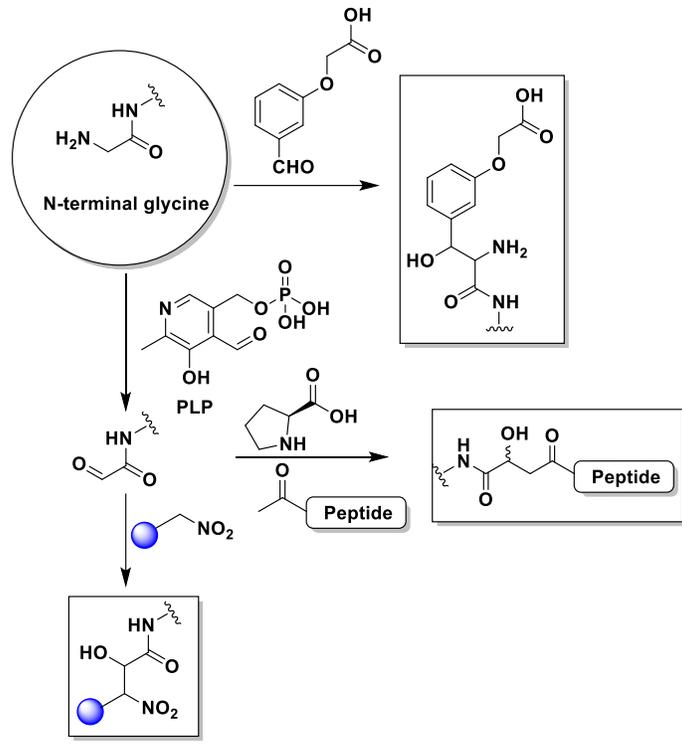
a. Site selective N-terminal modification of N-terminal tryptophan, cysteine, and proline



b. Site selective N-terminal modification of N-terminal serine



c. Site selective N-terminal modification of N-terminal glycine



d. N-terminal modification techniques for universal amino acids

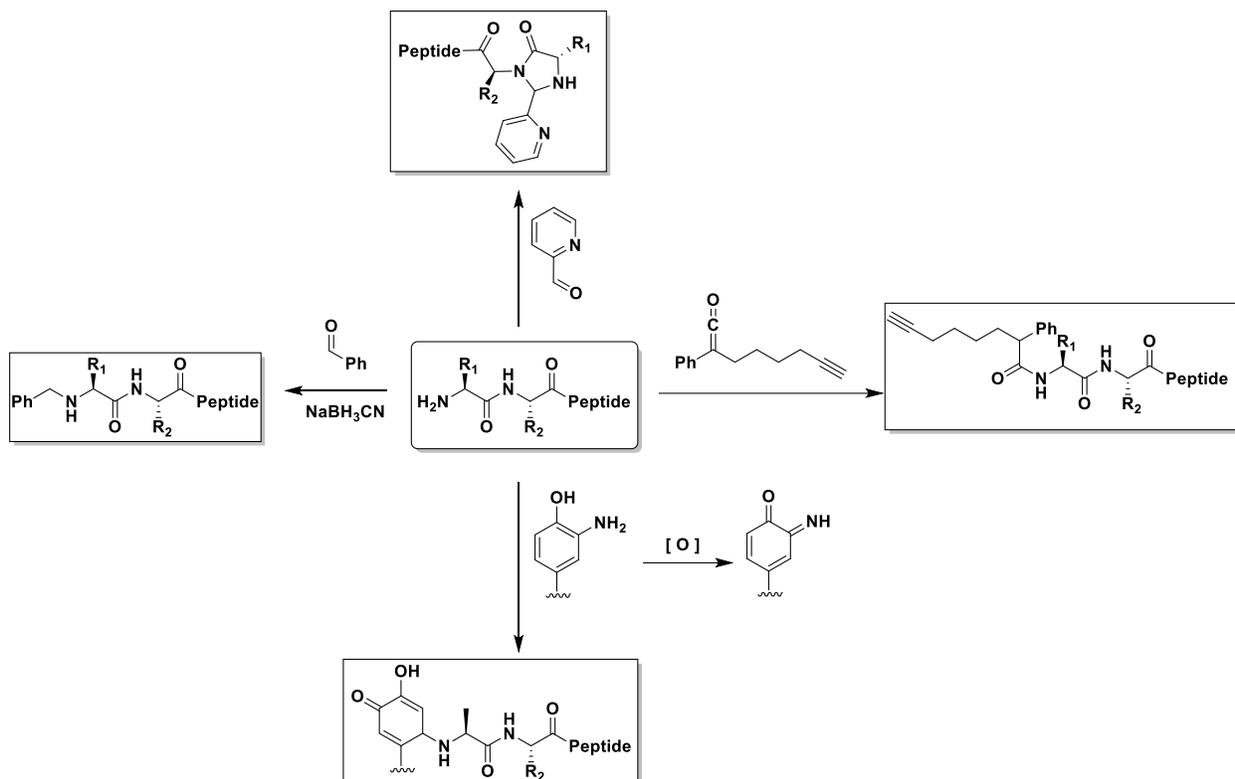


Figure 2. Selective technology for N-terminus labeling a) N-terminal tryptophan, cysteine, and proline. b) N-terminal serine. c) N-terminal glycine. d) N-terminal modification techniques for universal amino acids.

The current method to increase the detection sensitivity of bioconjugates involve the addition of mass sensitive probes but in a non-selective manner¹⁶. This leads to the formation of heterogeneous mixture which makes the analysis difficult. Multiple methods are usually needed to achieve all these goals, for example, one method is required for the site-selective modification and a second method for the addition of chemical tag that enhances mass sensitivity of the bioconjugate. As a result, a simple, one-step method capable of achieving all the above mentioned goals with the broad scope of structurally and chemically varied peptides and proteins

would be highly beneficial. This would enable efficient analysis of the modified peptides in a complex mixture and aid in the detection of low abundant peptides.

1.1.2 Lysine residue modification:

Lysine shows rich chemistry through its nucleophilic amine group and is abundant in various active and allosteric sites. Lysines also catalyze multiple reactions and regulate various biological processes. Moreover, lysines are frequent sites for posttranslational modifications and regulate the structure and functions of proteins. Lysines have a high frequency in human proteins (& 6% of all residues)¹⁶. Together, these properties make lysine residues desirable targets with covalent drugs; however, the low nucleophilicity of lysine as compared to cysteine makes selective targeting by covalent probes a highly daunting task. Several small molecule covalent ligands for selective proteome profiling of cysteine has been reported; however most of the small molecule electrophiles for lysines such as dichlorotriazines¹⁷, imidoesters¹⁸, 2-acetyl- or 2-formyl-benzeneboronic acids¹⁹, isothiocyanates^{20,21}, pyrazolecarboxamidines^{22,23}, sulfonyl fluorides^{24,25}, and vinyl sulfonamides²⁶ also react with other amino acids such as serine, tyrosine and cysteine (Figure 3). Sulfonyl acrylate reagent has been reported for the regio- and chemoselective labeling of lysine on pure protein by using low equivalent of the probe but it has a potential to react with cysteine during profiling (Figure 3). Recently, activated esters such as NHS-ester²⁷ and STPyne²⁸ have been explored for ligandability of lysine in the human proteome (Figure 3). Although > 9000 ligandable lysine sites have been discovered using NHSesters²⁹ and STP-esters, none of these probes act as a covalent ligand/inhibitor for lysine in cells because of poor hydrolytic stability. Therefore for the discovery of covalent ligands for lysines, new amine reactive chemotypes are needed that are stable to enzymatic and non-enzymatic hydrolysis, non-

cytotoxic and generate stable covalent adducts with lysine. These lysine selective, hydrolytically stable covalent probes would significantly expand the druggable content of the human proteome.

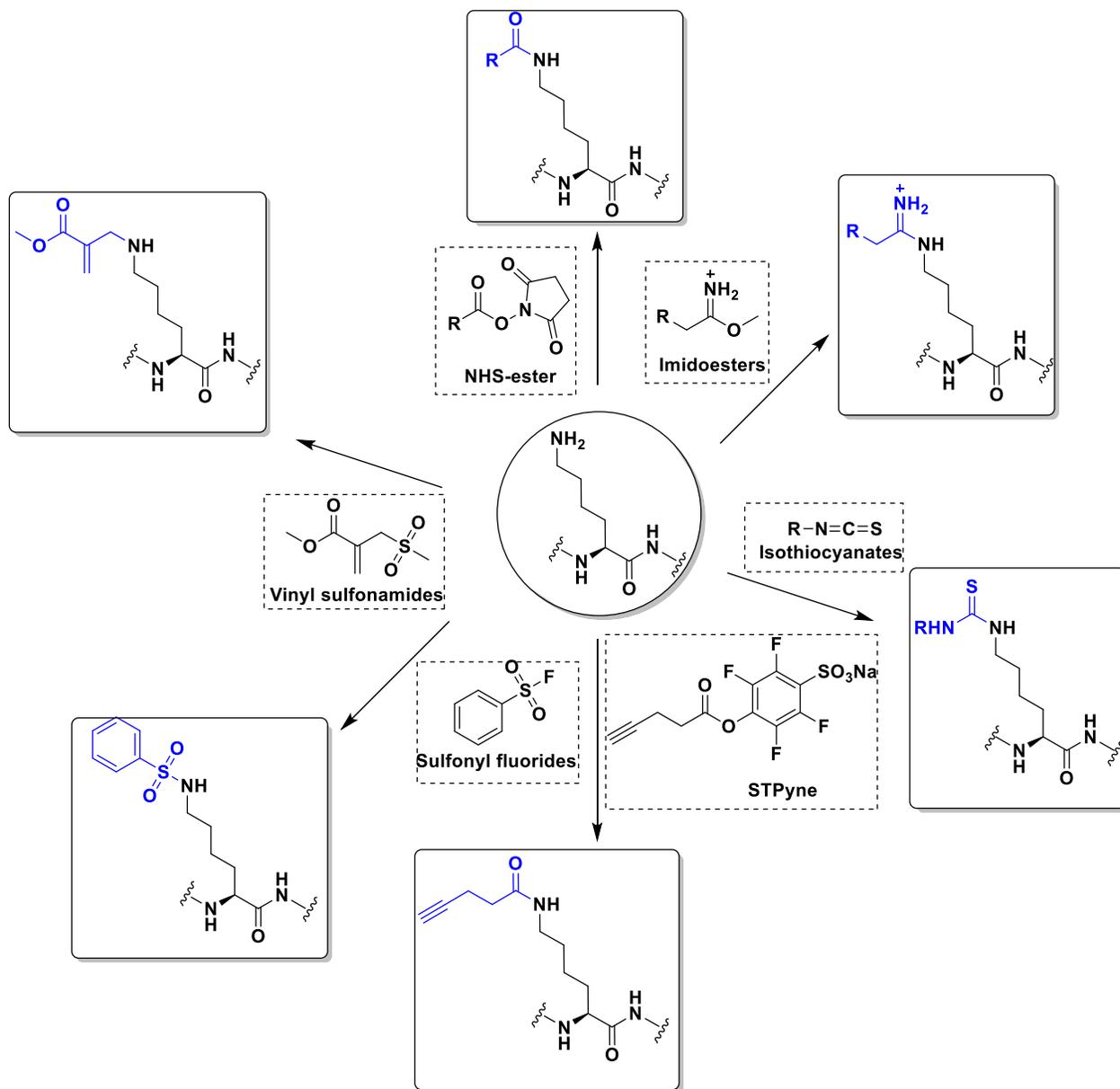


Figure 3. Reagent for chemo-selective modification of lysine

1.1.3 Cysteine side chain modification

Among the 20 natural amino acids, cysteine is especially attractive for site-specific chemical modification owing to its high nucleophilicity at physiological conditions and relatively low natural abundance³. It is often located in functionally important sites. In addition to stabilizing proteins through disulfide bonds, cysteine residues have diverse roles in metabolic processes, for example, in catalysis, allosteric regulation, and metal binding. Various electrophiles have been reported for tagging cysteines such as haloacetamides (IAA)³⁰, epoxides³¹, sulfonate esters³², chloro³³ and acyloxymethyl ketones³⁴, fluorobenzene³⁵, aryl halides³⁶, Michael acceptors^{37,38} and heteroaromatic sulfones^{39,40,41} (Figure 4). However, poor hydrolytic stability of probes or resulting conjugates, low reactivity and cross-reactivity with other amino acids lead to the development of several metal-free and transition metal-mediated cysteine bioconjugation approaches by Davis, Bernardes, Pentelute, Wong, and others⁴²⁻⁴⁵ (Figure 4). Bioconjugation reactions with cleavable linkers have recently gained considerable attention due to their wide application in many research fields such as protein immobilization, drug development, and proteomics. Only a few compounds are available for cleavable cysteine-specific modification, including Ellman's reagent^{46,47}, bromomaleimides⁴⁸, bromopyridazinediones⁴⁹, electron deficient acetylenes⁵⁰, 5-methylene pyrrolones⁵¹, 4-substituted cyclopentenones⁵² and recently discovered isoxazoliniums⁵³. Despite all these advances, only $\approx 17\%$ of cysteine^{54,55} in the entire proteome has been identified so far. Moreover, the resulting bioconjugates exhibit poor mass detection sensitivity limiting their applications in the identification of low abundant cysteine in the proteome. Therefore, there is a great need to develop new highly efficient cysteine selective modification methods with cleavable linkers using easily accessible reagents with high stability, distinct and tunable selectivity and reactivity that increases the mass detection sensitivity of the

resulting conjugates that would aid in the identification of low abundant cysteines in the proteome.

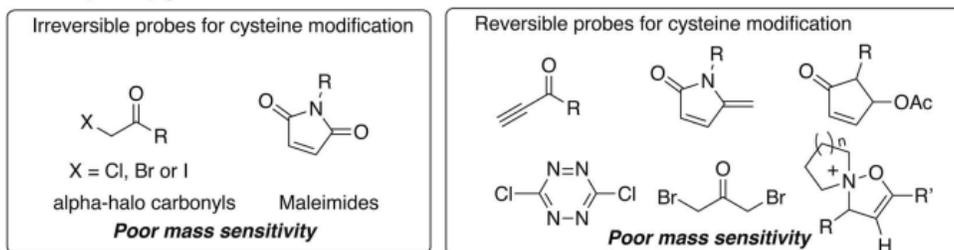


Figure 4. General cysteine labeling method

1.2 Results and Discussion

1.2.1 Screening of azoline, azole, and azolinium with standard peptide

In order to discover the new series of warheads that are hydrolytically stable and highly specific for the site-selective modification of the proteins, we have synthesized different types of methylthiol azole, azoline, and azolinium since they are stable under physiological conditions. Moreover, the reactivity of those probes is tunable based on a variety of heteroaromatic rings and methylation levels. (Figure. 5). With those probes in hand, we incubated those small molecules with the standard peptide FKVCF **2a**, followed by the LC-MS/MS analysis to identify the tagging site of the pentapeptide (Supplementary Figure. 1). We found four of our probes **1a**, **1i**, **1m**, and **1o** are selective to cysteine due to the high nucleophilicity of cysteine, probes **1b**, **1d**, **1n** are specific to lysine residue because the cysteine-adducts of those probes are reactive electrophiles towards lysines and unstable towards hydrolysis, **1c** and **1j** can react both with cysteine and lysine, the oxazoline (**Ox1**) probe is specific to N-terminus under optimized conditions. The rest of the probes do not show significant reactivity toward any amino acid residue due to the low aqueous solubility.

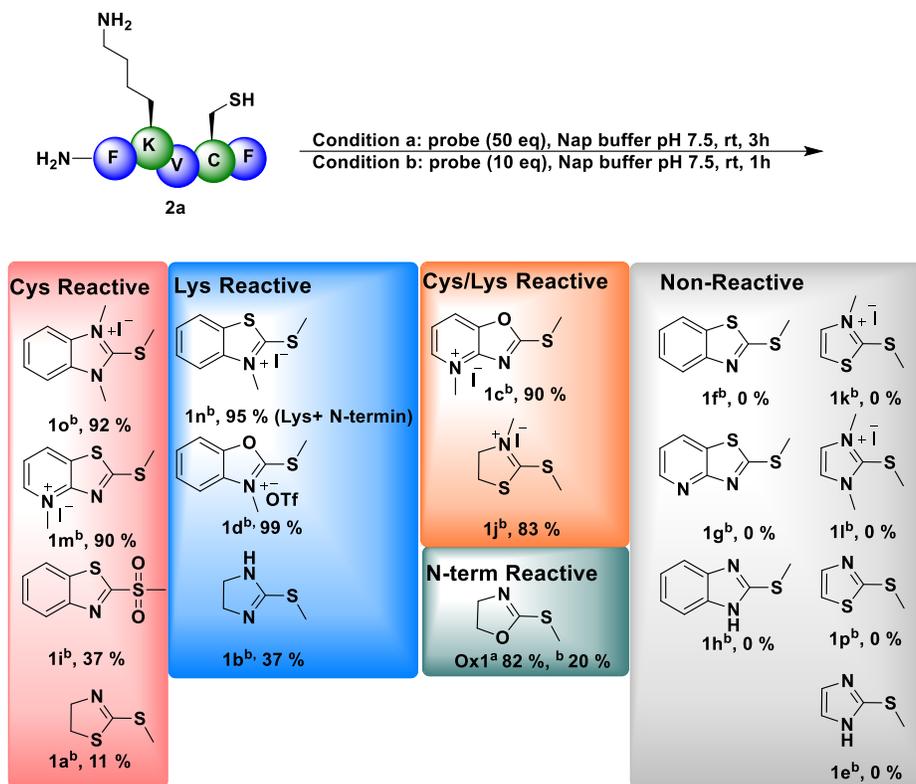


Figure 5. Screening of different azole, azoline, and azolinum probes with peptide 2a

1.2.2 Synthesis of Azoline, Azole, and Azolinum and their Analogs

Synthesis of Ox1, 1a, and 1b:

The simple azole thioethers⁵⁶ were synthesized from the corresponding amino alcohol (**1**) or diamine (**4**) with proper thio reagent (carbon disulfide or thiocarbonyldiimidazole) under basic conditions. The corresponding thione intermediate (**2**, **3**, **5**) was methylated by CH₃I to obtain the azole thioether (**Ox1**, **1a**, **1b**).

Synthesis of Ox2:

The N-terminal amine of L-serine was protected by tert-butyloxycarbonyl protecting group to produce intermediate **7**, then the amide coupling was performed without protecting of the free

alcohol⁵⁷. After installation of the alkyne group, the intermediate **8** was treated under acid condition to remove the Boc group. The deprotected intermediate was treated with CS₂ to obtain thione intermediate **9**, then the final methylation was achieved by MeI under basic conditions to produce the final product **Ox2**.

Synthesis of Ox5:

The alcohol intermediate **12** was synthesized by following Song's procedure⁵⁸, then the intermediate **12** was treated with DSC to generate the pyrrolidine carboxylate intermediate **13** for the amide bond formation. The intermediate **13** was coupled with Biotin ethylenediamine to obtain the **Ox5**.

Synthesis of Ox3, and Ox4:

The alcohol intermediate **12** was treated with TsCl to generate precursor **15**, then the substitution reaction of Ts intermediate with NaN₃ was carried out to obtain **Ox3**. **Ox4** was synthesized by the propargylation of the alcohol intermediate **12**.

Synthesis of 1j, 1l, and 1k:

The corresponding thioether (**1a**, **16**, **1e**) was refluxing with MeI to obtain the corresponding azolium (**1j**, **1l**, **1k**).

Synthesis of 1n and 1m:

Compound **1n** was synthesized according to Hara's procedure⁵⁹. The synthetic conditions for **1m** are modified from Bethge's procedure⁶⁰.

Synthesis of 1o-N3:

Compound **1o-N3** was synthesized from the commercially available 5-Nitro-2-benzimidazolinone. First, compound **20** was treated with MeI to afford the double methylated intermediates **21_a** and **21_b** as isomers. Without further purification, the nitrobenzene isomer mixtures was reduced under Fe/ammonium chloride condition to obtain the intermediate mixtures **22_{a,b}**. Next, the aniline intermediates **22_{a,b}** was treated with sodium nitrite under acid condition, followed by the addition of NaN₃ to obtain the aryl azide intermediates **23_{a,b}**. After the final methylation of intermediate **23_{a,b}**, compound **1o-N3** was generated as an exclusive product.

Synthesis of **1d** and **1c**

The synthetic conditions⁶⁰ for compound **1c** and **1d** were modified from Bethge's protocol.

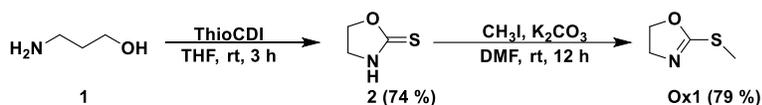
Synthesis of **1d-yne**:

Compound **31** was synthesized by Imaizumi's protocol⁶¹. The intermediate **32** was synthesized by the methylation of mercaptobenzoxazole intermediate **31**. The alcohol intermediate **34** was generated by the hydrolysis of ester intermediate **32** following by the reduction of the anhydride precursor⁶². After the propargylation of the alcohol intermediate **34**, the final methylation was achieved by MeOTf to obtain **1d-yne**.

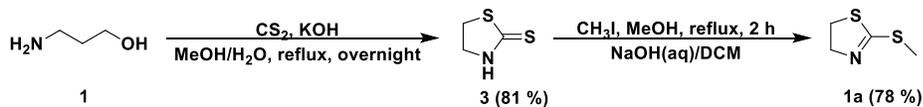
Synthesis of **1c-yne**:

The probe **1c-yne** was synthesized through mercapto formation following by the methylation of mercapto group and the propargylation of the pyridine.

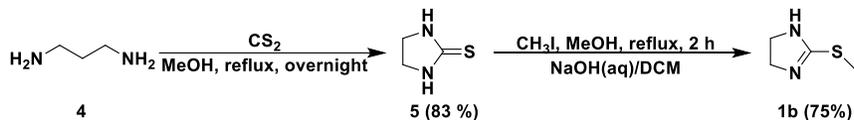
Synthesis of Ox1



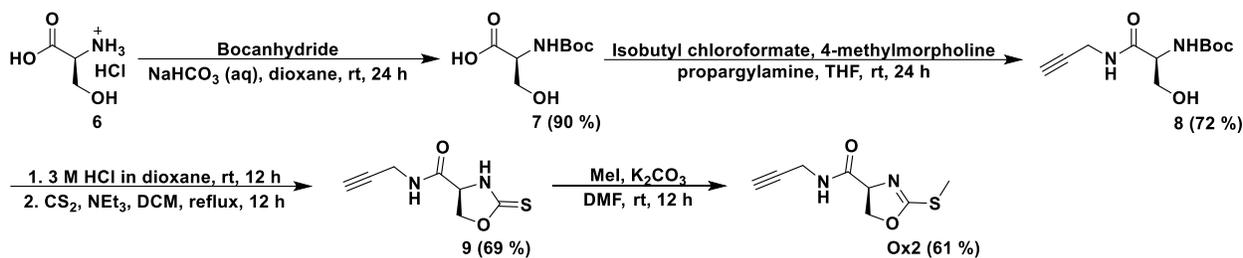
Synthesis of 1a



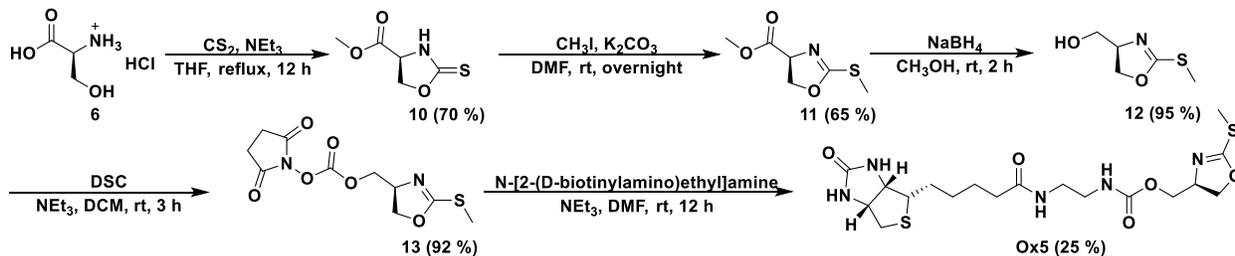
Synthesis of 1b



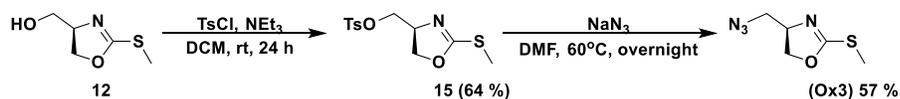
Synthesis of Ox2



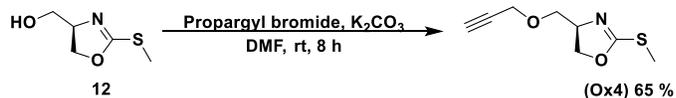
Synthesis of Ox5



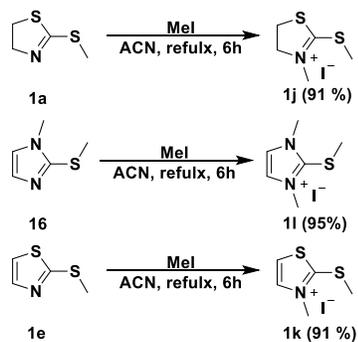
Synthesis of Ox3



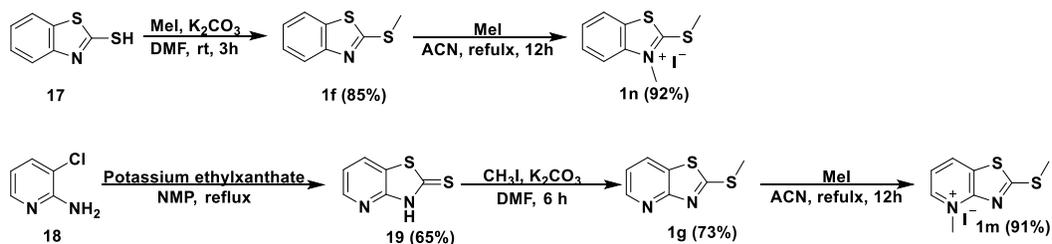
Synthesis of Ox4



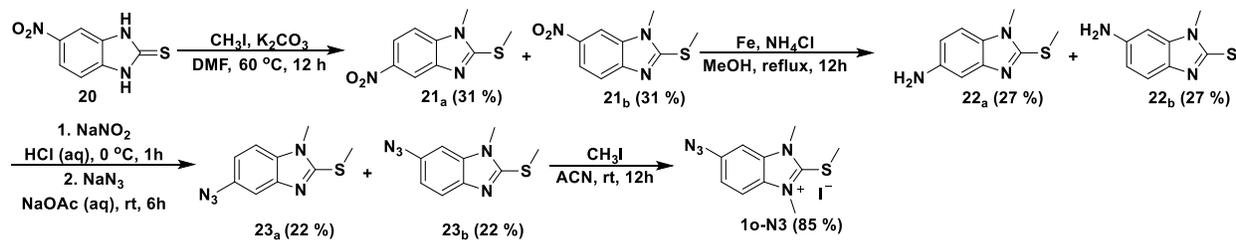
Synthesis of 1j, 1l, and 1k



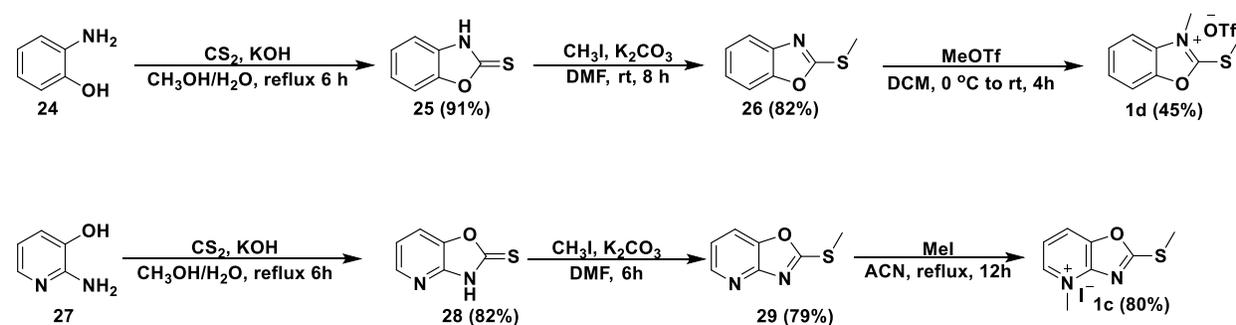
Synthesis of 1n and 1m



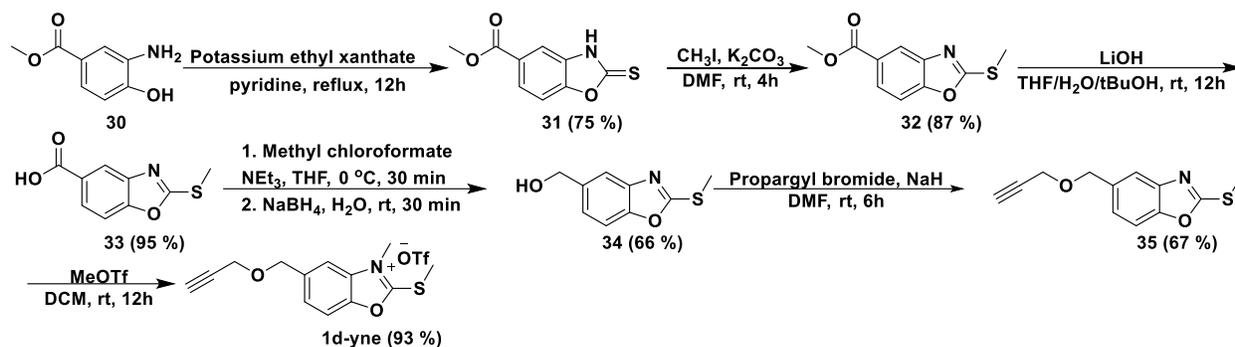
Synthesis of 1o-N3



Synthesis of 1d and 1c



Synthesis of 1d-yne



Synthesis of 1c-yne

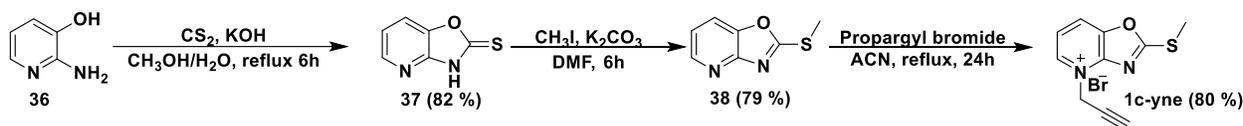


Figure 6. Synthesis of various azole, azoline, azolinium, and their analogs.

1.2.2 One-Step Azolation Strategy for Site-and Chemo-Selective Labeling of Proteins with Mass-Sensitive Probes

1.2.2.2 Studying Universal Sequence Compatibility of N-terminus Azolztion

Next, we evaluated universal sequence compatibility of N-terminus azolztion chemistry by screening a panel of peptides with varying reactive amino acid residues at the N-terminus. The incubation of peptides (XAF, 2c–2j X = D, E, F, G, M, P, T, Y (6.25 mM) with 2- methylthio oxazoline, **Ox1** (312.5 mM) in 10 mM phosphate buffer at pH 7.5 for 3 h yielded N-terminus Ox1- modified peptides 3c–3j with 50-99 % conversion as determined by HPLC and LC-MS (Figure 7, Supplementary Figure. 2). The conversion of azolation reaction with different N-terminal amino acid was determine by the steric hindrance of N-terminal amino acids. No modifications of the reactive side-chains were observed under the reaction conditions thus reaction is highly specific toward the N-terminus.

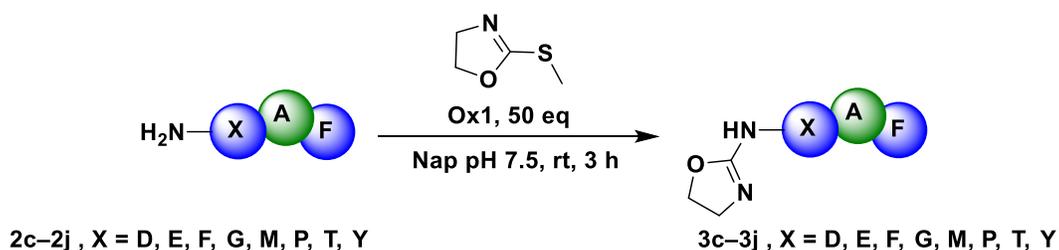


Figure 7. Compatibility of N-terminus azolztion with universal sequence. Various of N-terminus amino acids were used to confirm the high compatibility of azolztion reaction.

The N-terminal oxazolation of peptides GAF **2f** and APF **2k** with glycine at the N-terminus and proline at the second position proceeded smoothly to generate N-terminal oxazoline-modified peptides (**3f-3k**) in a quantitative manner (Figure 8). This is in contrast to the **2PCA** method resulting in the double modification with glycine at the N-terminus and no modification of a peptide with proline at the second position. These studies showed that oxazolation is compatible with all nature of amino acid sequences thus exhibit universal sequence compatibility (Figure 8, Supplementary Figure. 3)

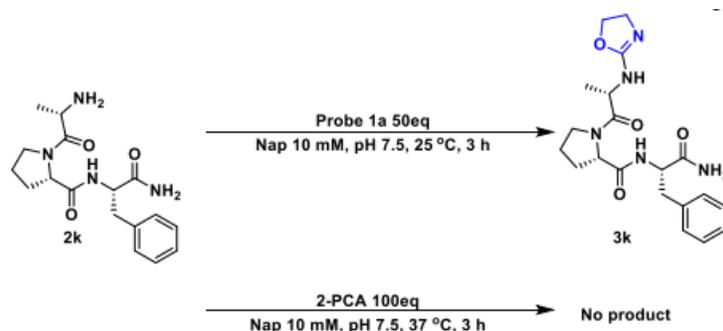


Figure 8. Comparison of oxazolation Vs 2- PCA methods using peptide APF 2k. The high conversion of azolztion with APF (80 %) showed the high compatible property of the azolztion reaction. The 2-PCA method is incompatible with peptide or protein, which has a proline at the second position.

1.2.2.3 Protein modification-reaction optimization

With those results at hand, we started our initial investigation on protein myoglobin Mb. Detailed optimization studies with myoglobin Mb showed that the best conditions are 2-methylthio oxazoline **Ox1** in 10 mM phosphate buffer at pH 7.5 and 25 °C for 12 h (Figure 9, Table 1, Entry 2). Long reaction time is needed for high labeling yield (70-99%) but moderate amounts of labeling are observed even in shorter reaction time (6 h) (50–80%). The concentration of protein (0.75 mM–3 mM) did not influence the extent of modification (Figure 9, Table 1, Entry 3, 4, and 5). Excellent conversion (70%) to N-terminal oxazoline-modified myoglobin Ox1-Mb was obtained at the more neutral pH as confirmed through protein digestion followed by the MS/MS analysis (Supplementary Figure. 4). Low pH 6.5 reduced the conversion to modified protein (43%, Figure 9, Table1, Entry 1). An increase in pH 8.5 increased the percentage conversion to modified protein (52%) but also resulted in the lower selectivity, leading the small amounts of modification on the side chain of lysine (30%) (Figure 9, Table 1, Entry 3). The increase in the probe equivalents from 50 to 200 lead to the multiple modifications of protein Mb. High site-selectivity for the N-terminus was observed with 50 equiv of the probe (Figure 9, Table 1, Entry 2).

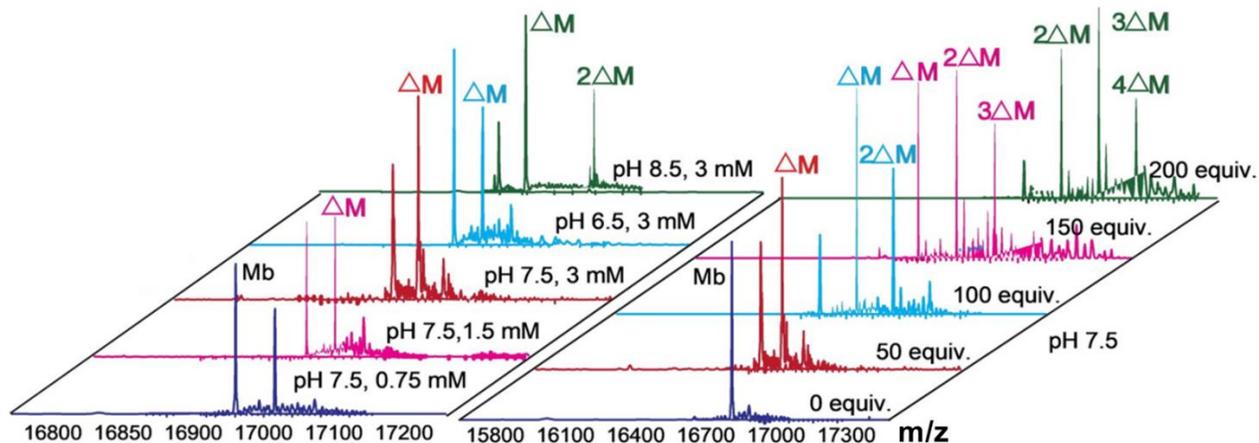


Figure 9. HR-MS of oxazolation on protein Myoglobin Mb. Optimization studies with myoglobin Mb revealed that the most efficient conditions are 2-methylthio oxazoline 1 a Ox1 in 10 mM phosphate buffer at pH 7.5 and 25 8C for 12 h. Long reaction time is needed for high labeling yield (70-99%) but moderate amounts of labeling are observed even in shorter reaction time (6 h) (50–80%). The concentration of protein (0.75 mM–3 mM) did not influence the extent of modification.

Table 1 Optimization of the oxazolation on protein Myoglobin Mb

^a Entry	Equiv. of probe	Temperature	pH	Total volume (conc.)	^b Conversion (%)
1	50	25 °C	6.5	200 μL (3 mM)	43 % mono
2	50	25 °C	7.5	200 μL (3 mM)	70 % mono
3	50	25 °C	8.5	200 μL (3 mM)	52 % mono, 30 % double,
4	50	25 °C	7.5	400 μL (1.5 mM)	65 % mono
5	50	25 °C	7.5	800 μL (0.75 mM)	45 % mono
6	100	25 °C	7.5	200 μL (3 mM)	51 % mono, 33 % double
7	150	25 °C	7.5	200 μL (3 mM)	35 % mono, 37 % double, 25 % triple
8	200	25 °C	7.5	200 μL (3 mM)	6 % mono, 31 % double, 42 % triple, 21 % quadruple
9	50	40 °C	7.5	200 μL (3 mM)	13 % double, 29 % triple, 36 % quadruple, 22 % quintuple

^aCondition: Protein Mb (0.6 μmol) in 10 mM phosphate buffer of pH 7.5 (200-800 μL) and 1a, Ox1 (50-200 equiv.) was incubated overnight. ^bThe conversion was calculated based on

the relative peak intensity of native protein and labeled protein in the deconvoluted mass spectrum.

1.2.2.4 Synthesis of functionalized Oxazoles for N-terminal modification

1.2.2.5 Protein Scope for N-terminal Oxazolation

We used several functionalized oxazolines (50-300 equiv) to oxazolate a variety of commercially available proteins, including cytochrome C (Cy), alpha-lactalbumin, ubiquitin (Ub), and insulin (Ins) (3 mM). The reactions provided high selectivity and good to excellent yields (50-99%) (Figure 10). To our surprise, the ubiquitin with seven lysine residues produced a single modification product at the N-terminus. Due to reactions at both N-termini, double modification products were produced by insulin with two N-termini (Figure 10). Proteins with various functional group handles will be modified to varying degrees, depending on the size of the handles and how accessible the N-terminus is. Higher labeling efficiency was seen when the handle size was smaller, with > 99 percent labeling of Cytochrome (Cy) with **Ox2**, > 99 percent labeling of Insulin (Ins) with **Ox2** and **Ox4**, and 90% labeling of Myoglobin (Mb) with **Ox2** (Figure 10). Because **Ox5** is bulky, proteins are labeled less frequently (between 40 and 70 percent) (Figure 10). These illustrations showed that the reaction is adaptable to a variety of molecular weights, three-dimensional architectures, and structurally essential disulfide connections. The high nucleophilicity of the cysteine thiol group at neutral pH makes it extremely difficult to modify proteins selectively when cysteine is present. In numerous N-terminal bioconjugation techniques, the cross-reactivity with cysteine is frequently seen. When reduced insulin with six free cysteines was used to oxazolate proteins, no such modification of cysteine was seen (Supplementary Figure. 5). The unique feature of our approach is that it leads

to the single-site N-terminal modification of proteins in the presence of highly nucleophilic cysteines under physiological conditions.

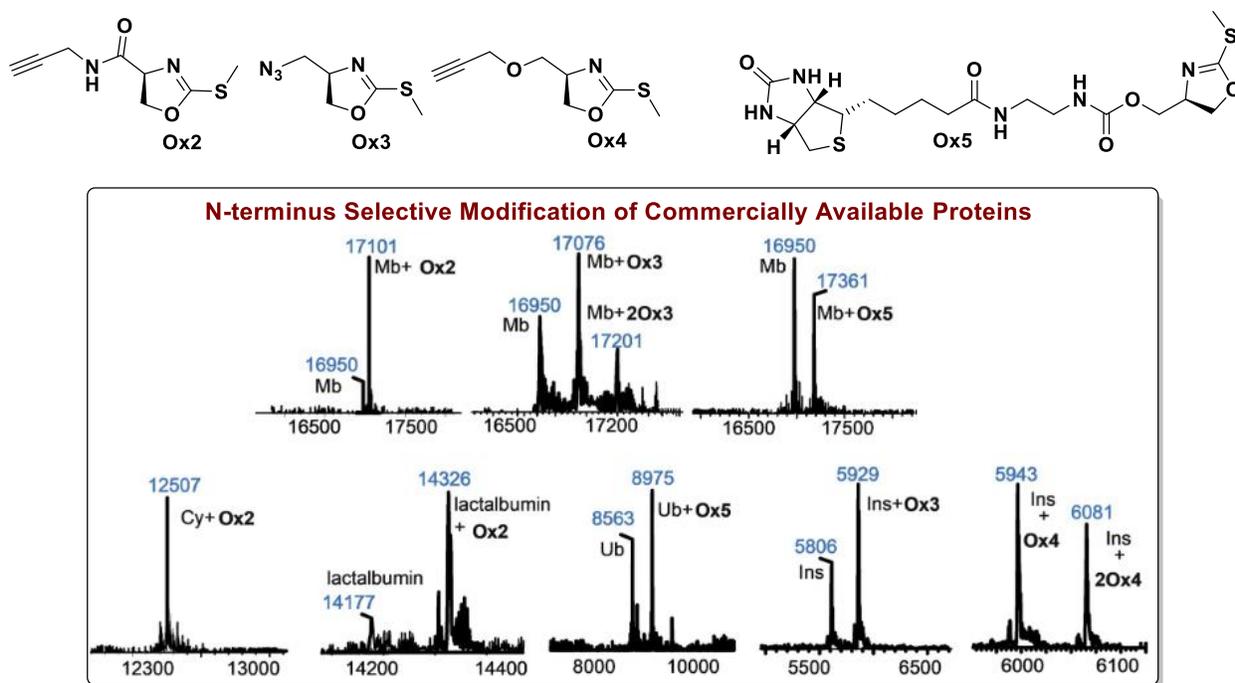


Figure 10. Site-specific attachment of oxazoline and its derivatives to a variety of different proteins. Conditions for 3 mM protein, 50–300 equiv oxazoline, 10 mM phosphate buffer at pH 7.5 and 40 °C. The samples were incubated for 12 h. The conversion with the N-terminus depends on the size of affinity targets; for the small affinity tags Ox2-Ox4, we observed a higher conversion (> 90 %); however, the bulky affinity tag Ox5 showed moderate reactivity (ca. 50 %).

1.2.2.6 Rate and Stability Studies of Oxazolation

Time-course studies on the linear peptide GAF-OMe **2 f-OMe** were conducted to understand better reaction rates and the products formed. In order to conduct quantitative monitoring for this experiment, samples were injected for HPLC analysis at regular intervals. Over 60 min, the oxazolation of tripeptide GAF-OMe **2 f-OMe** (2 mM) with 2-methylthio oxazoline **Ox1** (100

mM) at pH 7.5 was observed (Figure 11). These results demonstrate that the initial rate of production of Ox1-GAF-OMe conjugate **3 f-OMe** is quite quick, with 70% conversion in just 10 min (Figure 11). The rate studies of the oxazolation between tripeptide GAF-OMe **2 f-OMe** (1 mM) and 2-methylthio oxazoline **1 a** (1 mM) at pH 7.5 were monitored over 60 min and showed that the reaction follows a second-order rate constant ($k = 1.66 \times 10^{-2} \text{ M}^{-1} \text{ S}^{-1}$) for this process (Figure 11). For bioconjugation reactions, the stability of protein conjugates is a critical concern. So, for 48 hours at 25 °C and 37 °C, we tested the stability of the oxazole-peptide conjugate Ox1-GAF-OMe **3 f-OMe** in aqueous buffers with pH ranging from 3.5 to 10.5. High stability of Ox1-GAF-OMe **3 f-OMe** at various pH settings was shown by HPLC analysis (Figure 12).

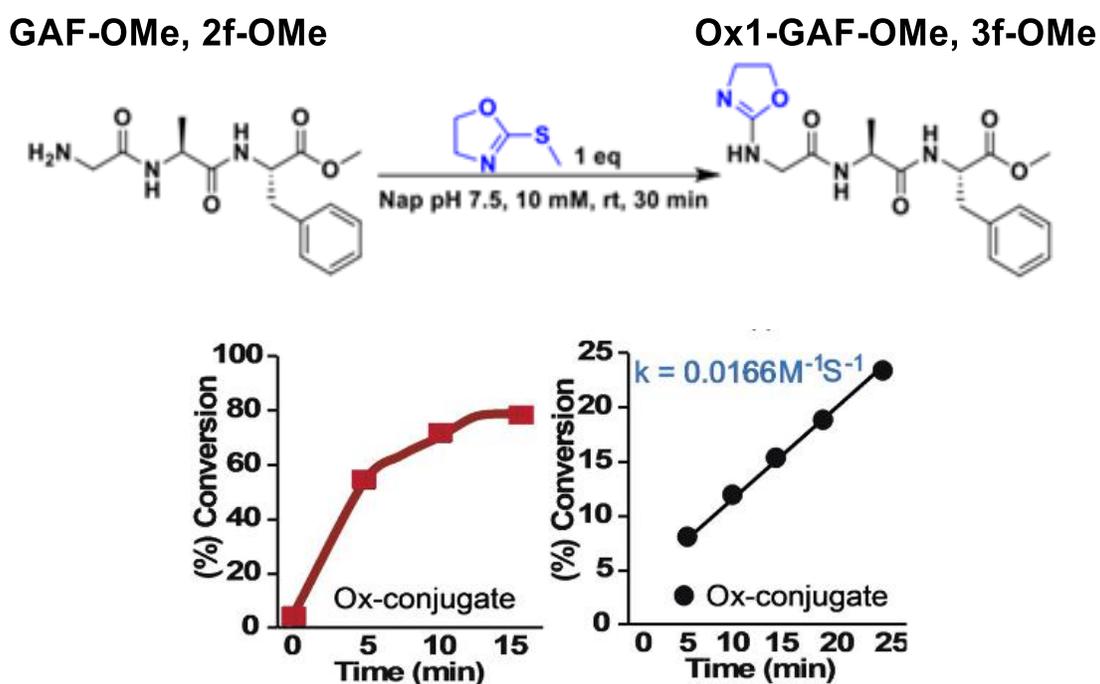
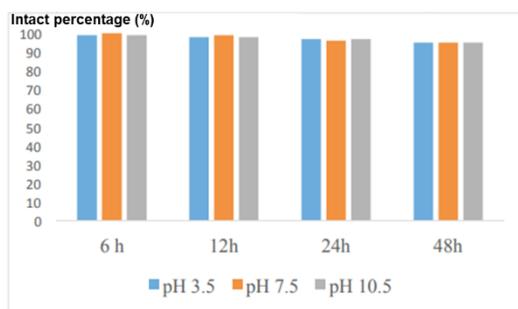


Figure 11. Rate study of azolation bioconjugation. Rate studies for synthesis of Ox-GAF-OMe conjugate **3 f-OMe**. The oxazolation of tripeptide GAF-OMe **2 f-OMe** (1 mM) with **1 a** (1 mM) at pH 7.5 follows a second order rate constant. Time-course study with 50

equivalents of Ox1 (left plot), rate studies to determine the reaction rate by using 1 equiv. of Ox1.

a. 25 ° C under different pH conditions



b. 37 ° C under different pH conditions

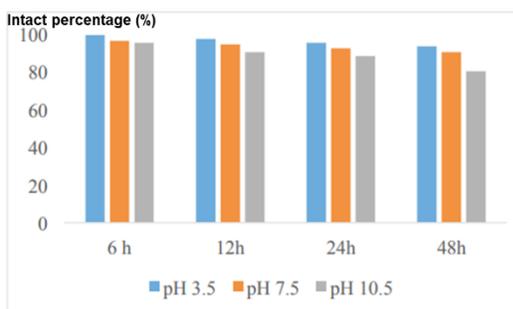


Figure 12. Stability study of Ox1-GAF-OMe under different pH conditions at different temperatures; a. 25 °C. b. 40 °C. Ox-GAF-OMe 3f-OMe peptide (6.25 mM) was incubated in 10 mM phosphate buffer (Nap) at different pH ranging from 3.5 to 10.5 at room temperature and at 37 °C. A sample (50 µL) was taken from the mixture and directly injected into HPLC. The reaction was monitored by injecting samples in HPLC after regular intervals of time 6 h, 12 h, 24, and 48 h. The small amount of modified peptide Ox-GAF-OMe 3f-OMe was degraded in pH 10.5 after 48 h at 37 °C.

1.2.2.7 Labeling of Proteolytic Fragments in the Complex Mixture

Oxazolation was carried out with both **Ox1** and **Ox4** on a combination of peptides with different amino acids at the N-terminus (XAF (2e-2f, 2l-2n) X = F, G, A, V, L) to see if it could be used to simultaneously tag the N-termini of several peptides and proteins in a complex system. According to the results of the quantitative MS analysis, all of the peptides had been fully N-terminally oxazolated (Figure 13, 14). In proteomics research, the capacity to tag every N-terminus of the complicated mixture of peptides could be quite useful.

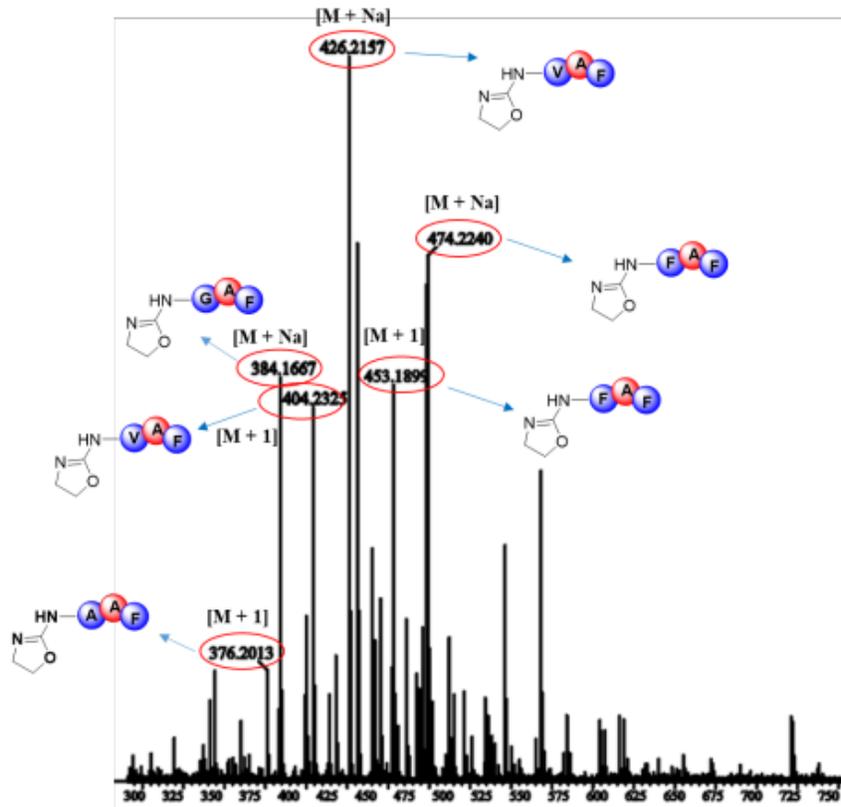
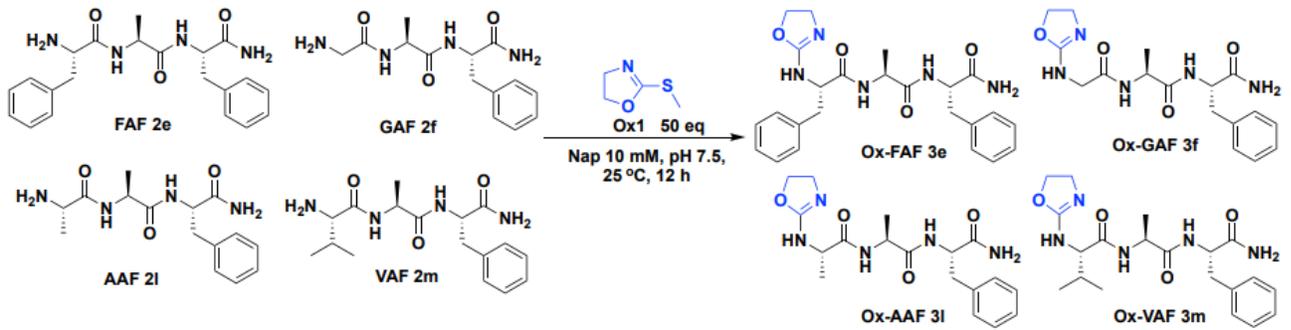


Figure 13. The ESI-MS spectrum of the reaction mixture for the Ox1-FAF, Ox1-GAF, Ox1-AAF and Ox1-VAF N-terminal modification of mixed peptide fragments.

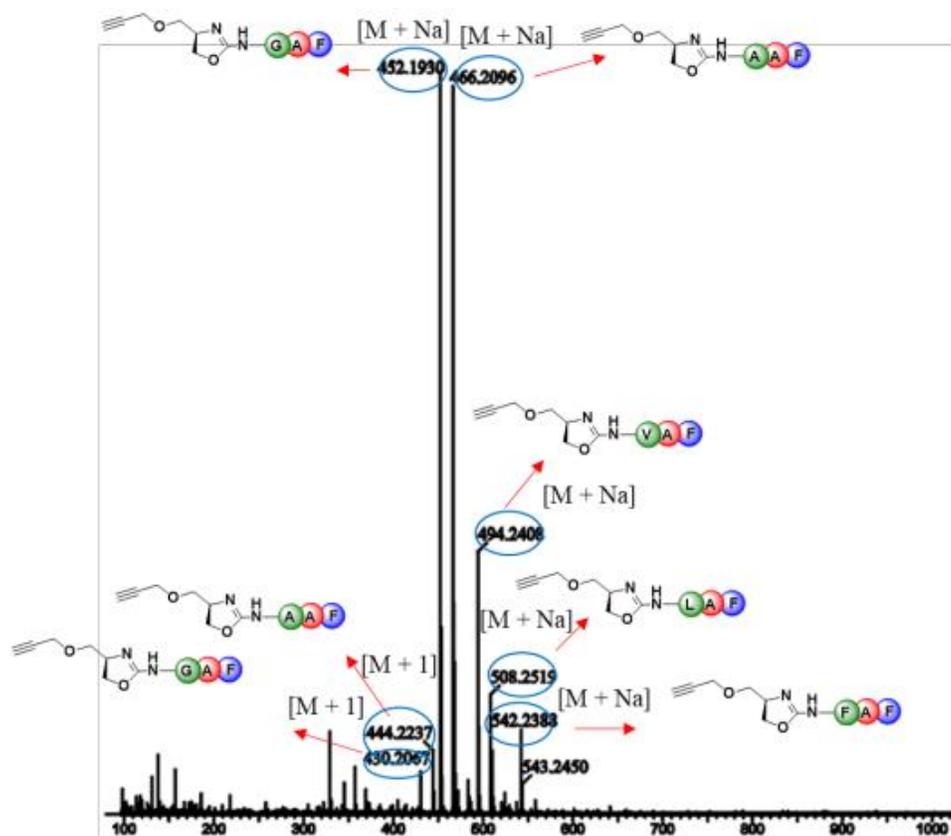
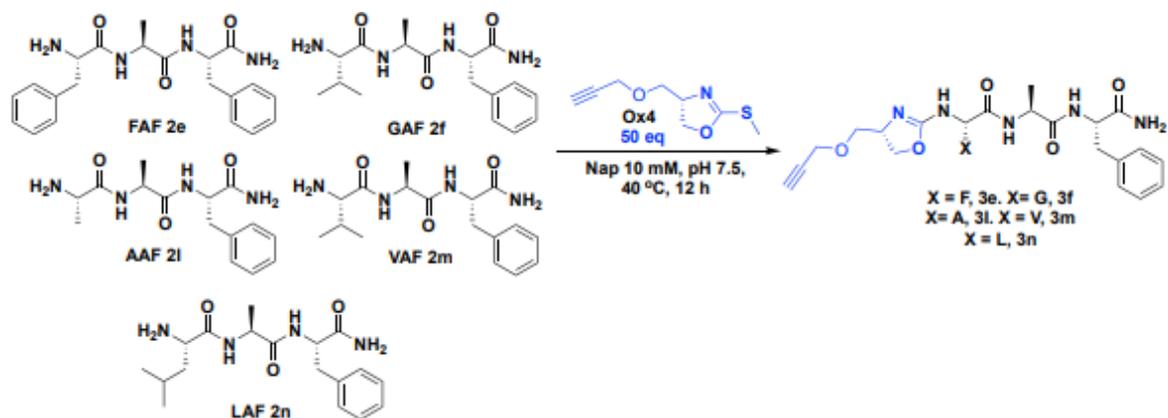


Figure 14. Selective N-terminal labelling of all the peptides by oxazoline Ox4 in the complex mixture of a variety of peptides.

1.2.2.8 Oxazoline as Mass Sensitivity Booster

The difficulty characterizing the resultant bioconjugates is one of the main drawbacks of existing N-terminal labeling techniques. LC-MS is frequently used to determine labels and discover the

site of modification. The hydrophobic tags obstructing the free N-terminus are to blame for the ineffective ionization of N-terminally labeled peptides. This restriction manifests itself more clearly in complex mixes. Therefore, approaches that can improve the detection of tagged proteolytic fragments are needed urgently. Including chemical tags that boost the ionization of the labeled peptides in the MS is one method of achieving this goal. Some strategies for improving the ionization ability of peptides have been published, but they modify various amino acids, thus making the analysis difficult. Our method is distinctive in that it increases the ionization of the labeled fragments by introducing the oxazoline group, which has a sp² nitrogen, in addition to selectively labeling the N-terminus. In order to determine the impact of the oxazoline tag on mass sensitivity, we mixed equal concentrations of several unlabeled peptides with methyl ester protected C-termini and rich in hydrophobic groups (XAF-OMe (2e, 2 fOMe, 2lOMe-2mOMe), X = F, G, A, and V) with corresponding N-terminal oxazoline-labeled peptides (Ox-X (Figure 15). MS analysis showed that the oxazoline moiety affects the detection sensitivity (Figure 15). Regardless of the peptide sequence, oxazoline-tagged peptides (Ox-XAF-OMe, 3e, 3f, and 3lOMe-3mOMe) significantly improved signal enhancement (Figure 15).

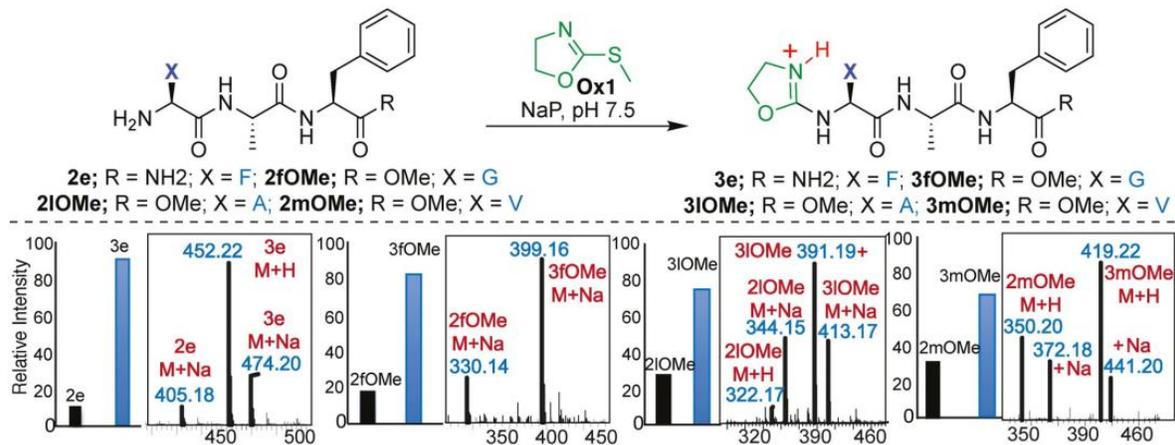


Figure 15. Oxazoline tagging of hydrophobic N-terminal peptides increases the mass detection sensitivity as compared to untagged peptides. The oxazoline-peptide bioconjugates, showed remarkable improvement in the signal enhancement irrespective of the unmodified sequence of the peptides.

Inspired by these results, we investigated the mass sensitivity of proteins and their oxazoline-labeled bioconjugates. In order to investigate the potential of oxazoline as a sensitivity enhancer, we used cytochrome C and myoglobin. In this study, myoglobin and cytochrome C were digested with cyanogen bromide and then incubated with 2-methylthio oxazoline, **Ox1**. Without any purification, the proteolytic fragments before and after the N-terminal oxazoline tagging were analyzed by MS (Figure 16, Supplementary Figure. 6). Low MS sensitivity was observed for unlabeled proteolytic fragments of myoglobin, and some fragments were not visible (Figure 16). All of the myoglobin tagged proteolytic fragments were found to have remarkably high mass intensities (Figure 16). We saw a significant difference between the myoglobin proteolytic fragments that were oxazoline-labeled and those that weren't. Very low MS sensitivity was observed for all the unlabeled myoglobin proteolytic fragments. For the labeled proteolytic fragments of myoglobin, all the fragments were found to have remarkably high mass intensities (Figure 16). These investigations demonstrated increased sensitivity of the tagged proteolytic

fragments in the complex mixture, making them very significant and overcoming the detection limitations of the current N-terminal bioconjugation approaches. The ability to modify all the N-termini of digested proteolytic fragments in the complex mixture could be precious in proteomics studies. We also demonstrated that the N-terminally modified myoglobin by **Ox1** greatly improved in terms of mass sensitivity when compared to the unmodified intact myoglobin (Supplementary Figure. 7). Additionally, the N-terminal oxazolation technique could be utilized to identify the distinctive proteolytic fragments produced by chemotherapy-induced cell death, leading to the identification of novel cell death biomarkers.

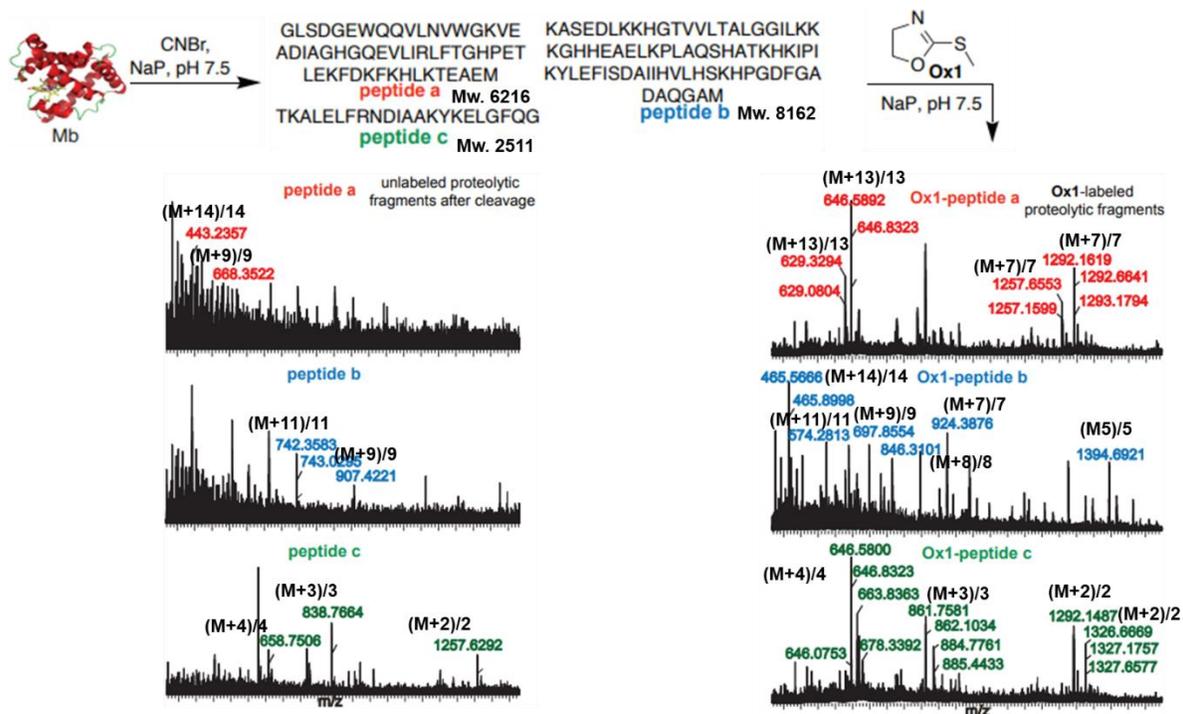


Figure 16. All the proteolytic fragments of myoglobin, peptides a, b, and c are more ionized and clearly visible after tagging with oxazoline. Oxazoline tagging improves the detection of proteolytic fragments in a complex mixture. Proteolytic fragment a is barely visible after cleavage of myoglobin (left MS trace). All the proteolytic fragments of myoglobin, peptides a, b, and c are more ionized and clearly visible after tagging with oxazoline (right MS trace).

1.2.2.9 Unmodified Myoglobin and Modified Myoglobin Bioactivity Assay

The activity assay of Ox1-Mb for its capacity to convert 2,3-diaminophenazine from o-phenylenediamine with hydrogen peroxide further demonstrated the addition of oxazoline to the N-terminus did not affect the activity of Mb (Figure 17). All of these findings confirmed that the structure of modified protein and bioactivity were preserved as oxazolation placed a probe specifically at a single location.

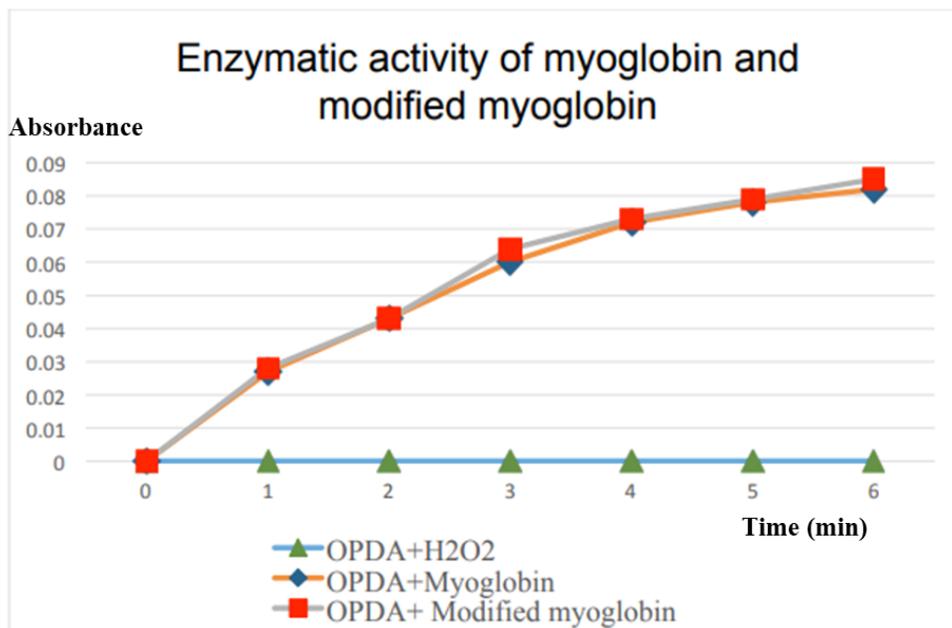


Figure 17. Enzymatic assay of myoglobin activity before and after the labeling was checked by oxidation of o-phenylenediamine with hydrogen peroxide. Oxidation of o-phenylenediamine (OPDA) to 2,3-diaminophenazine was monitored at 426 nm (A₄₂₆) using nanodrop. Citric acid-Na₂HPO₄ buffer (pH 5.6) was prepared by mixing of 0.1 M citric acid and 0.2 M Na₂HPO₄. All solutions were made in aqueous buffer. The myoglobin, oxazoline labeled myoglobin (Ox1-Mb), ophenylenediamine, and hydrogen peroxide (30 %) solutions were used for the assay.

1.2.2.10 Conclusion

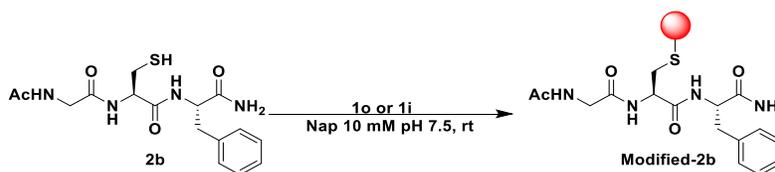
In conclusion, we have created a unique oxazolation strategy, a one-step method for selectively tagging proteins with mass-sensitive probes in physiological conditions. Due to the enhanced availability of deprotonated alpha amino groups and the lack of interference from lysine and highly nucleophilic cysteine side chains, this technique is very selective for the N-terminus. Although the nature and order of amino acids can have an impact on some N-terminal approaches, no such effect was observed with this method. The selective labeling of several

peptides and proteins with various amino acid compositions, as well as the reaction with proteolytic fragments in complicated combinations, as illustrated in Figures, provide excellent evidence of the potency of oxazolation strategy. The N-terminal labeling methods often suffer from low detection sensitivity of the resulting bioconjugates due to the blockage of the free N-terminus by the hydrophobic group, and multiple derivatizations are needed for the complete analysis. The N-terminal oxazolation produced bioconjugates with a remarkably high level of detection sensitivity. This is caused by the oxazoline's sp²-hybridized nitrogen, which improves the ionization of the tagged peptides. This method offers a fantastic tool for clearly identifying bioconjugates in a complicated combination. One of the main benefits of our approach is that, unlike other mass-sensitive boosters, it performs site-selective modification of the intact protein at the N-terminus, increasing its mass sensitivity without reducing its activity. Our method differs from previous mass sensitivity booster techniques in that it is extremely selective and produces only one modified fragment, as opposed to the combination of several modified fragments produced by the use of commercial mass sensitivity reagents because of their non-selective character. It is quite challenging to analyze the mass data when it contains various modifications. The excellent stability of the N-terminally oxazoline-modified peptides and proteins over a range of pH levels demonstrates the potential applicability of this chemistry in the conjugation of biological probes and pharmaceutically active molecules. We believe that this technique will be extremely helpful in many scientific fields due to its straightforward setup, chemoselective character, use of readily derivatized mass sensitivity boosters, and effectiveness in labeling proteolytic fragments from a complicated mixture.

1.2.3 Tunable heteroaromatic azoline thioethers (HATs) for cysteine profiling

1.2.3.1 Rate and Stability Study of HAT Probe

In order to monitor the reaction between the peptide Ac-GCF **2b** (3 mM) and the HAT probe **1o** (5-25 equiv.) under optimal conditions (NaP, pH 7.5, 10 mM), we observed the reaction throughout time at regular intervals (Figure 18). The formation of the coupling product was analyzed using HPLC and MS. With 25 equivalents of the **1o** probe, the reaction progressed quickly, and more than 80% of the conversion to the cysteine modification product was observed in 5 min (Figure 18). The comparison experiment with IAA (25 equiv.) showed a decreased product formation (30%) in 5 min. In order to understand how the substituent affects the rate of the reaction, we lastly conducted rate tests using the azide derivative of **1o**, (**N3-1o**). A similar reaction rate to **1o** was found in the reaction with **N3-1o** (25 equiv.). (Figure 18). Our initial theory on the adjustment of the core structure to get more predictable reactivity with different probe-derivatives is confirmed by the same reactivity profile of **1o** and **N3-1o**. Then, in a buffer solution (NaP, pH 7.5, 10 mM) at room temperature, we evaluated the kinetics of the reaction between HAT probe **1o** (0.973 mM) and its oxidized sulfone analog **1i** (0.973 mM). The findings indicated that **1o** is 10 times more reactive than **1i** ($k = 23.43 \text{ M}^{-1} \text{ S}^{-1}$) at $k = 236.77 \text{ M}^{-1} \text{ S}^{-1}$ (Figure 18, Supplementary Figure. 8).



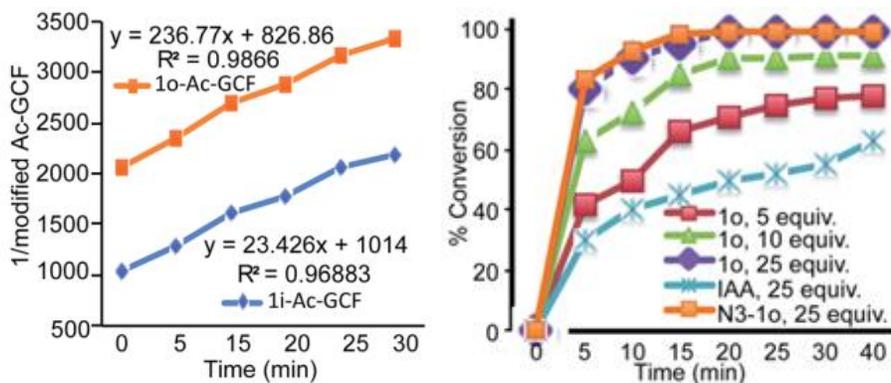


Figure 18. Rate study of 1o, N3-1o, and IAA with peptide 2b. Observed rate of the modification of cysteine of the peptide Ac-GCF 2b (0.003 mM) with 1o (5–25 equiv.), its azide analog N3-1o (0.6 mM, 25 equiv.) and IAA (0.6 mM, 25 equiv.) in 10 mM phosphate buffer (pH 7.5, 25 C) at different time intervals (Left). Kinetics study comparison of probe 1o (0.973 mM) and 1i for labelling peptide AcGCF 2b (0.973 mM) showed a 10 fold faster rate of probe 1o for cysteine modification as compared to sulfone analog 1i (0.973 mM) under physiological conditions (NaP (10 mM), pH 7.5, 25 C) (Right).

Highly reactive probes are typically thought to have low hydrolytic stability. Therefore, we compared the HAT probes 1o and N3-1o to the sulfone analog 1i in order to test their stability. The incubation of these probes (38.75 mM) in aqueous phosphate buffer (pH 7.5) at room temperature was observed periodically by HPLC. Surprisingly, the less reactive 1i probe degraded by 24 percent over the course of 6 hours whereas the more reactive **1o** and **N3-1o** probes were relatively stable. Probes **1o** and **N3-1o** exhibited hydrolysis of just 14–17% in 24 hours, whereas 1i showed hydrolysis of 40% in 24 hours (Figure 19). According to these investigations, HAT probes 1o and N3- 1o are hydrolytically more stable than 1i and demonstrate strong reactivity and excellent selectivity towards cysteine.

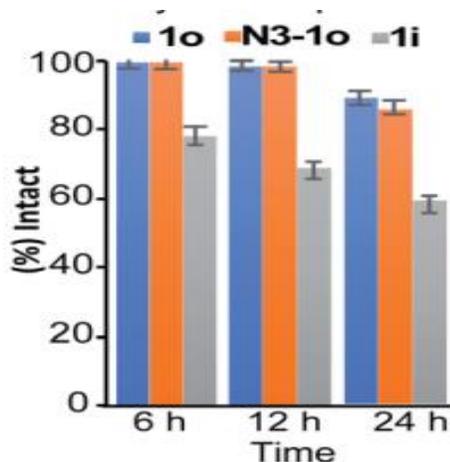
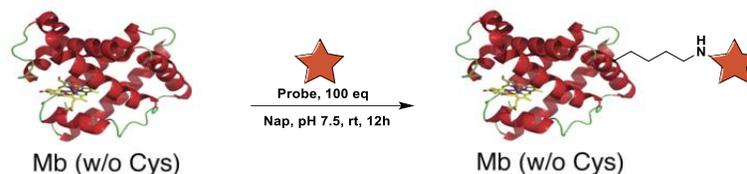


Figure 19. Stability study of probe 1o, N3-1o, and 1i. High stability of probe 1o (38.75 mM) and its azide analog N3-1o (38.75 mM) as compared to sulfone probe 1i (38.75 mM) under physiological conditions (NaP (10 mM), pH 7.5, 25 C).

1.2.3.2 Protein Modification with HAT

Previous studies with carbon electrophiles such as chloroacetamide and sulfonate esters showed that the solution reactivity of electrophiles with peptides is often not predictive of reactivity observed with proteins due to the unique protein environment, which modulates the pKa and reactivity of amino acid side chains. We tested the reaction of all HAT probes (1a-1o) with myoglobin (Mb) (w/o Cys) in order to assess the selectivity and reactivity of HAT probes towards Cys with proteins. Since Mb lacks the Cys residue, the modification of Mb was seen with all of the reactive probes (1a, 1b, Ox, 1j, 1n, and IAA) except for 1i, 1m, and 1o, clearly demonstrating their strong selectivity for Cys and correlating with the results of peptide screening (Figure 20). The majority of the additional HAT probes responded with additional Mb nucleophilic amino acids, like Lys (Figure 20). Due to their low reactivity and reduced water solubility, which is similar to peptide screening results, the probes 1d-1h and 1k-1l do not react with Mb (Figure 20, Supplementary Figure. 9).



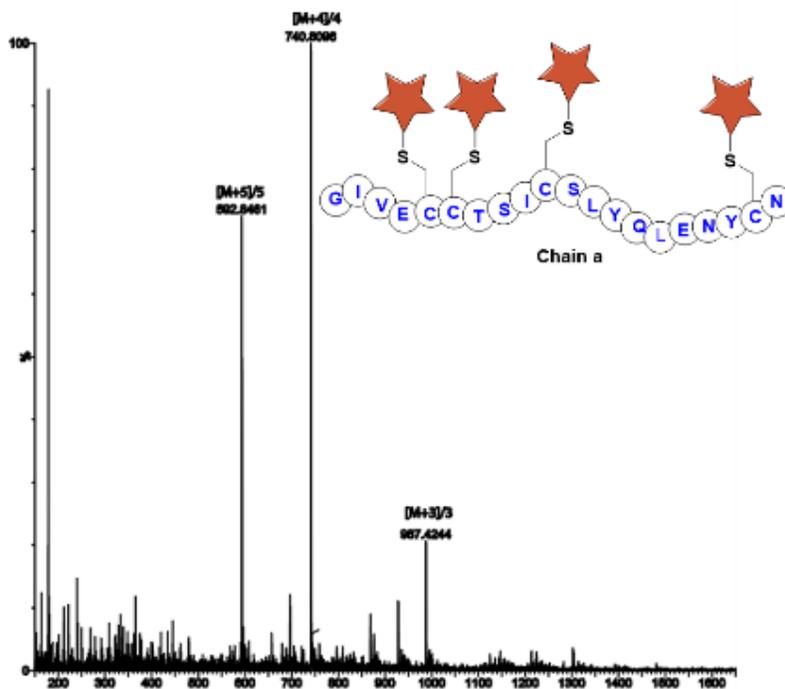
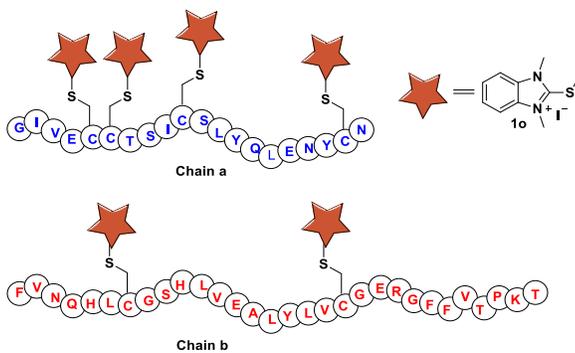
Probe	Conversion (%)						
1a	N-terminal (33)	1f	No modification	1j	Lysine (56)	1n	Lysine (99)
1b	Lysine (88)	1g	No modification	1k	No modification	1o	No modification
Ox	N-terminal (70)	1h	No modification	1l	No modification	1p	No modification
1e	No modification	1i	No modification	1m	No modification	IAA	Lysine (88)

Figure 20. Selectivity studies of probes with myoglobin

We performed the reaction with both natural insulin (w/o free Cys) and reduced insulin containing six free cysteine residues in order to determine the reactivity and selectivity of **1o** and **1m** for proteins (two in chain A and four in chain B, Figure 21, Supplementary Figure. 10). Under the same conditions (50 equivalents of **1m**, 12 hours, pH 7.5), we observed full modification of all cysteine residues in both chains A and B of reduced insulin by **1o** (>99 percent conversion) as compared to **1m** (chain A - 55 percent and chain B - 40 percent) (Figure 22, Supplementary Figure. 11). Both **1o** and **1m** did not modify the native insulin, demonstrating the remarkable selectivity of **1o** and **1m** for Cys. Next, we studied and compared the selectivity of IAA for Cys using myoglobin (Mb). Under physiological conditions (NaP, pH 7.5, 10 mM), the reaction of Mb with IAA (100 equiv.) generated the Mb-conjugate with multiple modifications (Figure 20). This demonstrated that IAA is not primarily selective for Cys and also results in the modification of other reactive nucleophiles on Mb, such as Lys and the N-terminus, correlating with published reports. These experiments demonstrated that in comparison to the widely used IAA and all other HAT probes, HAT probe **1o** is more reactive and more selective for cysteine conjugation. Following a thorough investigation of the effectiveness, chemoselectivity, range, and stability of the cysteine modification caused by the HAT probe **1o**, we further investigated its suitability for additional protein bioconjugation.



10.50 eq
 Nap 10 mM, pH 7.5 / 0.05 M TCEP pH 7.4
 rt, 8 h



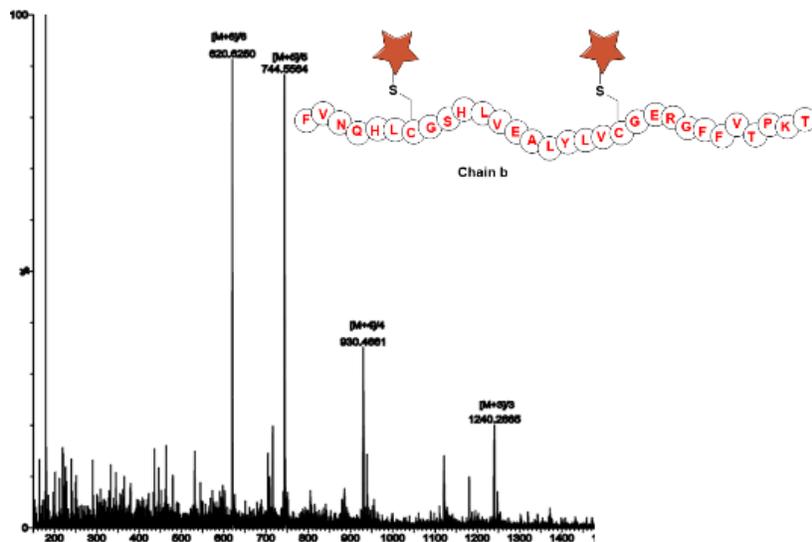
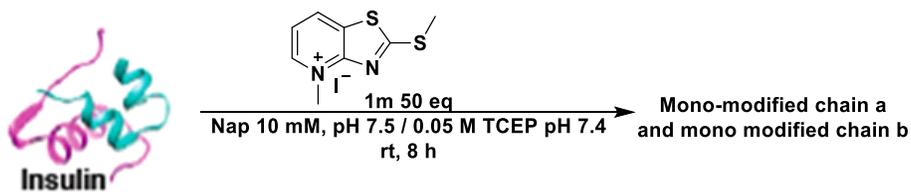


Figure 21. Labeling of free cysteine in reduced insulin by probe1o. Selective modification of cysteine in reduced insulin with six free cysteines (4 cysteine modifications observed in chain a, 2 cysteine modifications observed in chain b).



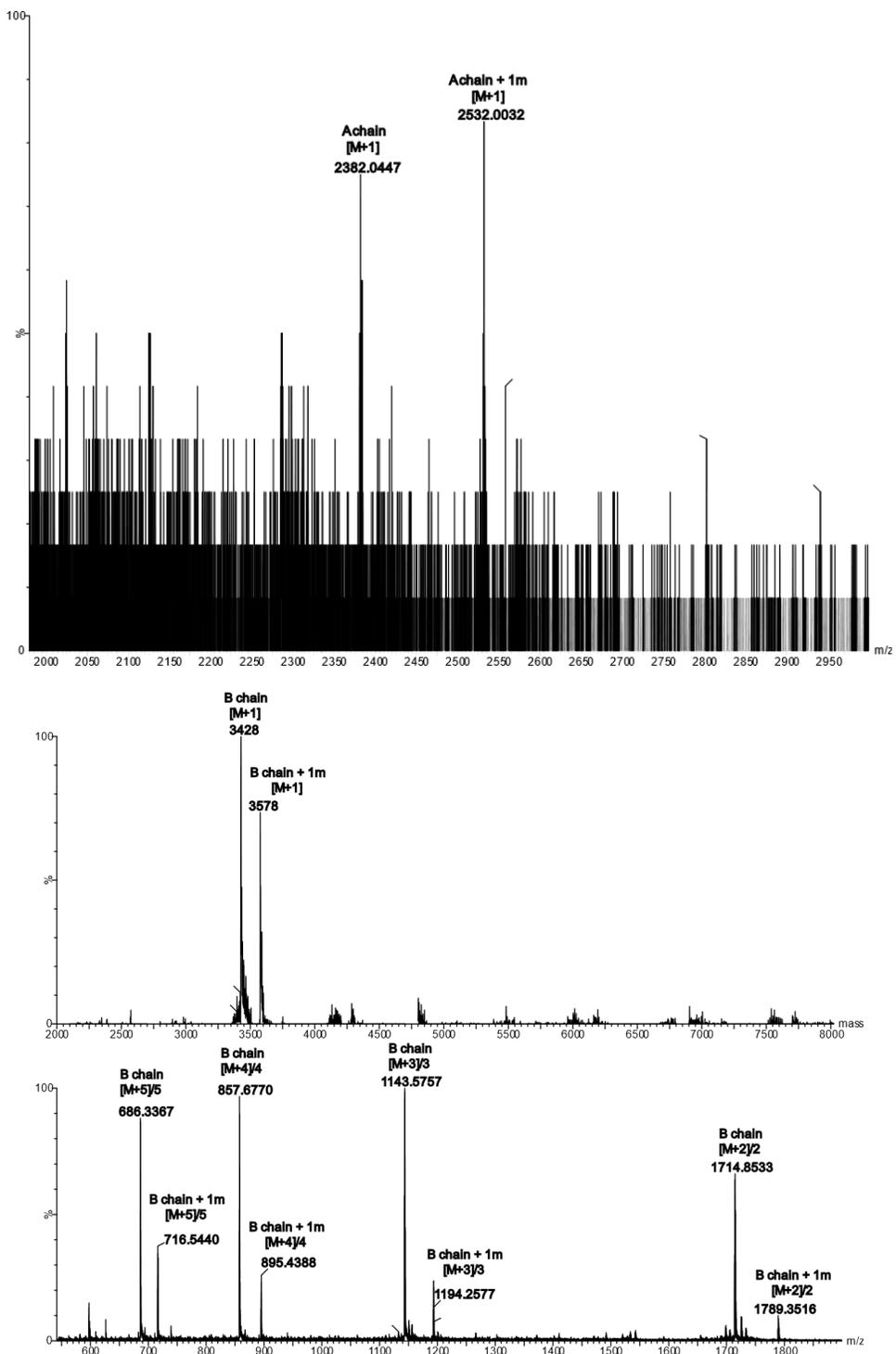


Figure 22. Modification of free cysteine in reduced insulin by probe 1m. Selective modification of cysteine in reduced insulin with six free cysteines (1 cysteine modifications observed in chain a, 1 cysteine modifications observed in chain b).

For bioconjugation, bovine serum albumin (BSA) with a single free cysteine residue was used. By LC-MS analysis, the modified protein **BSA-1o** was obtained after treatment of BSA (0.15 mM) with HAT probe 1o (300 equiv.) in Nap 7.5 buffer at 25 °C for 8 h had a conversion rate of >99 percent (Fig 23). Under the same reaction conditions, no modification was found in the lysozyme protein, which contains oxidized cysteine. Reaction with 1o modified 1-3 cysteine residues in reduced lysozyme produced 8 free cysteines as a result of lysozyme reduction (Nap 7.5 buffer at 25 °C for 8 hours, >99 percent conversion) (Figure 23, Supplementary Figure. 12). The simple surface accessibility of the probe and the microenvironment influence on the pKa of specific cysteines in lysozyme contribute to the high reactivity of probe towards those cysteines. These findings suggested that proteins could be modified using the HAT probe in a very efficient and chemoselective manner.

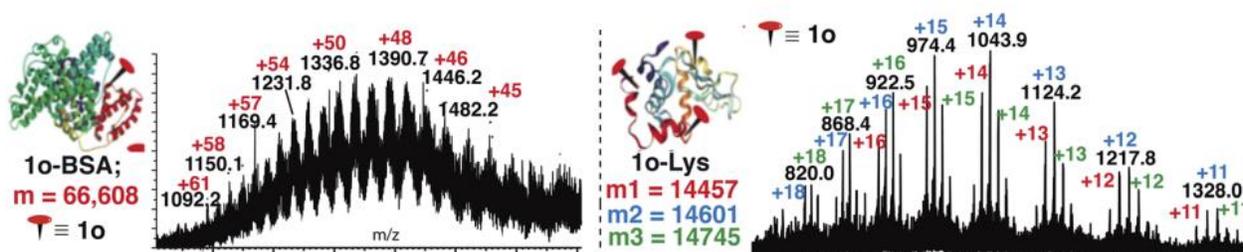
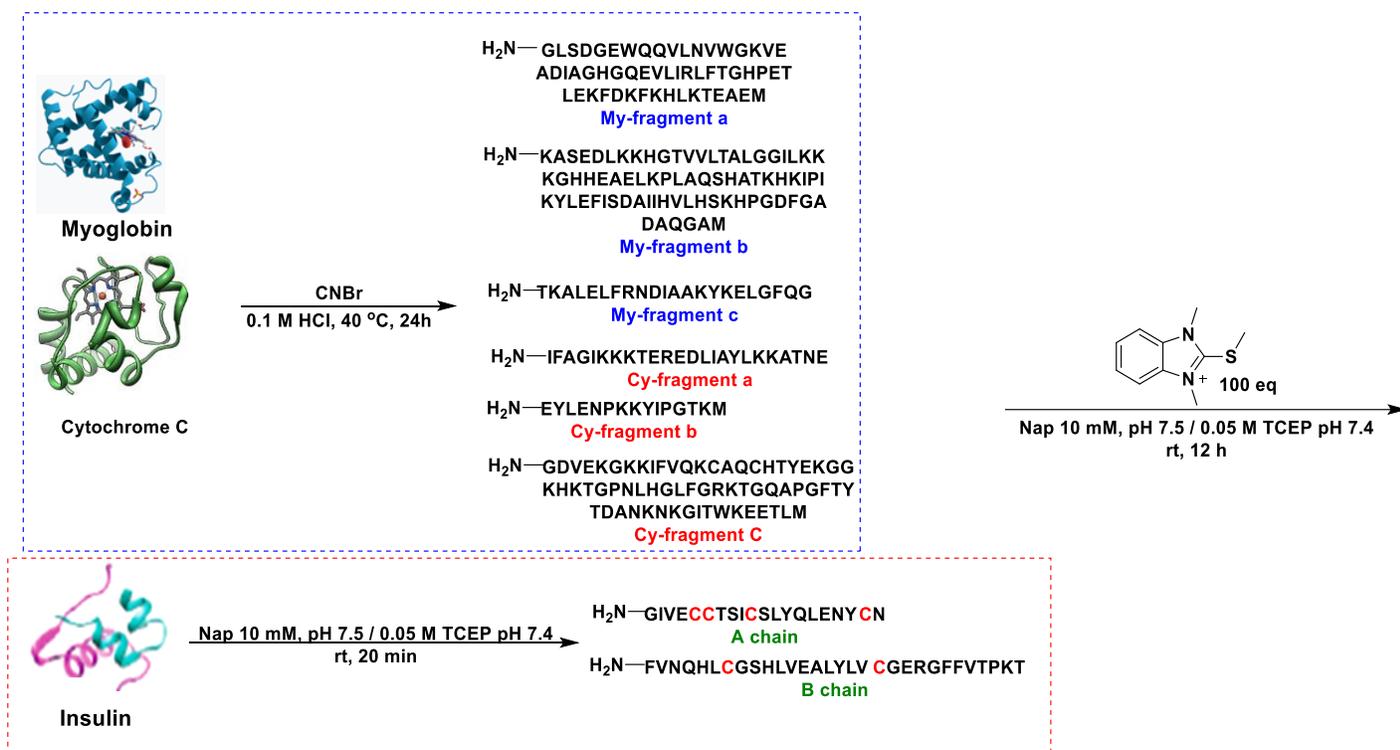


Figure 23. Selective modification of cysteine in BSA and reduced lysozyme by probe 1o. Selective modification of cysteine in a protein BSA with one free cysteine and reduced lysozyme (1–3 cysteine modifications observed).

1.2.3.3 Selective tagging of Cys in a complex mixture

We tried tagging numerous proteolytic fragments in the same solution to explore the possibility of our approach for enrichment in a complex mixture as a further illustration of the strong selectivity of HAT probe 1o for Cys. Under optimal reaction conditions, the mixture of

proteolytic fragments produced by CNBr cleavage of Mb, cytochrome C, and reduced insulin was incubated with HAT probe 1o for 12 hours. Data from the LCMS analysis of the reaction showed that only free Cys containing proteolytic fragments were marked with 1o. (Figure 24, Supplementary Figure. 13).



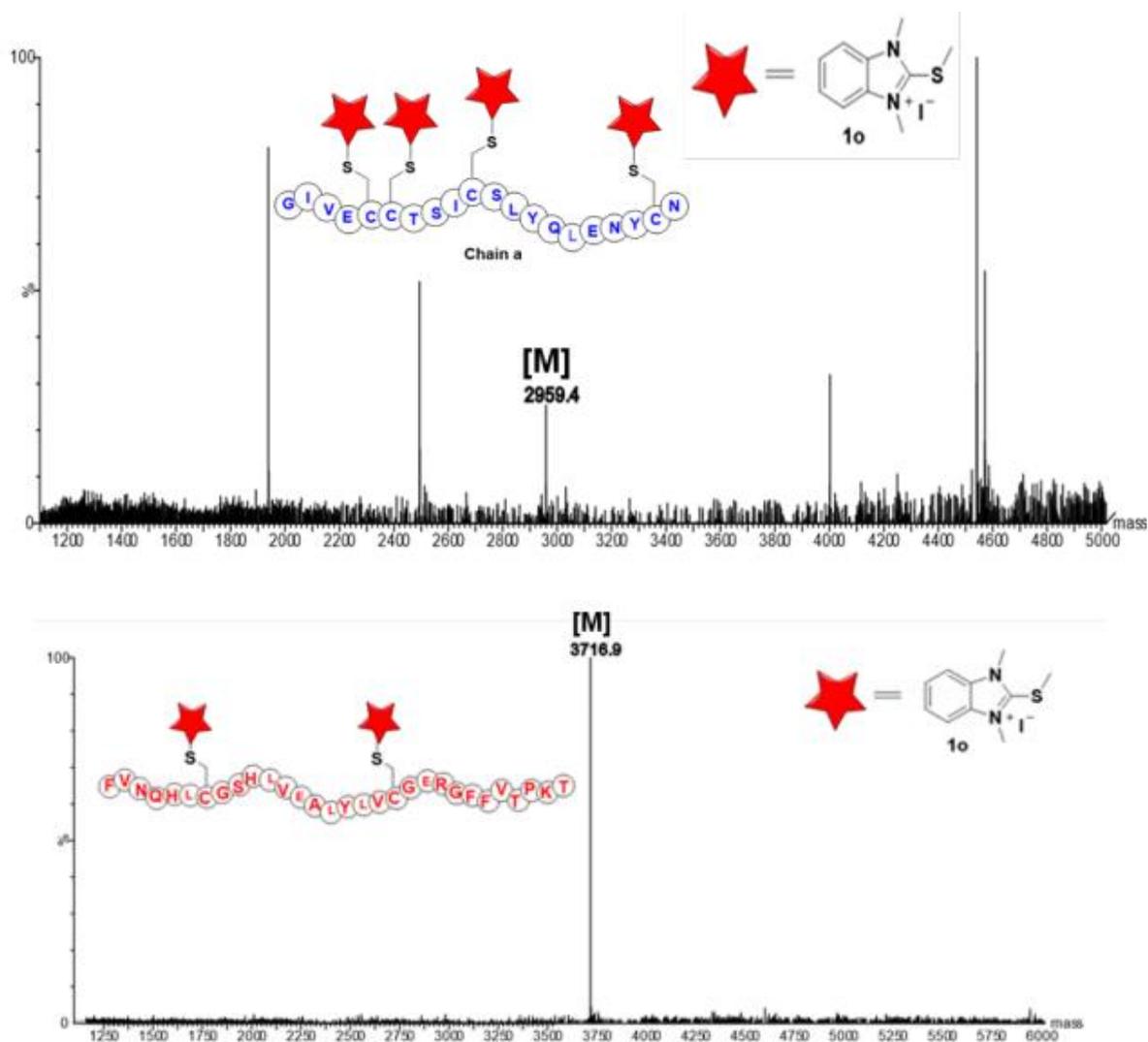


Figure 24. Enrichment of cysteine containing peptides with 1o in mixture of proteolytic fragments. Probe 1o is able to selective tag the cysteine containing peptides from all the proteolytic fragments of myoglobin, Cytochrome C and reduced insulin in complex mixture.

1.2.3.4 Reversibility and stability of the HAT–cysteine conjugation

The stability of the 1o-conjugated peptide 1o-Ac-GCF **3b** was then evaluated by HPLC monitoring under various reaction conditions. We showed that the peptide conjugate 1o-Ac-GCF

3b is stable for 48 hours at low pH 3.5 at both room temperature and 40 °C and that just 10% degradation was seen in 24 hours at pH 7.5. (Figure 9). Tris(2-carboxyethyl)phosphine (TCEP), a potent protein disulfide reducer, does not degrade peptide conjugate 1o-Ac-GCF **3b** when exposed to it for 48 hours (Figure 25, Supplementary Figure. 14).

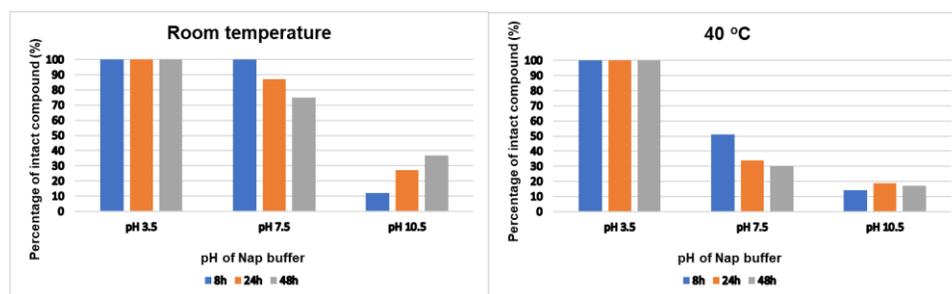


Figure 25. Stability of Ac-GCF-1o at 25 °C and 40 °C under different pH conditions. The bioconjugate product showed high stability under pH 3.5 at both room temperature and 40 °C. We observed high stability of the conjugate for 24h under physiological conditions (pH 7.5) at room temperature.

The HAT-peptide combination is stable under physiologically relevant conditions, but we wanted to see whether there was a way to decouple it without leaving any traces in order to get around any restrictions on irreversible protein inhibition. To do this, it is possible to transform the nucleophilic Cys residue into electrophilic DHA, enabling reversibility through the assault of a nucleophile on the unlabeled starting material. In fact, HPLC and MS analysis showed that the peptide conjugate 1o-Ac-GCF **3b** rapidly converted to the unmodified peptide Ac-GCF **2b** in 5 min when exposed to reduction conditions in the presence of sodium borohydride (10 equiv., 25 °C, in NaP pH 7.5). (Figure 26).

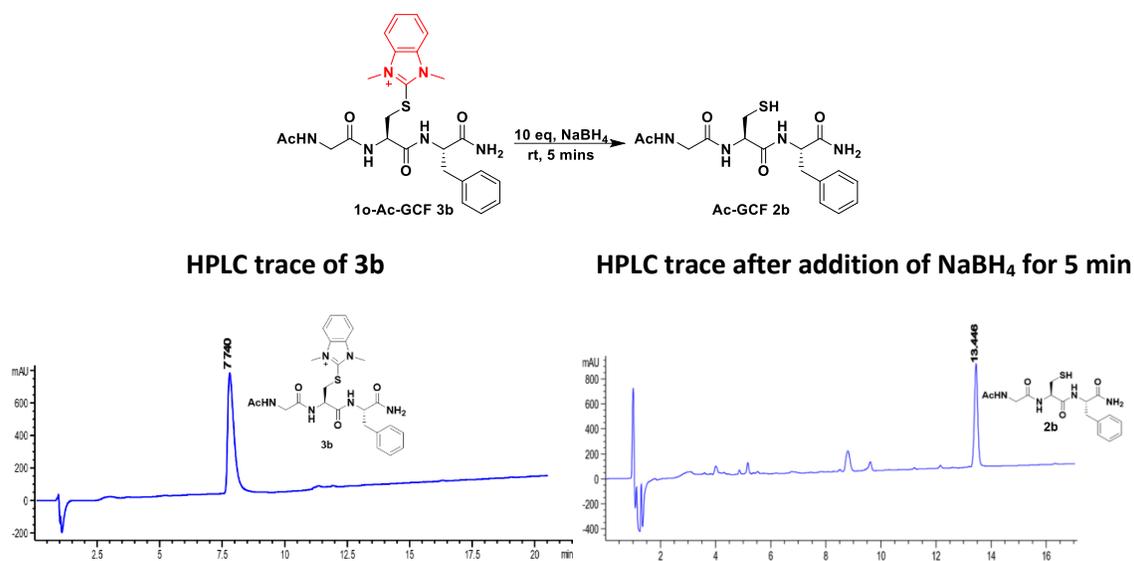


Figure 26. Reversible study of Cys-HAT biconjugate with NaBH₄. The resulting biconjugate is easy to decouple with NaBH₄ under mild and rapid conditions.

Next, we applied the reversibility approach for insulin modified with **1o**; we subjected 1o-modified insulin chains A and B to reducing conditions and within 5 min observed the complete reversibility to unmodified insulin chains A and B with full conversion (>99% Figure 27, Supplementary Figure. 15). It is notable that the reduction produced the original protein with no modifications. The HAT probes' exceptional benefits in bioconjugation and proteome profiling are highlighted by their high selectivity and reactivity for Cys under physiological conditions, as well as their capacity for rapid reversal in a traceless manner to regenerate unmodified protein. This reduces the drawbacks of producing irreversibly modified proteins.

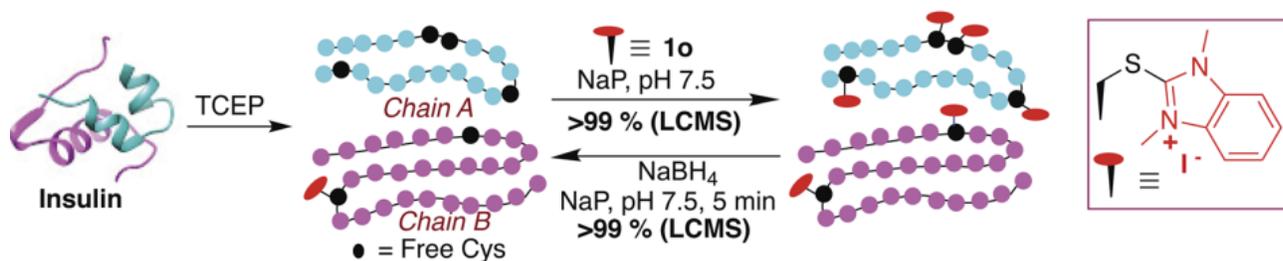


Figure 27. Reversibility studies of 1o-insulin conjugate.) Reversibility of the 1o-protein conjugate (0.15 mM) in the presence of sodium borohydride (1.5 mM) in NaP (pH 7.5) in 5 min.

1.2.3.5 Reactivity inversion by HAT probes

Many protein conjugation methods utilize the inherent nucleophilicity of Cys and carry out reactions with electrophiles. By reversing the reactivity of nucleophilic cysteine residue into electrophilic DHA, one can achieve a novel method for cysteine modification, allowing for modification by diverse nucleophiles. By creating a covalent link between the lysine of the interacting protein partner and the DHA on a peptide or protein, this method may also be used to capture the proteins that interact with other proteins. By subjecting Boc-Cys-OMe **2c** to basic conditions (NaP, pH 10.5, 37 C), we hoped to accomplish this. We then observed dehydroalanine Dha **3c** was formed directly via the 1o-Cys modified intermediate as shown by NMR and LCMS (86 percent, 8 hours, Figure 28, Supplementary Figure. 16).

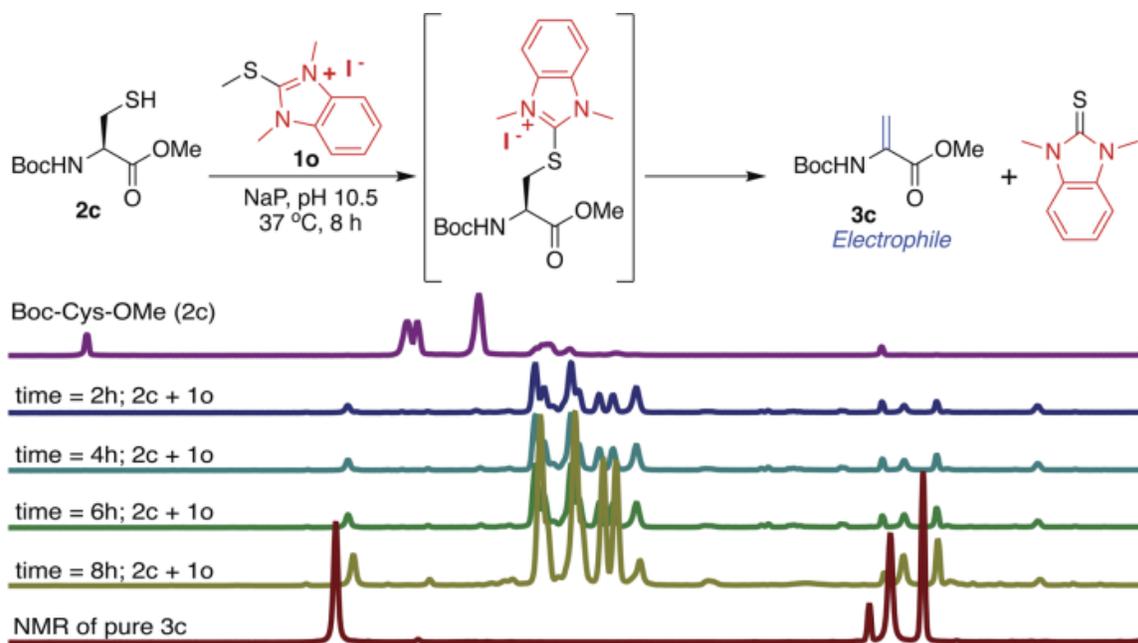


Figure 28. Conversion of peptide Boc-Cys-OMe 2c directly to dehydroalanine DHA 3c. Conversion of peptide Boc-Cys-OMe 2c directly to dehydroalanine DHA 3c under basic conditions (NaP, pH 10.5, 8 h) with >99% conversion. ¹H NMR shows the formation of dehydroalanine (DHA) 3c from the HAT-Cys conjugate in a one step process.

By reacting with nucleophilic amines and thiols, the spontaneous elimination of the 1o-Cys conjugate under basic conditions yields a type 2 alkene, dehydroalanine Dha, which in theory serves as a handle for further conjugation with different cargoes like polyethylene glycol (PEG) polymers, fluorophores, or affinity reagents. Using HAT probes, we changed the nucleophilic Cys from Ac-GCF **2b** to Ac-G(Dha)F **3d** (88 percent conversion), which was then subjected to reactions with various nucleophiles, including mercaptoethanol and benzylamine, to create the thiol addition product **3e** and conjugated amine **3f**, both of which had a conversion rate of more than 99 percent (Figure 29, Supplementary Figure. 17).

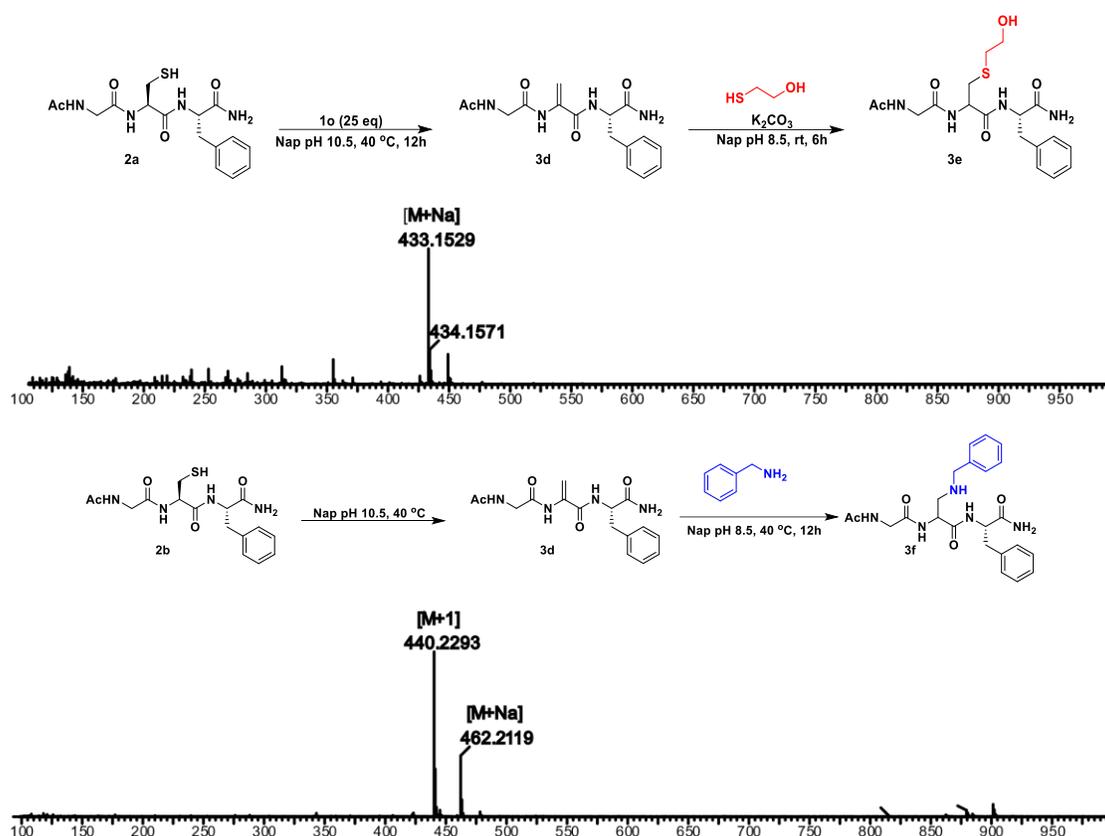


Figure 29. Aza-Michael addition and thiol-ene reaction of dehydroalanine. The dehydroalanine 3d can be further modified by Aza-Michael addition and thiol-ene reaction to achieve the site-specific modification.

Next, we converted the nucleophilic Cys on reduced protein lysozyme to DHA by treatment with probe 1o at pH 10.5. Three cysteines from the reduced lysozyme were converted to DHA with a conversion rate of more than 95%, as we saw. Next, we used the aza-Michael reaction to label DHA-modified lysozyme with benzylamine to create amine-modified lysozyme with a 78 percent conversion as measured by LCMS (Figure 30, Supplementary Figure. 18).

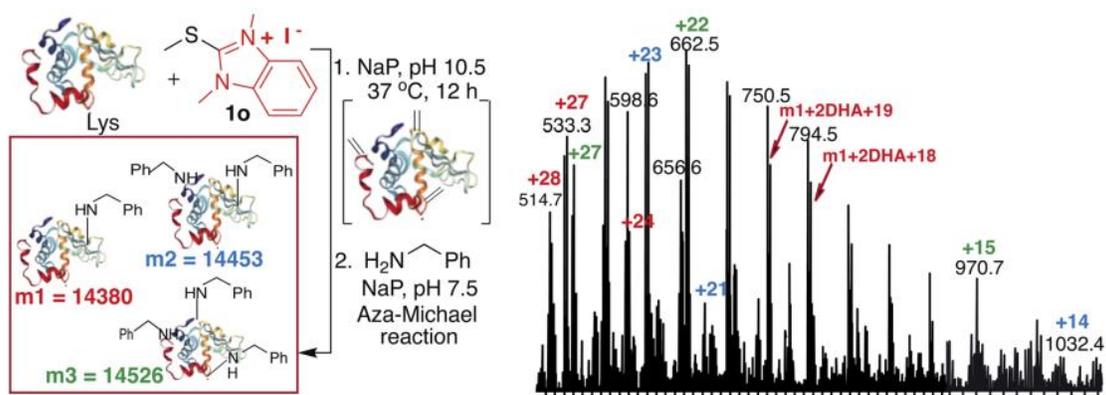
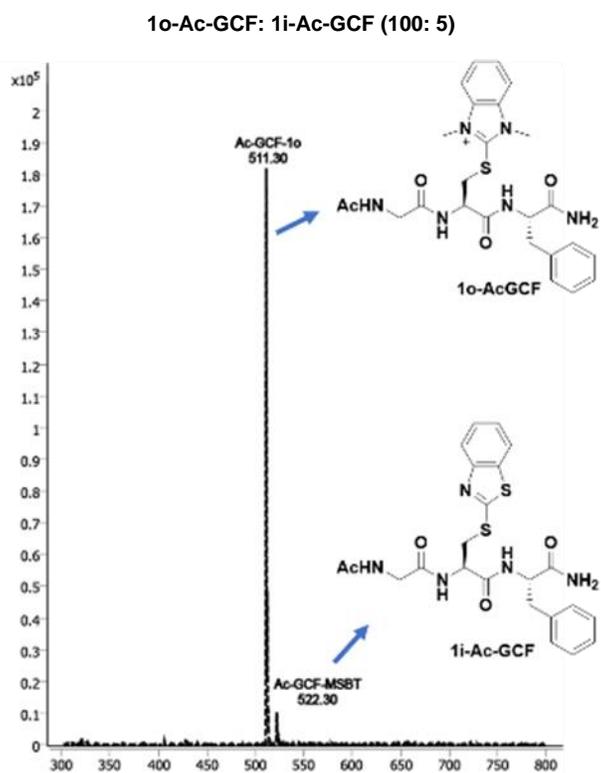


Figure 30. Aza-Michael reaction on DHA to generate amine labeled protein. Conversion of free cysteine of protein lysozyme (0.15 mM) to DHA (conversion >99%) by incubating with probe 1o (45 mM) at pH 10.5 for 12 h followed by the addition of amine (7.5 mM) by the aza-Michael reaction on DHA at pH 8.5 for 12 h to generate amine labeled protein (conversion 78%).

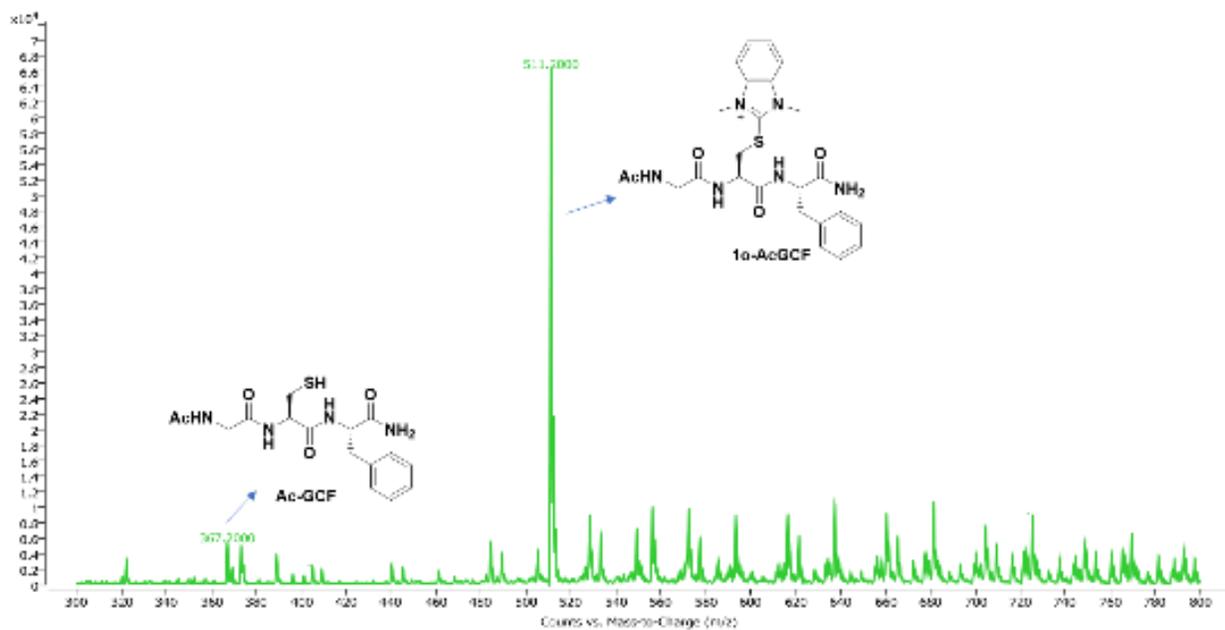
1.2.3.6 HAT as mass sensitivity booster

One of the limitations with current methods of selective labeling of Cys is the difficulty in the characterization of resulting bioconjugates by MS due to the poor ionization of the labeled fragments. Due to their limited mass sensitivity, low abundant protein biomarkers for the detection of early-stage infections are hampered by this constraint, which is especially pronounced in complex mixtures. Therefore, techniques are urgently needed to increase the detection sensitivity of tagged fragments. We applied the MS analysis on a mixture of free peptide Ac-GCF (5 mM), sulfone probe Ac-GCF (5 mM), IAA-labeled peptide Ac-GCF (5 mM), and HAT probe 1o labeled peptide Ac-GCF (5 mM) and unlabeled peptide Ac-GCF (5 mM) in water at equal concentrations to compare the ionization efficiency and mass sensitivity enhancement capabilities of the greatest notable signal amplification was provided by the 1o tag

on the peptide Ac-GCF (1o-Ac-GCF: 1i-Ac-GCF, 100: 5, 1o-Ac-GCF: IAA-Ac-GCF, 100: 8, and 1o-Ac-GCF: Ac-GCF, 100: 1.8). (Figure 31, Supplementary Figure. 19).



1 α -Ac-GCF: IAA-Ac-GCF (100: 8)



1 α -Ac-GCF: Ac-GCF (100: 1.8)

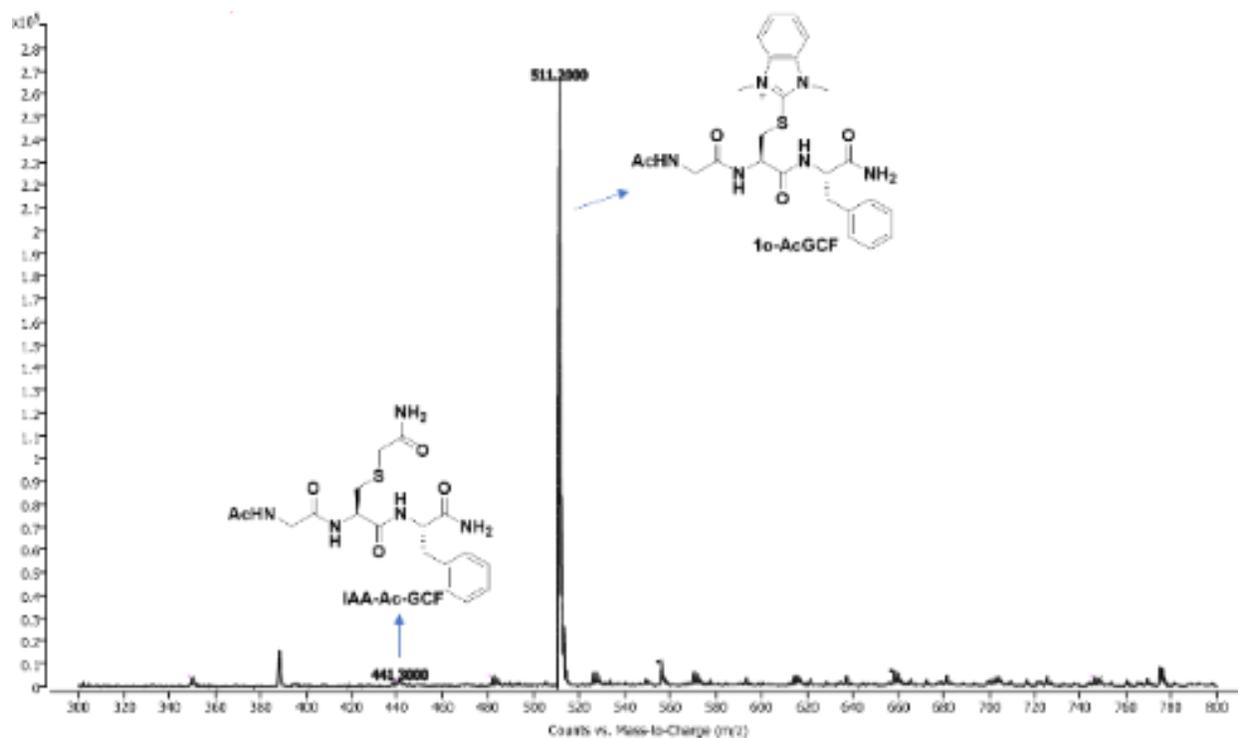


Figure 31. Mass sensitivity booster capability of 1o-modified peptides. To determine the ionization efficiency and mass sensitivity enhancement capability of HAT probe 1o in comparison with free peptide, sulfone probe 1i and the IAA probe, we carried out MS of the mixture of 1o-labeled peptide Ac-GCF (5 mM) and 1i-labeled peptide Ac-GCF (5 mM), 1o labeled peptide Ac-GCF (5 mM) and IAA-labeled peptide Ac-GCF (5 mM) and 1o labeled peptide Ac-GCF (5 mM) and unlabeled peptide Ac-GCF (5 mM) in water in equal concentrations.

The tagged peptide 1o-Ac-GCF was then detected up to 0.5 nanomolar concentration, according to a concentration assay (Figure 32, Supplementary Figure. 20). According to MS, without any purification, HAT probe N3-1o additionally significantly improved the detection sensitivity of the labeled protein fragments of the decreased insulin as compared to IAA (Figure 33, Supplementary Figure. 21). For unlabeled and IAA-labeled reduced insulin fragments, very poor sensitivity was observed for both chains A and B of insulin. In fact, chain A was undetectable by MS. Only one of chain B's free cysteines was labeled by IAA, whereas all four of chain A's free cysteines were labeled by **N3-1o** (Figure 33, Supplementary Figure. 21). Significantly high mass intensities of both chains A and B were present in the reduced insulin fragments that had been HAT **N3- 1o** tagged. Additionally, it makes chain A in MS, which is commonly undetected and may be highly significant in proteomics investigations, easier to identify.

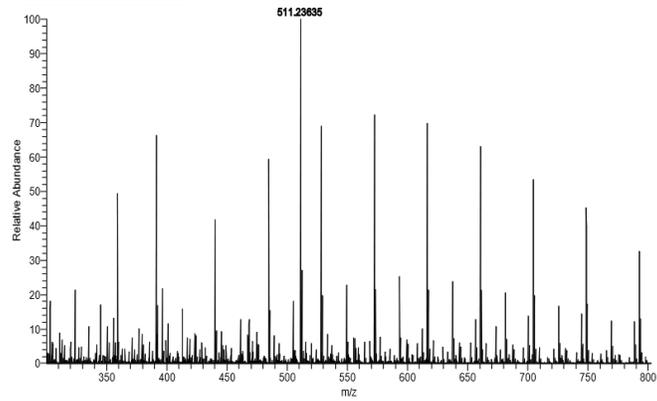


Figure 32. Mass sensitivity of 10-Ac-GCF at 0.5 nM. A concentration assay confirmed that the tagged peptide 10-Ac-GCF is detected up to 0.5 nanomolar concentration.

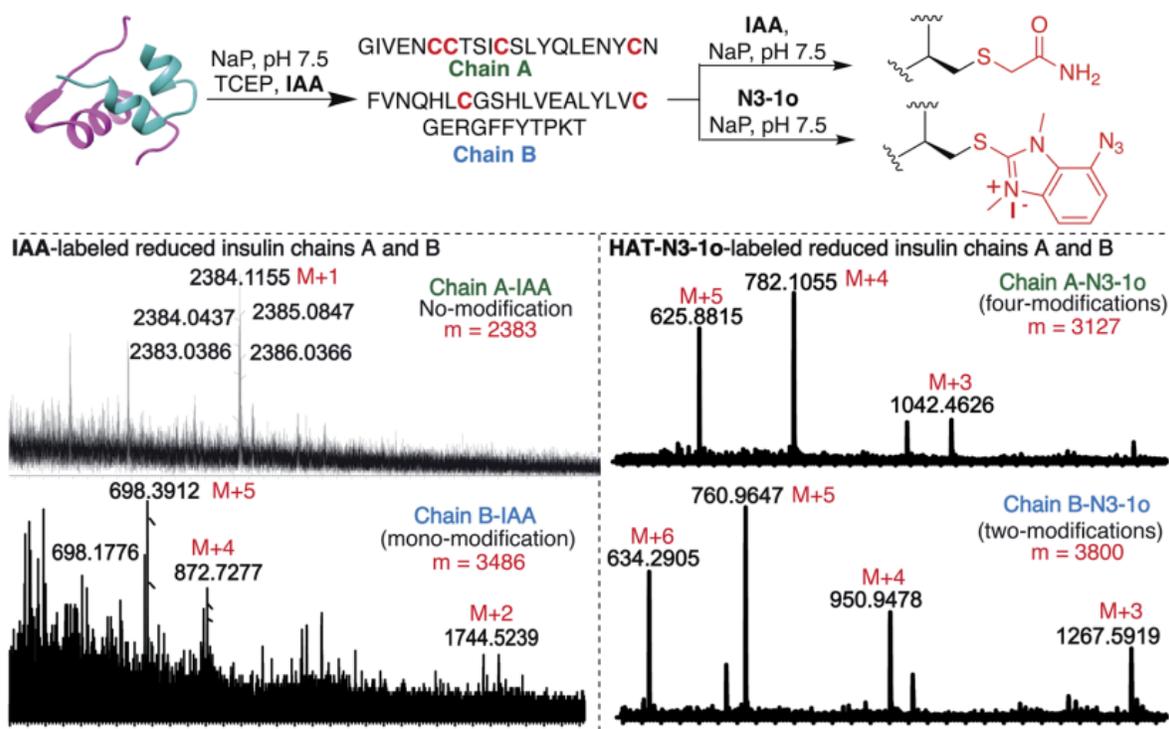


Figure 33. Mass intensity enhancement of N3-1o-reduced insulin bioconjugate products.)

HAT probe 1o significantly increases the mass detection sensitivity of a protein as compared to IAA tagged protein. Tagging with HAT probe N3-1o improves the detection of chain A of reduced insulin significantly (right MS trace). Chain A is not modified after TCEP reduction and labeling with IAA (left MS trace). Both chains A and B of reduced insulin are visible after tagging with the HAT probe, N3-1o (right MS trace). 1o modified all 4 cysteines and 2 cysteines of chains A and B respectively. IAA modified only one cysteine of chain B under identical conditions. Reaction conditions: reduced insulin (0.15 mM), probes 1o or IAA (50 equiv.) in NaP buffer at pH 7.5, room temperature for 8 h.

1.2.3.7 HAT probes for gel-based ABPP studies

Finally, because of its excellent selectivity and reactivity toward cysteine, as well as the modest size of the HAT group, which enables access to a wide range of proteins, we focused on the

utilization of the HAT **N3-1o** reagent for activity-based protein profiling (ABPP) applications. We tested the HAT probe **N3-1o** by gel-based competitive ABPP with NHS-tetramethylrhodamine (NHS-Rh) and iodoacetamide-tetramethylrhodamine (IA-Rh) using HEK293T cell lysate to determine the proteome reactivity of this family of electrophiles (Figure 34, Supplementary Figure. 22). The diminished in-gel fluorescence signal, which is a result of blocking IA-Rho labeling by pre-treating with high quantities of **N3-1o**, is consistent with cysteine-labeling by N3-1o (Figure 34, Supplementary Figure. 22). In the gel analysis experiment, large amounts of N3-1o are needed to produce a similar labeling to IAA, despite the fact that the kinetic investigation of probes revealed that **1o** and **N3- 1o** are substantially more reactive than IAA (Figure 18). This is due to the N3-1o cysteine conjugate products' poor stability in the gel-based competitive ABPP assays, which required heating at 37 °C and 95 °C. (SI). **N3-1o** provided no considerable blockage of proteome tagging by the lysine-reactive probe NHS-Rh, supporting its selectivity for cysteine labeling (Figure 34, Supplementary Figure. 22). These gel-based experiments prove that the HAT **N3-1o** probe's high specificity for identifying recombinant proteins extends to complicated cell lysates (Figure 34, Supplementary Figure. 22).

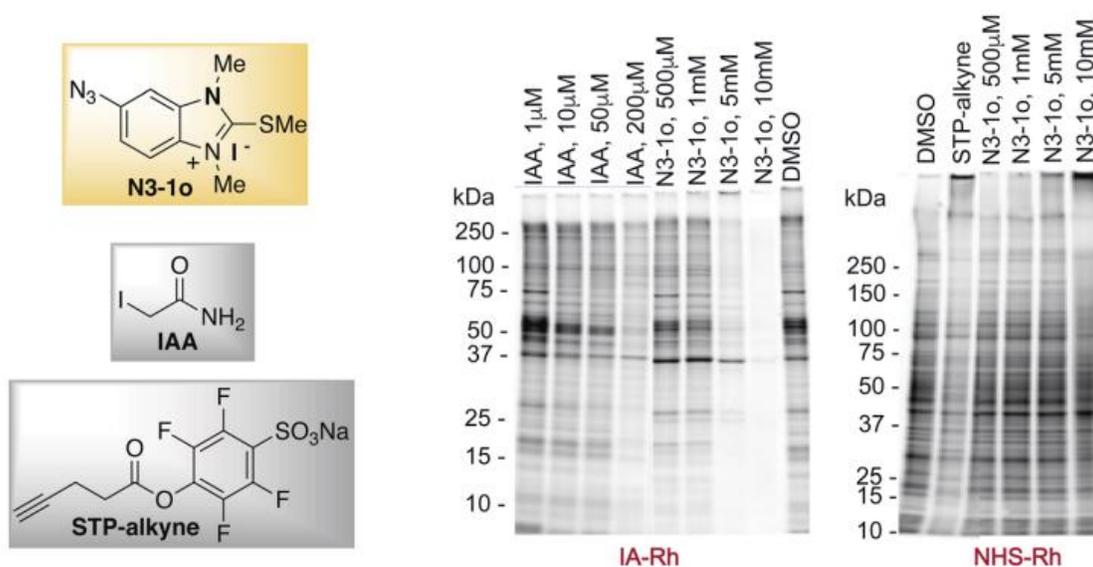


Figure 34. Chemoproteomic studies of the HAT N3-1o probe in HEK293T cell lysate. Chemoproteomic studies of the HAT N3-1o probe by gel-based competitive activity-based protein profiling (ABPP). In-gel fluorescence analysis of the HAT N3-1o probe at different concentrations (500 μM to 10 mM) and comparison with the cysteine reactive IAA probe (1–200 μM) followed by incubation and detection with IA-Rh in HEK293T cell lysate. Fading of bands with an increase in the N3-1o concentration showed labeling with cysteine (Left). In-gel fluorescence analysis of the HAT N3-1o probe at different concentrations (500 μM to 10 mM) and comparison with the lysine reactive STP-alkyne probe (1 mM) followed by incubation and detection with NHS-Rh. No fading of bands with increasing concentration of N3-1o indicates no reactivity with lysine (Right).

1.2.3.8 Conclusion

In conclusion, HAT offers a unique and stable chemotype for chemoselective cysteine modification free of cross-reactivity with other amino acids. HAT's reactivity is adjusted by varying the type of aromatic ring, the heteroatom attached to the ring, the oxidation state, and the methylation state of the heteroatom. The generated HAT probes are resistant to hydrolysis and

have a high degree of cysteine reactivity. One of the unique features of HAT probes is their ability to be quickly reversed by external stimuli to generate unmodified units in a traceless manner, indicating the potential utility of this approach in many research fields, such as protein immobilization, proteomics, and current drug discovery efforts avoiding the permanent modification of proteins. Unexpectedly, HAT probes allow for the reversal of the reactivity of nucleophilic cysteine to electrophilic dehydroalanine under mildly basic conditions (pH 10.5), allowing for the modification of proteins at the cysteine site by diverse nucleophiles such as thiols and amines. The ability of HAT probes to increase the mass sensitivity of the resulting bioconjugates by 100 fold allows for the simple detection of cysteine conjugates in a complex mixture, which is highly significant in proteomics studies for the identification of protein fragments with low protein abundance. Finally, we predict that the HAT's selectivity will enable future research aimed at identifying and pharmacologically modifying functional cysteines in entire proteomes, as well as a starting point for therapeutic treatments by reversible covalent inhibition of the reactive cysteines. The toolbox for bioconjugation, proteome probing, and pursuing otherwise inaccessible protein targets will be significantly widened by these cutting-edge HAT tools for probing cysteine, which would supplement current detection techniques.

1.2.4 Tunable amine-reactive electrophiles for selective profiling of lysine

1.2.4.1 Chemoselectivity Reaction of TARE Probes with Proteins

We performed the experiment with Myoglobin (Mb) in order to assess the selectivity and reactivity of probes **1 d** and **1 n** with proteins (without Cys Figure 35). According to an MS/MS analysis of Mb-1d and Mb-1n protein conjugates, both probes **1 d** (10 equiv, 1 h) and **1n** (100 equiv, 12 h) detected numerous modifications of lysines on Mb (Figure 35). Ser, Thr, Tyr, Trp, Glu, and His were not modified on Mb even after applying an excess of the probe **1n** (100 equiv)

and for a longer reaction period (12 h), further demonstrating the significant reactivity and selectivity of **1 d** and **1 n** for lysine. The modification of the lysine amino acid only occurred in the protein conjugates of lactalbumin-1d and cytochrome C-1d, according to MS/MS analysis of the probes **1 d** used to examine the modification of other proteins, including lb and CyC. (Figure 35, Supplementary Figure. 23). We synthesized the alkyne and azide functionalized TAREs **N3-1n**, **1 c-yne**, and **1 d-yne** in order to develop an efficient reaction for the enrichment of lysine fragments from a complex mixture. The reaction of the azide- and alkyne-functionalized TAREs with proteins like Mb, lb, and CyC was investigated. Similar to the model probe investigations with **1 c-1n**, azide- and alkyne-modified Lys of proteins Mb, lb, and CyC were generated under optimal conditions by TAREs **N3-1n** (25 equiv, 12 h), and **1 c-yne** (1 equiv, 1 h). This modification was verified by MS/MS. The labeling of significant amounts of lysine residues (6-7 lysines) on pure proteins such as myoglobin, cytochrome C, and lactalbumin without any cysteine using high equivalents of the probe **1 c-yne** demonstrated that the modification of lysine residue is independent of the presence of the cysteine residue (50 equiv).

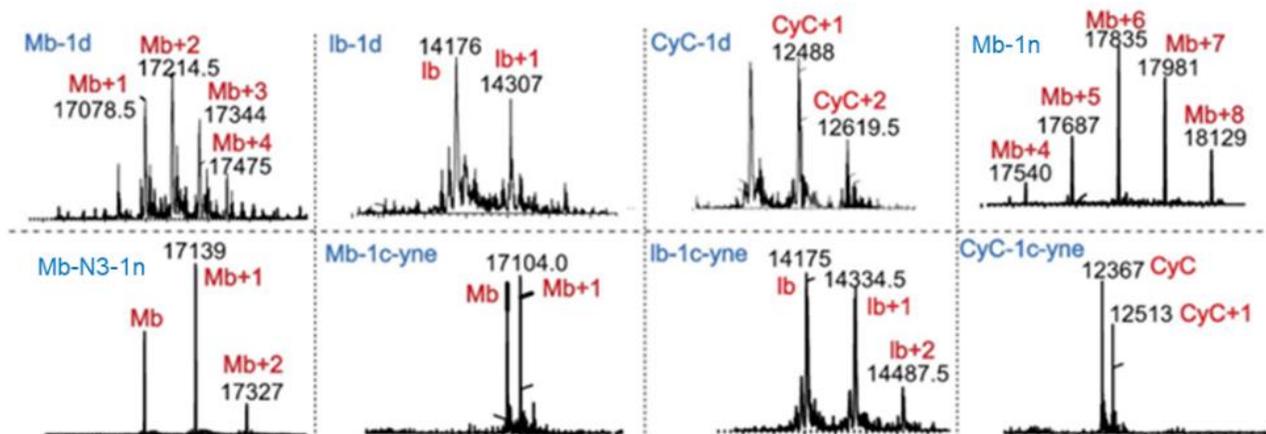
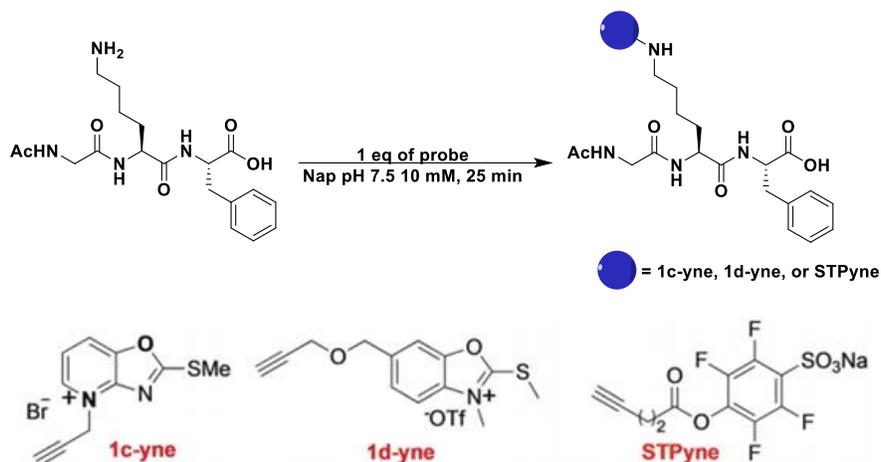


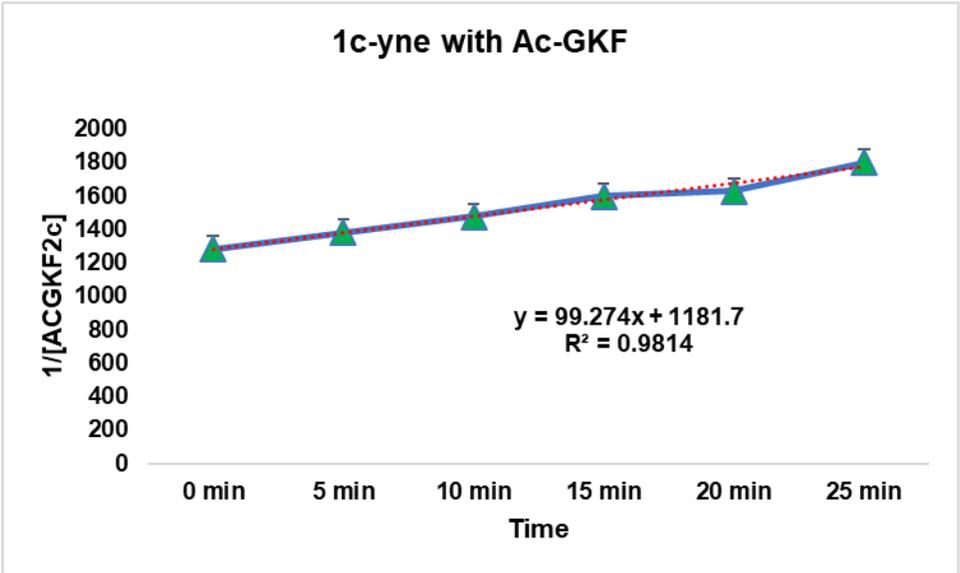
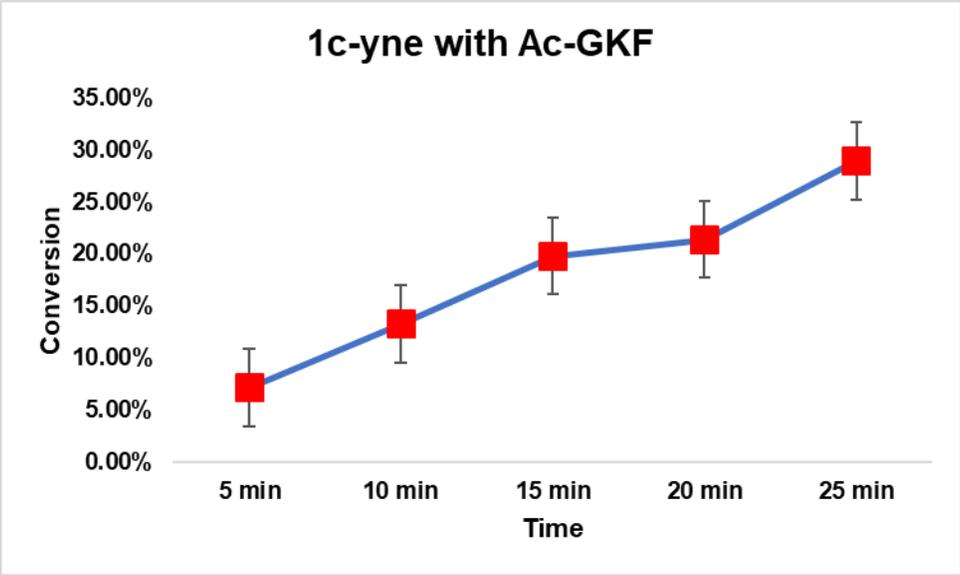
Figure 35. Chemselectivity reaction of probes with proteins. Modification of proteins with various TAREs. These proteins do not have any cysteine residue and TAREs showed high selectivity for lysine as analyzed by LC-MS/MS. Reaction Conditions: protein (3 mM in Nap (pH 7.5), probe 1 d (10 equiv, 30 mM), 1 e (100 equiv, 300 mM), room temperature for 1 h, detection wavelength 200 nm. For lb and CyC modification, probe 1 equiv of 1 d (3 mM) was used for 1 h.

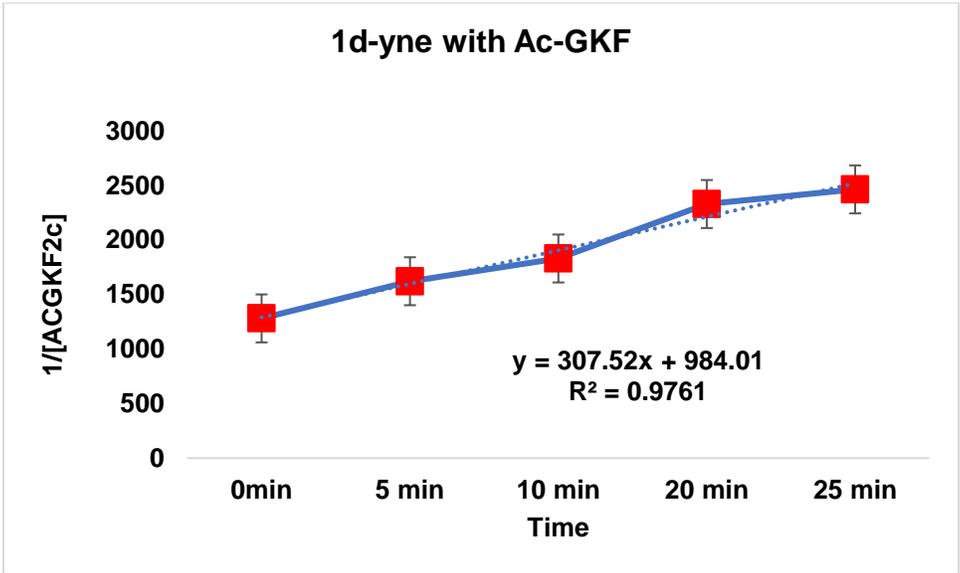
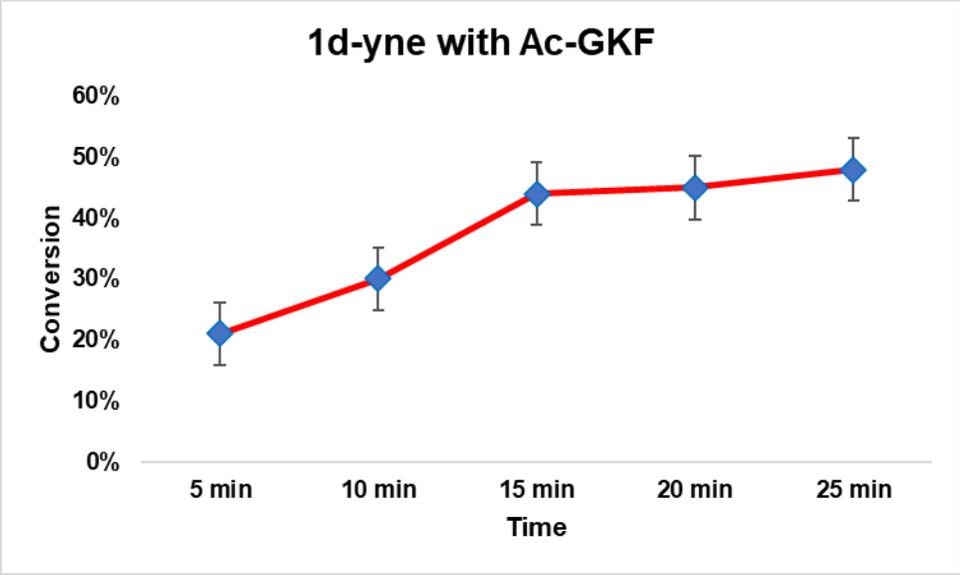
1.2.4.2 Rate, Stability and Reversibility of TARE Probes

Next, we compared the reactivity of the two most reactive TAREs **1 c-yne** and **1 d-yne** by carrying out the rate studies using a peptide Ac-GKF (GKF 4a). The reactions were monitored after regular intervals of time using HPLC and MS. The reaction with **1 d-yne** ($k = 307.52 \text{ M}^{-1} \text{ S}^{-1}$) showed 3-fold higher reactivity than **1 c-yne** ($k = 99.27 \text{ M}^{-1} \text{ S}^{-1}$) (Figure 36). The reaction rate of STPyne ($k = 190.92 \text{ M}^{-1} \text{ S}^{-1}$) with a peptide Ac-GKF **4a** showed lower reactivity compared to **1 d-yne** but high reactivity than **1 c-yne** (Figure 36, SI, Supplementary Figure. 24). Next, we sought to determine the stability of the reactive TAREs **1 c-yne** and **1 d-yne** towards hydrolysis and compare it with hydrolytic stability of NHS-ester and STPyne previously used for lysine profiling. We incubated probes in aqueous phosphate buffer (pH 7.5) under ambient conditions and monitored after regular intervals of time by HPLC and MS. Although of almost

similar reactivity, **1 c-yne** probe is more stable towards hydrolysis as compared to **1 d-yne** and only 30% degradation of **1 c-yne** probe was observed after 6 h and 50% of **1 c-yne** remained intact even after 12 h (Figure 37, SI, Supplementary Figure. 24). In contrast, NHS-ester showed 90% degradation in 2 h and completely degraded within 4 h (Figure 37). STPyne showed 90% degradation in 6 h. These studies showed the high stability of TARE **1 c-yne** as compared to NHS-ester and STPyne thus capable of acting as covalent inhibitors of lysine in cellular environment (Figure 37). Next, we conducted rate studies with the peptide Ac-GKF to evaluate the reactivity of the two most reactive TAREs, **1 c-yne** and **1 d-yne** (Ac-GKF 4a). By HPLC and MS, the reactions were observed at regular intervals (Figure 37). Compared to **1 c-yne** ($k = 99.27 \text{ M}^{-1} \text{ S}^{-1}$), the reaction with **1 d-yne** ($k = 307.52 \text{ M}^{-1} \text{ S}^{-1}$) shown a 3-fold higher degree of reactivity (Figure 36). STPyne's reaction rate with the peptide Ac-GKF **4a** ($k = 190.92 \text{ M}^{-1} \text{ S}^{-1}$) was lower than that of **1 d-yne** but higher than that of **1 c-yne** (Figure 36, Supplementary Figure. 24).







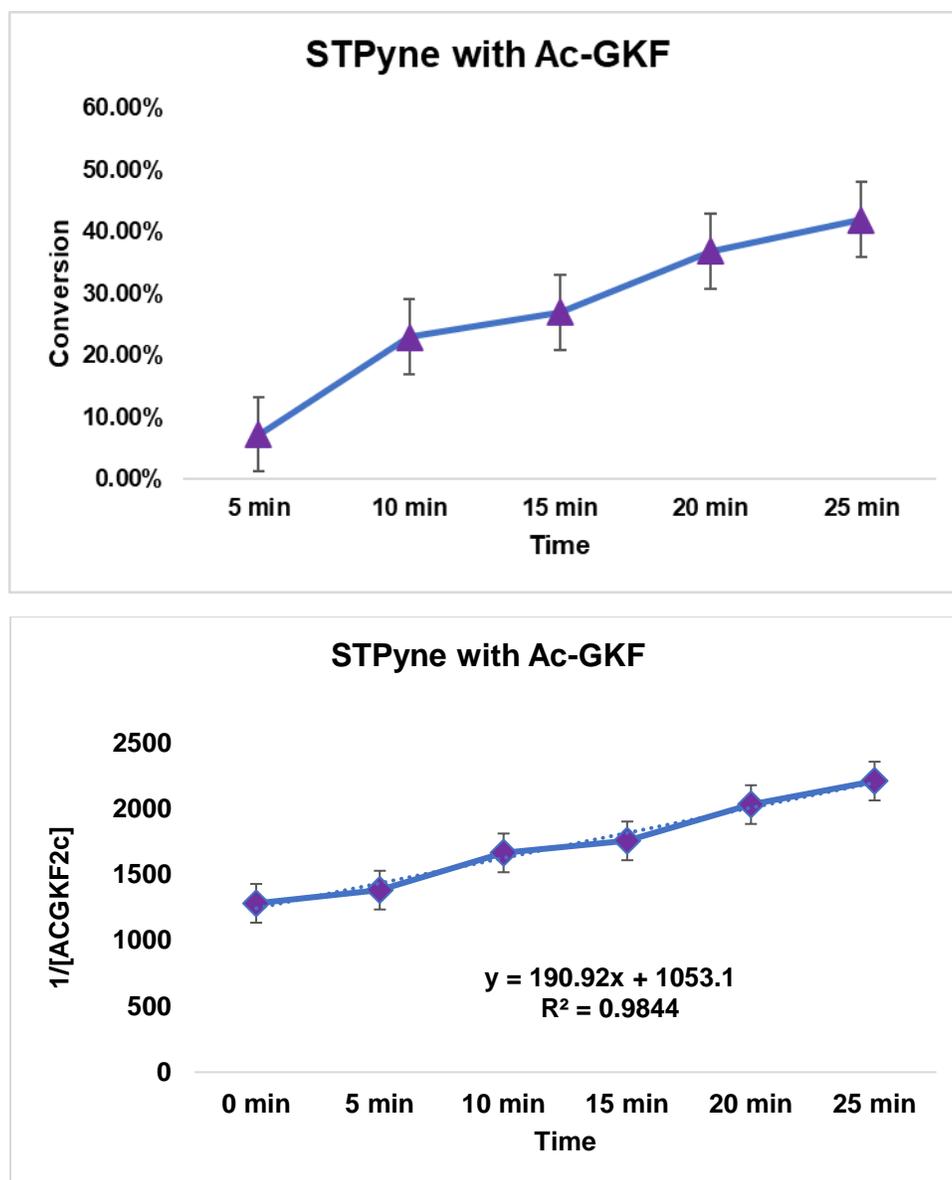


Figure 36. Rate study of 1c-yne, 1d-yne and STPyne with peptide OAc-GKF 2e. The result showed that 1c-yne, 1d-yne, STPyne bioconjugate reactions are second order reaction with $k = 99.27 \text{ M}^{-1} \text{ S}^{-1}$, $307.52 \text{ M}^{-1} \text{ S}^{-1}$, and $190.92 \text{ M}^{-1} \text{ S}^{-1}$ respectively.

The stability of the reactive TAREs **1 c-yne** and **1 d-yne** toward hydrolysis was then determined, and it was compared to the hydrolytic stability of NHS-ester and STPyne, which had previously been employed for lysine profiling. We incubated the probes in aqueous phosphate buffer (pH 7.5) at room temperature, and we used HPLC and MS to check the progress at fixed times.

Despite having practically comparable reactivity, the **1 c-yne** probe is more robust to hydrolysis than the **1 d-yne** probe; just 30% of the **1 c-yne** probe's degradation was seen after 6 hours, and 50% of the **1 c-yne** was still intact after 12 hours (Figure 37, Supplementary Figure. 25). In comparison, NHS-ester showed 90% degradation in 2 hours and 100% degradation in 4 hours (Figure 37, Supplementary Figure. 25). In 6 hours, STPyne showed a 90% degradation. These findings show that TARE **1 c-yne** is much more stable than NHS-ester and STPyne, and as a result, is able to function as covalent inhibitors of lysine in cellular environments (Figure 37, Supplementary Figure. 25).

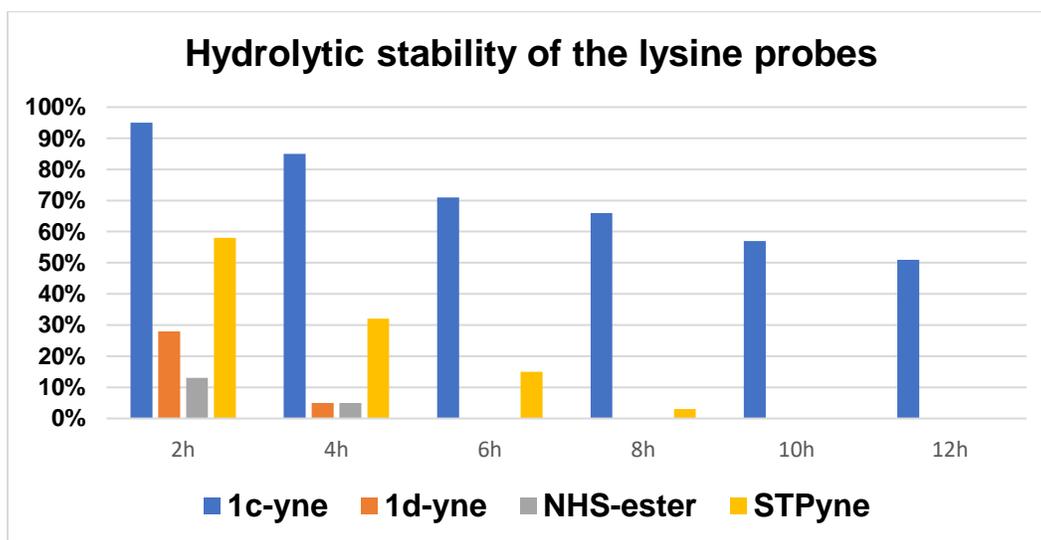
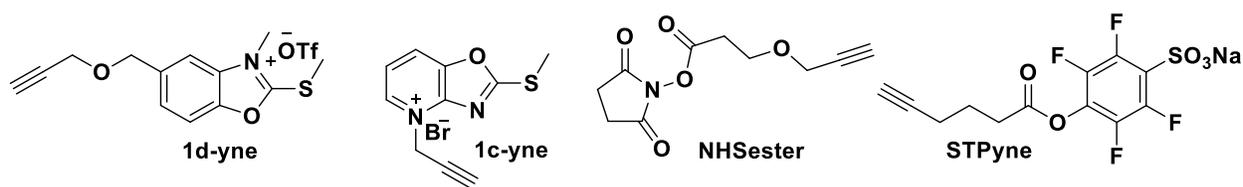
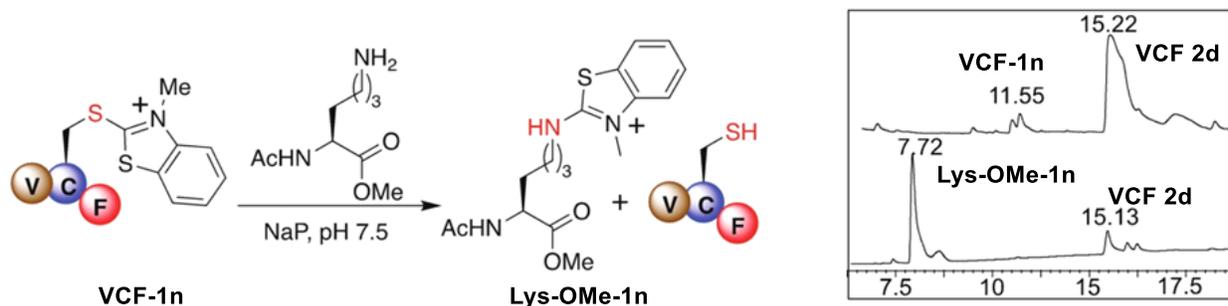


Figure 37. Stability study of 1c-yne, 1d-yne, NHS ester and STPyne. 1c-yne, 1d-yne, NHS ester, and STPyne (0.035 mmol) were incubated in 400 μ L of 10 mM Nap (pH 7.5) at room temperature. The result shows the probe 1c-yne is more hydrolytically stable than the widely used lysine labeling and profiling reagents NHS-ester and STP-yne.

We generated a probe 2-methylthio benzoN-methylthiozolinium ion **1n** with decreased reactivity by switching the O heteroatom for S in the 2-methylthio benzoN-methyloxazolinium ion **1d** in order to demonstrate the reversibility property of our probes. By reacting **1n** with the peptide AcVCF, we were able to isolate modest amounts of 1 n-Cysconjugate (VCF-1n), and we next treated the cysteine-adduct VCF-1 n with lysine methylester (Figure 38 a). According to the hypothesis, the reaction produced a stable adduct with lysine methylester (Lys-OMe1n), releasing the free peptide VCF **2d** for LCMS analysis (Figure 38 a, Supplementary Figure. 26). We created a 1d-thio-conjugate under non-aqueous conditions since 1d-Cys conjugate is very reactive and challenging to isolate under aqueous conditions. Complete modification with lysine methylester was shown in 1 hour under physiological conditions (NaP, pH 7.5, 10 mM) at room temperature (Figure 38 b). These experiments prove that the thioethers produced when Cys reacts with TAREs **1 c**, **1 n** go on to react with lysine and can be used to profile lysine specifically in the human proteome. We also observed full hydrolysis in 1 hour after incubating 1 d-thio-conjugate under buffer (NaP, pH 7.5, 10 mM) at room temperature (Figure 38 b, Supplementary Figure. 26), indicating the selectivity for lysine profile.

a



b.

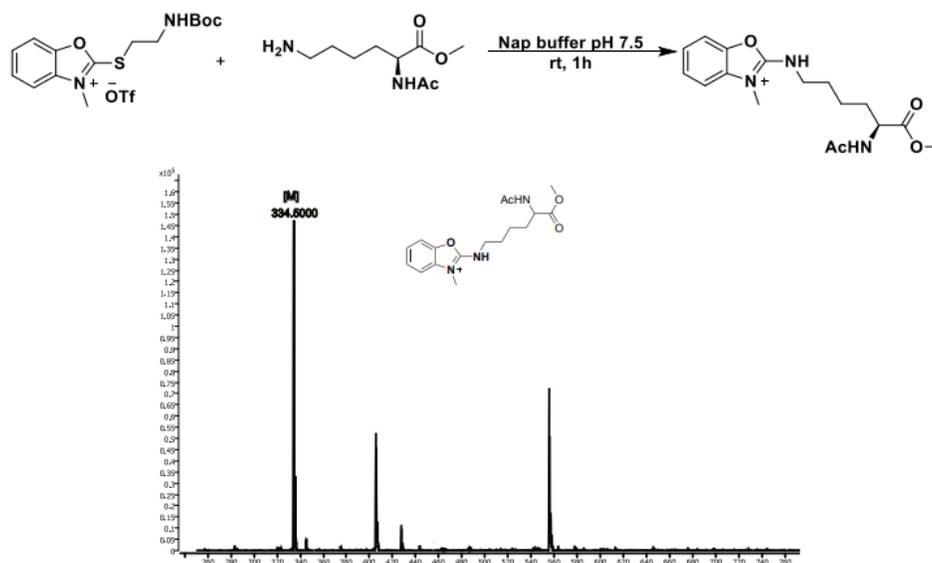
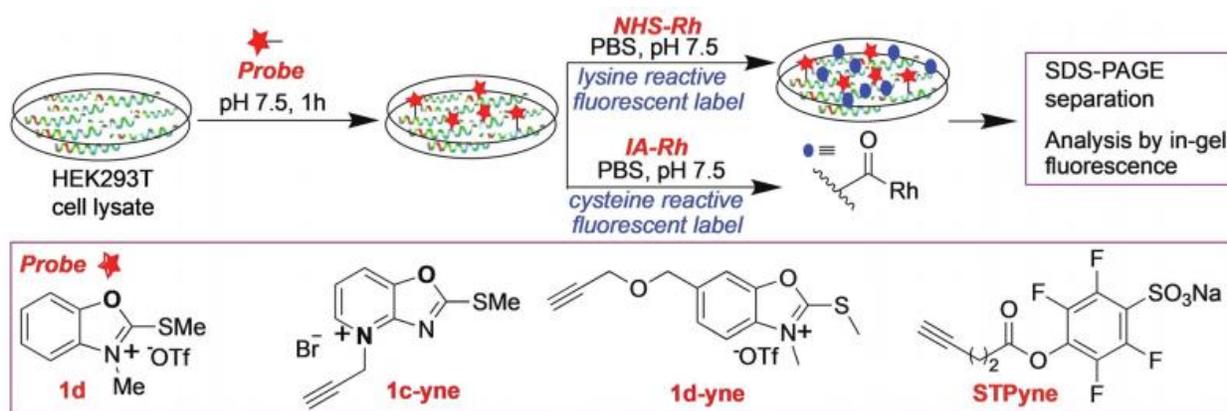


Figure 38. a). Reactivity of cysteine-TARE conjugate towards lysine to generate stable product. The cystein conjugated product is able to react with lysine residue to form a more stable lysine product. b). Reversibility study of 1d-thio-conjugate with lysine methylester. The cystein 1d-conjugated product is ready to react with lysine side chain to generate the hydrolytically stable lysine product.

1.2.4.3 Chemo-proteomic studies of TAREs

We then investigated whether this chemotype would perform well for activity-based chemoproteomic applications in complex cell lysates, encouraged by the lysine selectivity and enhanced reactivity observed in our TARE-protein and peptide-based labeling studies. We evaluated the relative cysteine- and lysine-reactivity of TAREs using a gel-based competitive activity-based protein profiling (ABPP) test. Iodoacetamide-tetramethylrhodamine (IA-Rh) or NHS-tetramethylrhodamine (NHS-Rh), respectively, or a pan-cysteine or pan-lysine reactive fluorescent probe were used to label HEK293T lysates after being first treated with probes **1 d**, **1 c-yne**, or **1 d-yne** (Figure 39). Gel-based ABPP analysis of compound **1 d** demonstrated enhanced selectivity for lysine (Figure 39, Supplementary Figure. 27), as indicated by

competition of labeling by NHS-Rh but not IA-Rh, whereas **1 c-yne** inhibited both cysteine and lysine labeling. On the other hand, the canonical lysine-reactive probe STPyne (1 mM) competed for labeling of both probes, consistent with off-target cysteine-reactivity, as reported before. Our next step was to assess the concentration range compatible with probe labeling because competitive inhibition of NHS-Rh labeling was only seen at high doses of **1 c** and **1 d** (e.g. 5 mM for **1 d**, Figure 39, Supplementary Figure. 27). We used an alkyne derivative of **1 d** termed **1 d-yne** for these investigations. Cell lysates were treated to **1 d-yne** labeling at the specified concentrations, CuAAC conjugation to biotin azide, and streptavidin blotting to detect the labeling (Figure 39, Supplementary Figure. 27). Fortunately, **1d-yne** demonstrated intense concentration-dependent labeling, and a banding pattern resembling that of STPyne was compatible with lysine-directed reactivity. The same labeling intensity was seen when comparing 1 mM of **1 d** to 100 mM of STPyne, indicating that the **1 d** TARE is a relatively attenuated warhead. We used mass spectrometry-based chemoproteomics to further evaluate the **1 c** and **1 d** TAREs' proteome-wide reactivity patterns. Using our SP3 chemoproteomic sample preparation method, cell lysates were exposed to either probe **1 c-yne** or **1 d-yne**, labeled proteins were attached to biotin azide using CuAAC, and the samples were prepared and analyzed (Figure 39, Supplementary Figure. 27).



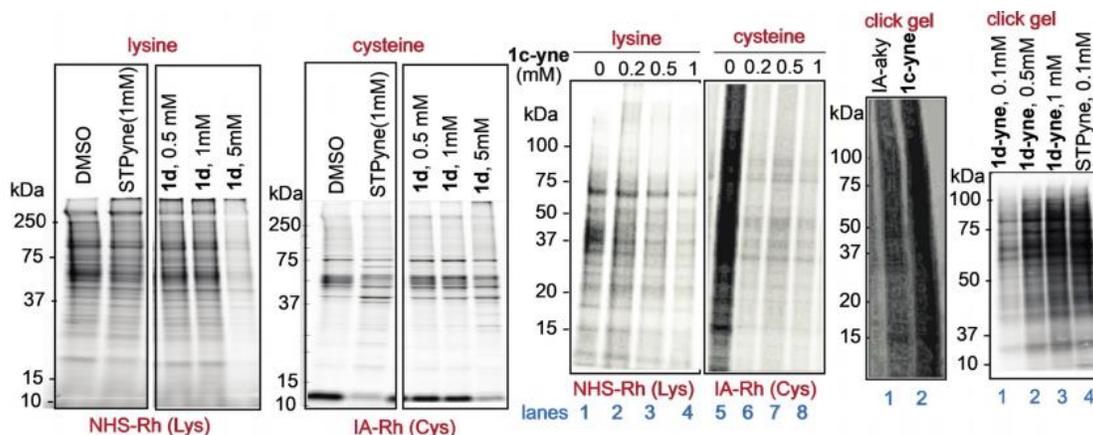


Figure 39. General protocol for proteome profiling by different probes and structure of probes. In-gel fluorescence analysis of STPyne (1 mM) and 1 d probes at different concentrations of 1 d (0.5 mM to 5 mM) followed by detection with NHS-Rh for lysine reactivity and IA-Rh for cysteine reactivity. 1 d is more selective for lysine as compared to STPyne. In-gel fluorescence analysis of 1 c-yne at different concentrations (0 to 1 mM) showed reactivity with both lysine and cysteine. Click gel assay of 1 c-yne and cysteine reactive IA-aky probes with cell lysate using fluorescent-biotin azide to determine the total labeled proteins. High labeling of cell lysate was observed with 1 c-yne. Click gel assay of 1 d-yne at different concentrations (0.1 mM to 1 mM) and STPyne (0.1 mM) probes with cell lysate using biotin-azide and streptavidin blot to determine the total labeled proteins. Dose-dependent labeling was observed for 1 d-yne with banding pattern similar to STPyne.

Briefly, magnetic beads with carboxyl coatings were used in single-pot solid-phase enhanced sample preparation (SP3) decontamination of TARE-labeled proteomes. Neutravidin was used to enrich biotinylated peptides from the SP3-resin tryptic digest and conduct LC-MS/MS analysis. Surprisingly, **1 c-yne** demonstrated nearly complete selectivity for lysine residues in contrast to its apparent cysteine-reactivity as reported by competitive gel-based tests (5124 total unique labeled lysine residues and 27 total unique labeled cysteine residues across two biological

replicate experiments (Figure 40, Supplementary Figure. 28). Similar to this, chemoproteomic analysis of **1 d-yne**-labeled lysates showed nearly complete lysine selectivity across three chemical doses examined (1595 total unique lysine residues and 54 total unique labeled cysteine residues). The **1 c-yne** datasets were re-searched using the program MSFragger, which has built-in PTMProphet for precise mass modification localization, and this discovered 1560 unique labeled lysines and 17 unique labeled cysteines, further confirming the specificity of the lysine residue (Figure 40, Supplementary Figure. 28). Consistent with our gel-based analysis, we observed a dose dependent increase in peptides identified as the concentration of **1 d-yne** was increased from 1 mM to 100 mM. **1 c-yne** labeled substantially more peptides (5151) than **1 d-yne** (1649), in contrast with the aforementioned kinetic analysis that revealed 3-fold higher reactivity for **1 d-yne** ($k = 307.52 \text{ M}^{-1} \text{ S}^{-1}$) as compared to **1 c-yne** ($k = 99.27 \text{ M}^{-1} \text{ S}^{-1}$) (Figure 36, Supplementary Figure. 28). Surprisingly high reactivity of **1 c-yne** in gel and proteomic studies as compared to **1 d-yne** might be due to the high hydrolytic stability of **1 c-yne** (95% intact in 2 h) as compared to **1 d-yne** (28% intact in 2 h) under the reaction conditions (Figure 37).

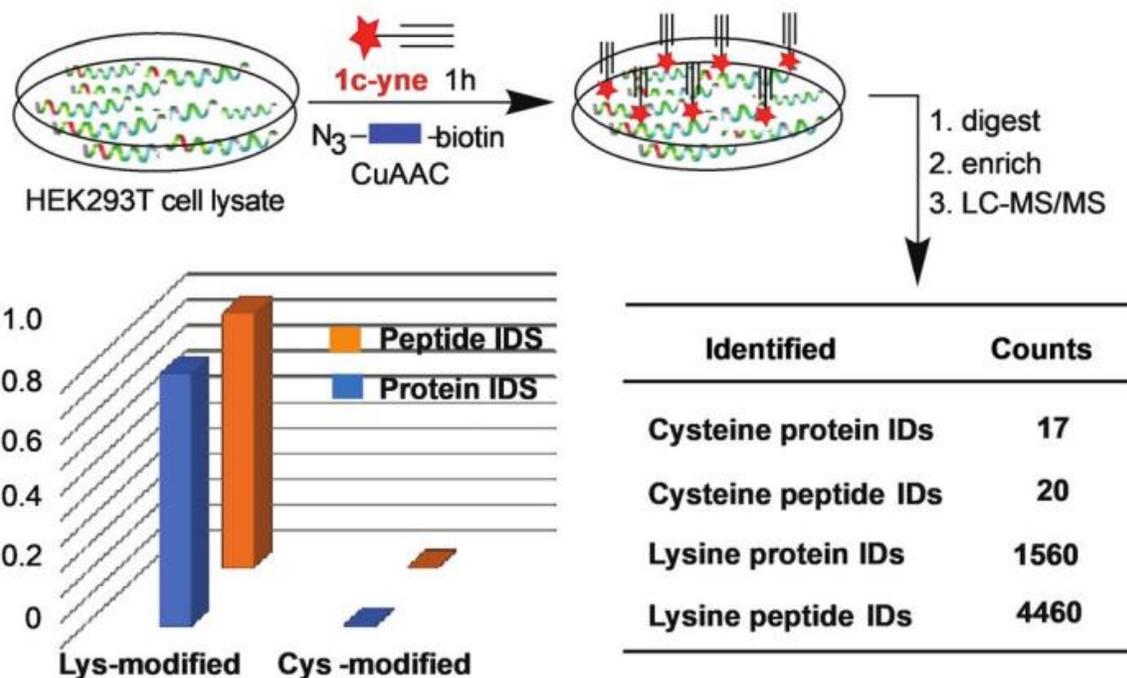


Figure 40. Percentage of unique peptides and proteins labeled on each nucleophilic amino acid by **1 c-yne** in HEK293T proteome. Cellular lysates are labeled with **1 c-yne** at different concentrations followed by conjugation with azide-biotin tags (blue) using CuAAC, enrichment of labeled proteins by neutravidin-conjugated beads and digested stepwise with trypsin to yield **1 c-yne**-labeled peptides for LC-MS analysis. Percentage of unique peptides and proteins labeled on each nucleophilic amino acid by **1 c-yne** in HEK293T proteome. Probe **1 c-yne** preferentially enrich lysine residues in human cell proteomes. Data represent means ± standard deviation for two experiments.

In contrast to our competing gel-based results, the absence of significant cysteine tagging observed with **1 c-yne** intrigued us. We hypothesized that this discrepancy might be caused by the lability of the cysteine adduct, which may result from the reducing CuAAC conditions. To verify this hypothesis, we used click chemistry to label a peptide GCF with **1 c-yne** both with

and without TCEP. Under click chemistry conditions, we only observed the modification of Cys in a peptide GCF when TCEP was absent; no modification of a peptide GCF was noticed when TCEP was present (Figure 41, Supplementary Figure. 29). Next, we isolated the Cys-modified GCF-1 c-yne adduct and incubated it in TCEP-containing buffer to further confirm. Within five min, we saw the GCF-1 c-yne adduct completely decompose into the unmodified peptide GCF (Figure 41, Supplementary Figure. 29).

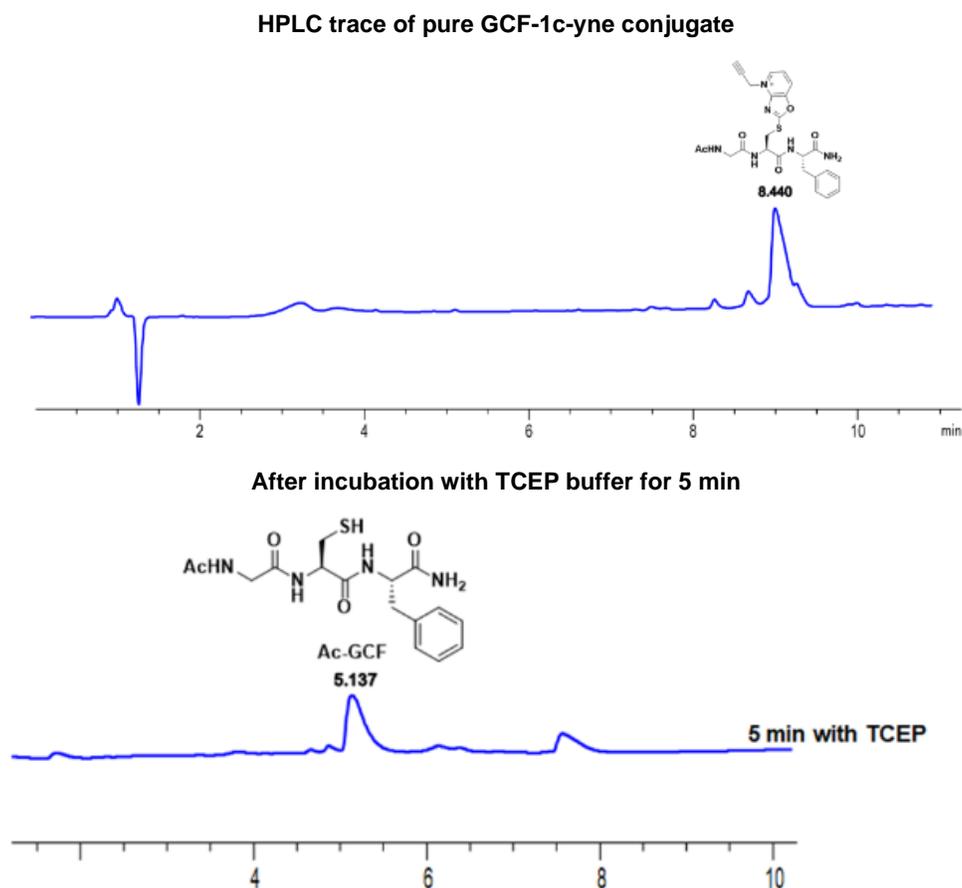
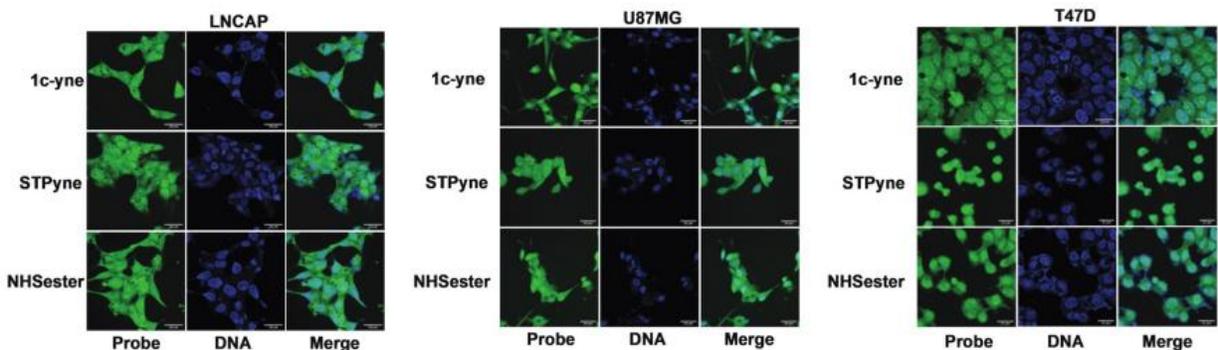


Figure 41. Stability study of GCF-1c-yne conjugate in TCEP buffer. The reaction between GCF and 1c-yne can be reversed under TCEP buffer.

1.2.4.4 Live cell label ligand amino-acid selectivity in proteome by probe 1c-yne

The ability of TAREs to profile the lysine proteome was further explored by treating living human cells with **1 c-yne**. Using progressively higher doses of **1 c-yne**, NHSester, and STPyne, we incubated three cancer cell lines—LNCaP, U87MG, and T47D—representing prostate, brain, and breast malignancies, respectively. Cells were fixed, permeabilized, and washed to get rid of the unreacted probe after 2 hours. The ability of each probe to mark proteins inside the cell was then directly imaged by conjugating the reacted probe to a fluorophore through CuAAC. Confocal fluorescence imaging demonstrated that, at concentrations ranging from 5 mM to 100 mM, all three cell lines had taken up **1 c-yne**, STPyne, and NHS-ester into the cytosolic and nuclear compartments (Figure 42, Supplementary Figure. 30). Protein labeling across molecular weights as shown by Western blot analysis of LNCaP cells treated with 100 mM of **1 c-yne**, NHS-ester, and STPyne and subsequent fluorescence labeling with CuAAC (Figure 42, Supplementary Figure. 30). The differences in band intensities may be due to 1 c-lysine yne's selectivity as opposed to NHS-ester and STPyne probes' cross-reactivity with cysteine. Last but not least, we looked at how quickly 1 c-yne may enter and mark distinct cellular compartments in living cells. Within the first five min

, intracellular labeling happens at **1 c-yne** concentration as low as 100 nM. (Figure 42, Supplementary Figure. 31). These findings demonstrate how effective **1 c-yne** is as a quick live cell labeling probe.



Live cell probe labeling 1c-yne labeling at drug-like concentrations

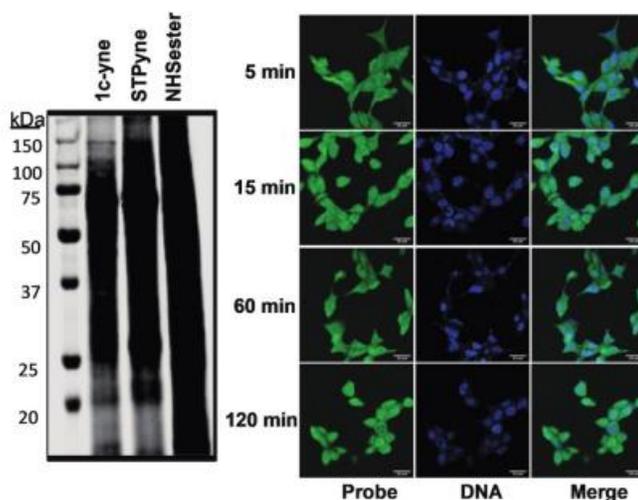
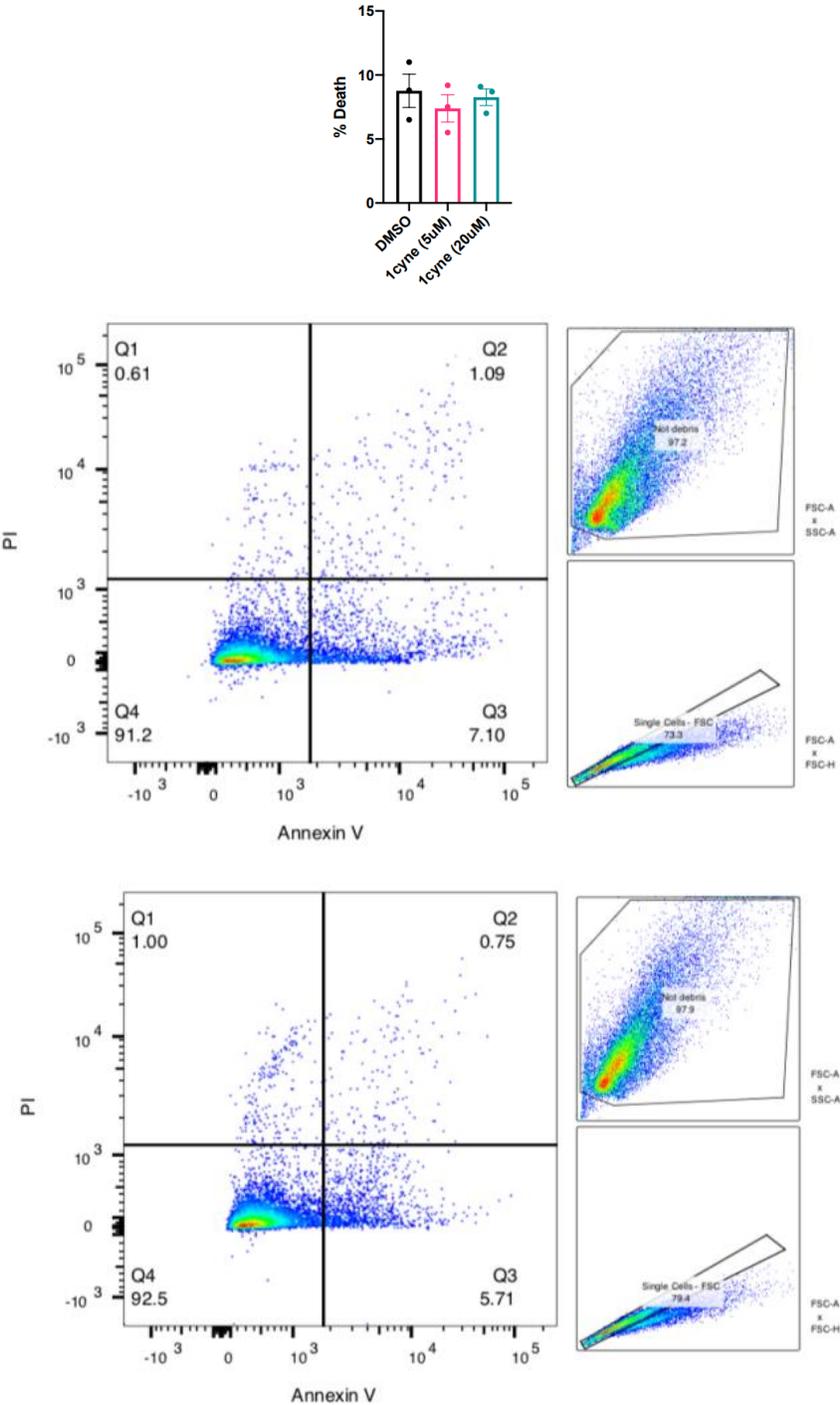


Figure 42. Intracellular probe labeling in LNCAP, U87MG, and T47D. Human LNCAP, U87MG and T47D cells treated with 5 mM 1 c-yne, STPyne, or NHSester for 2 h followed by fixing the cells, washing of unreacted probes and then conjugation with azide fluorophore tags using CuAAC show labeling in multiple cellular compartments. b) Western blot fluorescent analysis of LNCAP cells incubated with 100 mM 1 c-yne, STPyne, or NHSester for 2 h and then conjugated with azide fluorophore tags using CuAAC demonstrates protein labeling across molecular weights.

Then, for the following 24 hours, we performed cell viability tests with **1c-yne** utilizing T47D cells at two different concentrations (5 mM and 20 mM). By comparing the results of the flow

cytometry analysis to the DMSO control, we found no increase in apoptosis or necrosis (Figure 43, Supplementary Figure. 31).



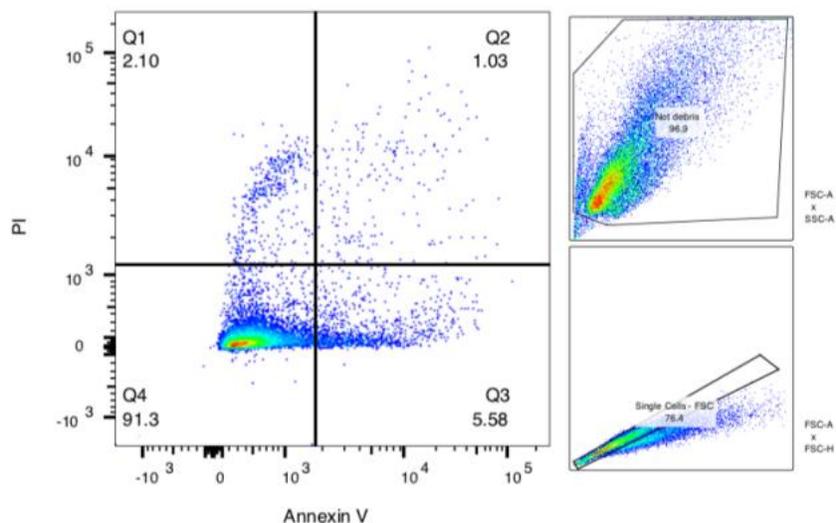


Figure 43. Cell viability studies with 1c-yne. T47D cells treated with indicated concentrations of 1c-yne for 24h did not show an increase in apoptosis/necrosis compared to DMSO control.

1.2.4.5 Computational study of probe 1c-yne

Next Dr. Houk's group carried out the DFT investigation on the reaction between **1 c-yne** and methyl thiolate in order to further evaluate the observed differences in cysteine- and lysine-reactivity for the TAREs. In order to better understand the observed differences in the amino acid reactivity profile, we chose compound **1 c-yne** as the candidate for our DFT experiments. This is due to its hydrolytic stability and observed reactivity with both cysteine and lysine residues. In contrast to the irreversibility of **1 c-yne** modification by amine nucleophiles, the DFT simulations using **1 c-yne** supported the reversibility of the reaction between **1 c-yne** and thiol nucleophiles (Figure 44, Supplementary Figure. 32).

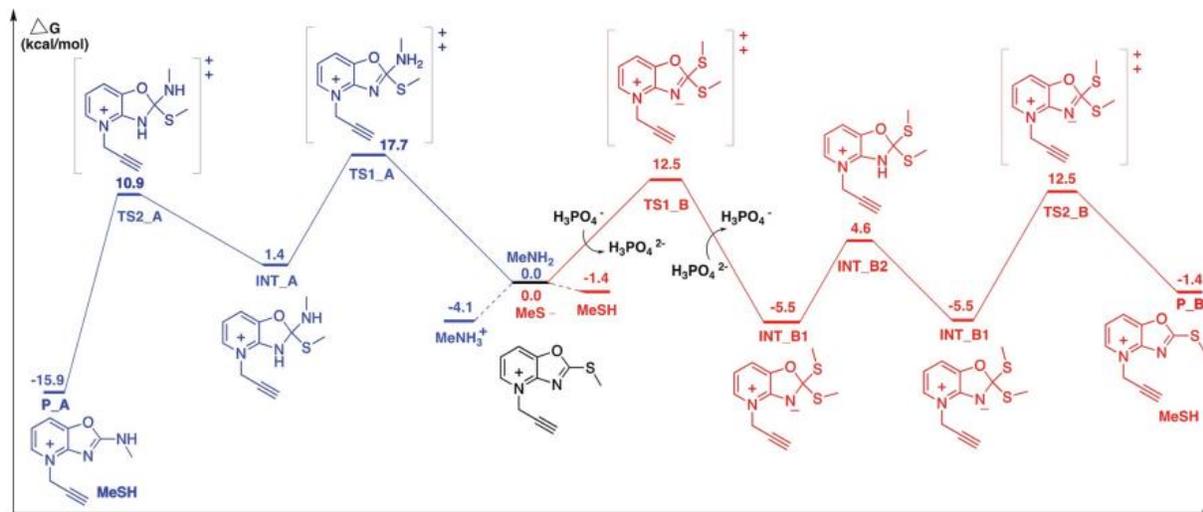


Figure 44. Free energy profiles for SNAr substitution of 1 c-yne with methylamine (blue) and methyl thiolate (red) in water, computed at the wB97X-D/6-311+ +G(d,p) level of theory in SMD water. DFT method: wB97X-D/6-311+ +G(d,p) SMD(H₂O)//B3LYP/6-31+G(D) SMD(H₂O), with Grimme correction for entropy and Head-Gordon correction for enthalpy in 298.15 K. All energies are in kcal mol⁻¹.

1.2.4.6 Conclusion

We have created TAREs that are highly reactive with both cysteine and lysine, but the reaction with cysteine generates a lysine-reactive electrophile, making these probes very selective for the enrichment and production of stable adducts with lysine exclusively. These TARE chemotypes are more resistant to hydrolysis than other lysine-reactive activated esters, such as STPyne and NHS esters. They hence have the potential to function as covalent ligands for lysine in the cellular environment. We demonstrate the remarkable tunability of TAREs, where reactivity and selectivity can be tailored for particular applications by adding various heteroatoms and methylating them. As was predicted, **1 d** and its analog **1 d-yne** were the most reactive, soluble, and selective for lysine due to the strong electron withdrawing nature of O as opposed to S and

greater electrophilicity and solubility of the charged methylation ions as compared to an uncharged moiety. Additionally, we demonstrated that **1 c** is more hydrolytically stable than other activated esters. The reactivity of **1 c** and **1 d** as well as their alkyne derivatives **1 c-yne** and **1 d-yne** are comparable, and both are reactive at low micromolar concentrations. The high permeability, non-toxicity, cellular stability, and cellular activity of TARE probes are shown by the labeling of both cytoplasmic and nuclear proteins in live cells, demonstrating their promise as a probe for quick live cell labeling and covalent inhibitor. As a result of the cysteine-adducts in both probes being reactive electrophiles towards lysines and unstable towards TCEP, both **1 c-yne** and **1 d-yne** enriched exclusively lysine peptide fragments from the cell lysate during the identification of the sites of modification for TAREs. Our experimental results demonstrating the reversibility of the reaction between **1 c-yne** and thiol nucleophiles and the irreversibility of the modification of **1 c-yne** by amine nucleophiles are supported by DFT calculations on **1c-yne**. The 2-methylthio pyridiniumoxazoline ion **1 c** and its alkyne analog **1 c-yne** offer an aromatic, synthetically tractable, non-cytotoxic, and hydrolytically stable electrophile to the arsenal of lysine reactive groups available for protein modification. With regard to protein modification, bioconjugation, material science, activity-based protein profiling, and covalent drug discovery for intractable human proteins, we anticipate that our design of novel probes, simple synthesis to a variety of derivatives, and thorough reactivity and selectivity studies with peptides, proteins, and live cells will encourage their use in a variety of applications.

Reference:

1. Hoyt, E. A.; Cal, P. M. S. D.; Oliveira, B. L.; Bernardes, G. J. L., Contemporary approaches to site-selective protein modification. *Nat. Rev. Chem.* **2019**, *3* (3), 147-171.
2. Boutureira, O.; Bernardes, G. J. L., Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115* (5), 2174-2195.
3. Spicer, C. D.; Davis, B. G., Selective chemical protein modification. *Nat. Commun.* **2014**, *5*.
4. Casi, G.; Huguenin-Dezot, N.; Zuberbuhler, K.; Scheuermann, J.; Neri, D., Site-Specific Traceless Coupling of Potent Cytotoxic Drugs to Recombinant Antibodies for Pharmacodelivery. *J. Am. Chem. Soc.* **2012**, *134* (13), 5887-5892.
5. Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E., Nucleophilic catalysis of hydrazone formation and transimination: Implications for dynamic covalent chemistry. *J. Am. Chem. Soc.* **2006**, *128* (49), 15602-15603.
6. Dirksen, A.; Hackeng, T. M.; Dawson, P. E., Nucleophilic catalysis of oxime ligation. *Angew. Chem., Int. Edit.* **2006**, *45* (45), 7581-7584.
7. Howard, T. S.; Cohen, R. D.; Nwajiobi, O.; Muneeswaran, Z. P.; Sim, Y. E.; Lahankar, N. N.; Yeh, J. T. H.; Raj, M., Amino-Acid-Catalyzed Direct Aldol Bioconjugation. *Org. Lett.* **2018**, *20* (17), 5344-5347.
8. Mahesh, S.; Adebomi, V.; Muneeswaran, Z. P.; Raj, M., Bioinspired Nitroalkylation for Selective Protein Modification and Peptide Stapling. *Angew. Chem., Int. Edit.* **2020**, *59* (7), 2793-2801.
9. Li, X. F.; Zhang, L. S.; Hall, S. E.; Tam, J. P., A new ligation method for N-terminal tryptophan-containing peptides using the Pictet-Spengler reaction. *Tetrahedron Lett.* **2000**, *41* (21), 4069-4073.

10. Obermeyer, A. C.; Jarman, J. B.; Francis, M. B., N-Terminal Modification of Proteins with o-Aminophenols. *J. Am. Chem. Soc.* **2014**, *136* (27), 9572-9579.
11. Purushottam, L.; Adusumalli, S. R.; Singh, U.; Unnikrishnan, V. B.; Rawale, D. G.; Gujrati, M.; Mishra, R. K.; Rai, V., Single-site glycine-specific labeling of proteins. *Nat. Commun.* **2019**, *10*.
12. Chan, A. O. Y.; Ho, C. M.; Chong, H. C.; Leung, Y. C.; Huang, J. S.; Wong, M. K.; Che, C. M., Modification of N-Terminal alpha-Amino Groups of Peptides and Proteins Using Ketenes. *J. Am. Chem. Soc.* **2012**, *134* (5), 2589-2598.
13. Chen, D.; Disotuar, M. M.; Xiong, X. C.; Wang, Y. X.; Chou, D. H. C., Selective N-terminal functionalization of native peptides and proteins. *Chem. Sci.* **2017**, *8* (4), 2717-2722.
14. MacDonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B., One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol.* **2015**, *11* (5), 326-U114.
15. Singudas, R.; Reddy, N. C.; Rai, V., Sensitivity booster for mass detection enables unambiguous analysis of peptides, proteins, antibodies, and protein bioconjugates. *Chem. Commun.* **2019**, *55* (67), 9979-9982.
16. Rosen, C. B.; Francis, M. B., Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* **2017**, *13* (7), 697-705.
17. Crawford, L. A.; Weerapana, E., A tyrosine-reactive irreversible inhibitor for glutathione S-transferase Pi (GSTP1). *Mol. Biosyst.* **2016**, *12* (6), 1768-1771.
18. Hunter, M. J.; Ludwig, M. L., Reaction of Imidoesters with Proteins and Related Small Molecules. *J. Am. Chem. Soc.* **1962**, *84* (18), 3491-&.

19. Bandyopadhyay, A.; Gao, J. M., Iminoboronate-Based Peptide Cyclization That Responds to pH, Oxidation, and Small Molecule Modulators. *J. Am. Chem. Soc.* **2016**, *138* (7), 2098-2101.
20. Wang, X. T.; Di Pasqua, A. J.; Govind, S.; McCracken, E.; Hong, C.; Mi, L. X.; Mao, Y. H.; Wu, J. Y. C.; Tomita, Y.; Woodrick, J. C.; Fine, R. L.; Chung, F. L., Selective Depletion of Mutant p53 by Cancer Chemopreventive Isothiocyanates and Their Structure-Activity Relationships. *J. Med. Chem.* **2011**, *54* (3), 809-816.
21. Zhang, Y. S.; Kensler, T. W.; Cho, C. G.; Posner, G. H.; Talalay, P., Anticarcinogenic Activities of Sulforaphane and Structurally Related Synthetic Norbornyl Isothiocyanates. *Pro. Natl. Acad. Sci. U. S. A.* **1994**, *91* (8), 3147-3150.
22. Musiol, H. J.; Moroder, L., N,N'-Di-tert-butoxycarbonyl-1H-benzotriazole-1-carboxamide derivatives are highly reactive guanidinylation reagents. *Org. Lett.* **2001**, *3* (24), 3859-3861.
23. Patra, A.; Patalag, L. J.; Jones, P. G.; Werz, D. B., Extended Benzene-Fused Oligo-BODIPYs: In Three Steps to a Series of Large, Arc-Shaped, Near-Infrared Dyes. *Angew, Chem., Int. Edit.* **2021**, *60* (2), 747-752.
24. Grimster, N. P.; Connelly, S.; Baranczak, A.; Dong, J.; Krasnova, L. B.; Sharpless, K. B.; Powers, E. T.; Wilson, I. A.; Kelly, J. W., Aromatic Sulfonyl Fluorides Covalently Kinetically Stabilize Transthyretin to Prevent Amyloidogenesis while Affording a Fluorescent Conjugate. *J. Am. Chem. Soc.* **2013**, *135* (15), 5656-5668.
25. Zhao, Q.; Ouyang, X.; Wan, X.; Gajiwala, K. S.; Kath, J. C.; Jones, L. H.; Burlingame, A. L.; Taunton, J., Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc.* **2017**, *139* (2), 680-685.

26. Asano, S.; Patterson, J. T.; Gaj, T.; Barbas III, C. F., Site-Selective Labeling of a Lysine Residue in Human Serum Albumin. *Angew, Chem., Int. Edit.* **2014**, *53* (44), 11783-11786.
27. Ward, C. C.; Kleinman, J. I.; Nomura, D. K., NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12* (6), 1478-1483
28. Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F., Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **2017**, *9* (12), 1181-1190.
29. Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; Chalker, J. M.; Corzana, F.; Jiménez-Osés, G.; Bernardes, G. J. L., Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **2018**, *140* (11), 4004-4017.
30. Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010**, *468* (7325), 790-795.
31. Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogoy, M., Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol* **2000**, *7* (8), 569-581.
32. Adam, G. C.; Cravatt, B. F.; Sorensen*, E. J., Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem. Biol.* **2001**, *8* (1), 81-95.
33. Berger, A. B.; Witte, M. D.; Denault, J.-B.; Sadaghiani, A. M.; Sexton, K. M. B.; Salvesen, G. S.; Bogoy, M., Identification of Early Intermediates of Caspase Activation Using Selective Inhibitors and Activity-Based Probes. *Mol. Cell* **2006**, *23* (4), 509-521.

34. Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F., Proteome-wide covalent ligand discovery in native biological systems. *Nature* **2016**, *534* (7608), 570-574.
35. Embaby, A. M.; Schoffelen, S.; Kofoed, C.; Meldal, M.; Diness, F., Rational Tuning of Fluorobenzene Probes for Cysteine-Selective Protein Modification. *Angew, Chem., Int. Edit.* **2018**, *57* (27), 8022-8026.
36. Shannon, D. A.; Banerjee, R.; Webster, E. R.; Bak, D. W.; Wang, C.; Weerapana, E., Investigating the Proteome Reactivity and Selectivity of Aryl Halides. *J. Am. Chem. Soc* **2014**, *136* (9), 3330-3333.
37. Wong, H. L.; Liebler, D. C., Mitochondrial Protein Targets of Thiol-Reactive Electrophiles. *Chem. Res. Toxicol.* **2008**, *21* (4), 796-804.
38. Motiwala, H. F.; Kuo, Y.-H.; Stinger, B. L.; Palfey, B. A.; Martin, B. R., Tunable Heteroaromatic Sulfones Enhance in-Cell Cysteine Profiling. *J. Am. Chem. Soc.* **2020**, *142* (4), 1801-1810.
39. Zambaldo, C.; Vinogradova, E. V.; Qi, X.; Iaconelli, J.; Suciu, R. M.; Koh, M.; Senkane, K.; Chadwick, S. R.; Sanchez, B. B.; Chen, J. S.; Chatterjee, A. K.; Liu, P.; Schultz, P. G.; Cravatt, B. F.; Bollong, M. J., 2-Sulfonylpyridines as Tunable, Cysteine-Reactive Electrophiles. *J. Am. Chem. Soc.* **2020**, *142* (19), 8972-8979.
40. Evans, M. J.; Saghatelian, A.; Sorensen, E. J.; Cravatt, B. F., Target discovery in small-molecule cell-based screens by in situ proteome reactivity profiling. *Nat. Biotechnol.* **2005**, *23* (10), 1303-1307.

41. Weerapana, E.; Simon, G. M.; Cravatt, B. F., Disparate proteome reactivity profiles of carbon electrophiles. *Nat. Chem. Biol.* **2008**, *4* (7), 405-407.
42. Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G., Facile Conversion of Cysteine and Alkyl Cysteines to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins. *J. Am. Chem. Soc.* **2008**, *130* (15), 5052-5053.
43. Toda, N.; Asano, S.; Barbas III, C. F., Rapid, Stable, Chemoselective Labeling of Thiols with Julia–Kocięński-like Reagents: A Serum-Stable Alternative to Maleimide-Based Protein Conjugation. *Angew. Chem., Int. Edit.* **2013**, *52* (48), 12592-12596.
44. Bernardim, B.; Cal, P. M. S. D.; Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Albuquerque, I. S.; Perkins, E.; Corzana, F.; Burtoloso, A. C. B.; Jiménez-Osés, G.; Bernardes, G. J. L., Stoichiometric and irreversible cysteine-selective protein modification using carbonylacrylic reagents. *Nat. Commun.* **2016**, *7* (1), 13128.
45. Kung, K. K.-Y.; Ko, H.-M.; Cui, J.-F.; Chong, H.-C.; Leung, Y.-C.; Wong, M.-K., Cyclometalated gold(III) complexes for chemoselective cysteine modification via ligand controlled C–S bond-forming reductive elimination. *Chem. Commun.* **2014**, *50* (80), 11899-11902.
46. Roberts, D. D.; Lewis, S. D.; Ballou, D. P.; Olson, S. T.; Shafer, J. A., Reactivity of small thiolate anions and cysteine-25 in papain toward methyl methanethiosulfonate. *Biochemistry* **1986**, *25* (19), 5595-5601.
47. King, T. P.; Li, Y.; Kochoumian, L., Preparation of protein conjugates via intermolecular disulfide bond formation. *Biochemistry* **1978**, *17* (8), 1499-1506.

48. Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R., Protein Modification, Bioconjugation, and Disulfide Bridging Using Bromomaleimides. *J. Am. Chem. Soc.* **2010**, *132* (6), 1960-1965.
49. Chudasama, V.; Smith, M. E. B.; Schumacher, F. F.; Papaioannou, D.; Waksman, G.; Baker, J. R.; Caddick, S., Bromopyridazinedione-mediated protein and peptide bioconjugation. *Chem. Commun.* **2011**, *47* (31), 8781-8783.
50. Shiu, H.-Y.; Chan, T.-C.; Ho, C.-M.; Liu, Y.; Wong, M.-K.; Che, C.-M., Electron-Deficient Alkynes as Cleavable Reagents for the Modification of Cysteine-Containing Peptides in Aqueous Medium. *Chemistry – A European Journal* **2009**, *15* (15), 3839-3850.
51. Zhang, Y.; Zhou, X.; Xie, Y.; Greenberg, M. M.; Xi, Z.; Zhou, C., Thiol Specific and Tracelessly Removable Bioconjugation via Michael Addition to 5-Methylene Pyrrolones. *J. Am. Chem. Soc.* **2017**, *139* (17), 6146-6151.
52. Yu, J.; Yang, X.; Sun, Y.; Yin, Z., Highly Reactive and Tracelessly Cleavable Cysteine-Specific Modification of Proteins via 4-Substituted Cyclopentenone. *Angew, Chem., Int. Edit.* **2018**, *57* (36), 11598-11602.
53. Deng, J.-R.; Chung, S.-F.; Leung, A. S.-L.; Yip, W.-M.; Yang, B.; Choi, M.-C.; Cui, J.-F.; Kung, K. K.-Y.; Zhang, Z.; Lo, K.-W.; Leung, Y.-C.; Wong, M.-K., Chemoselective and photocleavable cysteine modification of peptides and proteins using isoxazoliniums. *Commun. Chem.* **2019**, *2* (1), 93.
54. Yan, T.; Desai, H. S.; Boatner, L. M.; Yen, S. L.; Cao, J.; Palafox, M. F.; Jami-Alahmadi, Y.; Backus, K. M., SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome**. *ChemBioChem* **2021**, *22* (10), 1841-1851.

55. Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; Cravatt, B. F.; Gygi, S. P., Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **2021**, *39* (5), 630-641.
56. Aoyagi, N.; Endo, T., Synthesis of five- and six-membered cyclic guanidines by guanylation with isothiuronium iodides and amines under mild conditions. *Synth. Commun.* **2017**, *47* (5), 442-448.
57. Sanda, F.; Araki, H.; Masuda, T., Synthesis and Properties of Serine- and Threonine-Based Helical Polyacetylenes. *Macromolecules* **2004**, *37* (23), 8510-8516.
58. Gong, Z.; Liu, Q.; Xue, P.; Li, K.; Song, Z.; Liu, Z.; Jin, Y., Novel chiral thiazoline-containing N—O ligands in the asymmetric addition of diethylzinc to aldehydes: substituent effect on the enantioselectivity. *Appl. Organomet. Chem.* **2012**, *26* (3), 121-129.
59. Hara, Y.; Fujii, T.; Kashida, H.; Sekiguchi, K.; Liang, X.; Niwa, K.; Takase, T.; Yoshida, Y.; Asanuma, H., Coherent Quenching of a Fluorophore for the Design of a Highly Sensitive In-Stem Molecular Beacon. *Angew. Chem. Int. Ed.* **2010**, *49* (32), 5502-5506.
60. Bethge, L.; Jarikote, D. V.; Seitz, O., New cyanine dyes as base surrogates in PNA: Forced intercalation probes (FIT-probes) for homogeneous SNP detection. *Bioorg. Med. Chem.* **2008**, *16* (1), 114-125.
61. Imaizumi, T.; Kobayashi, A.; Otsubo, S.; Komai, M.; Magara, M.; Otsubo, N., The discovery and optimization of a series of 2-aminobenzoxazole derivatives as ChemR23 inhibitors. *Bioorg. Med. Chem.* **2019**, *27* (21), 115091.

62. Shibuya, K.; Kawamine, K.; Miura, T.; Ozaki, C.; Edano, T.; Mizuno, K.; Yoshinaka, Y.; Tsunenari, Y., Design, synthesis and pharmacology of aortic-selective acyl-CoA: Cholesterol O-acyltransferase (ACAT/SOAT) inhibitors. *Bioorg. Med. Chem.* **2018**, *26 (14)*, 4001-4013.

Chapter 2: Chemical Tool for Tagging Mono-Methyl Lysine and K-me-Directed Modification

2.1 General Introduction of Mono-Methyl Lysine

Despite decades of intensive research, the putative roles of lysine monomethylation (Kme1) PTMs in regulating epigenetics, chromatin assembly, gene expression, and other biological processes remain unclear and limited¹⁻⁴. The main cause of this uncertainty is the inability of conventional experimental methods to identify methylation lysine PTMs universally. This is mostly due to the fact that the addition of small methyl groups very slightly alters the physiochemical characteristics of proteins including mass, charge, and hydrophobicity^{5,6}. To date, attempts to identify and characterize lysine methylation have depended on the use of affinity reagents such as antibodies and methyl binding domains MBD^{7,8} (Figure 45). However, these affinity reagents are unable to detect all lysine methylation sites due to their sequence-specificity and inability to differentiate between mono-, di-, and trimethylated lysine⁹⁻¹¹. The fundamental disadvantage of MBDs is that they enrich methylation proteins without identifying the individual methylated residues. In addition, MBDs require flanking amino acids on both sides of the methylated lysine for detection, making them incapable of detecting methylation sites on trypsin digested fragments, which are frequently employed in proteomic study of other PTMs. Others have investigated functionalizing the methyl donor S-adenosylmethionine (SAM) with biorthogonal groups such as alkynes or azides, which are transferred to the substrate proteins in place of the methyl group¹²⁻¹⁶ (Figure 45). However, the unnatural SAM analogs exhibit cross-reactivity with natural SAM, and not all lysine methyltransferases (KMTs) accept the changed cofactors. Mass spectrometry is one of the most common techniques for lysine methylation detection (MS). Although robust MS analysis is challenged by analytical limitations,

such as a low natural abundance of methyl lysine PTMs in complex mixtures^{17,18}. In addition, the change in mass caused by one methyl group is equal to the substitutions of some amino acids (e.g., Val vs. Leu, Asn vs. Gln, Asp vs. Glu), resulting in false identification (Figure 45).

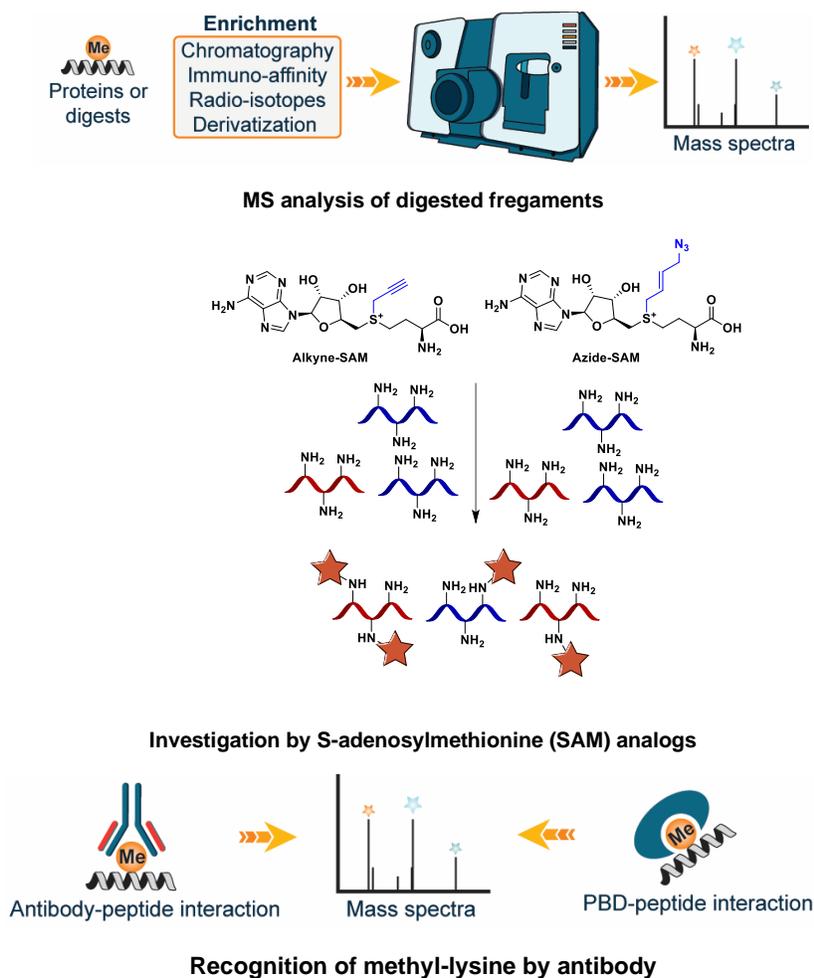


Figure 45. Current method for detecting methyl-lysine

Due to the large number of known methyltransferases KMTs and demethylases (KDMs) involved for lysine methylation, only 5000 Lys methylation sites (combined mono-, di-, or tri-) have been found thus far, despite evidence of their vast occurrence^{19,20}. Our group has recently developed STaR chemistry for selective labeling of monomethyl lysine Kme for identifying proteins with Kme1 (Figure 46).

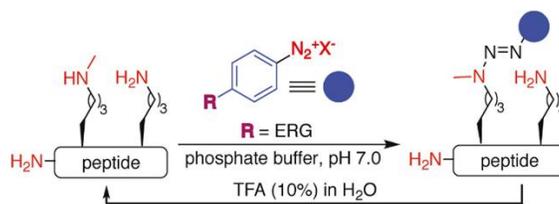


Figure 46. Selective triazene reaction (STaR) for tagging monomethyl lysine post-translational modifications

however, we are unable to (provide residue-specific information) determine the site of monomethyl lysine, which is necessary for addressing specific biological/fundamental questions. This was mostly due to the monomethyl lysine triazene product's sensitivity to mildly acidic conditions²¹, making it incompatible with the MS-based proteomics and fluorosequencing required for the detection of Kme sites (Figure 47). All together, none of the previous methods can be used to selectively label Kme1 sites in a pan-specific way. This means that a powerful chemical method is needed to label Kme1 in a specific way.

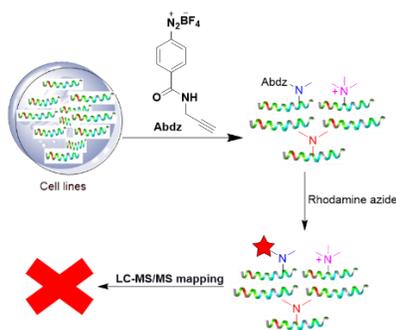


Figure 47. False identification of labeling site after profiling

2.2 Coarctate Cyclization for Selective labeling of Monomethyl Lysine Posttranslational Modifications

The Kme-triazene product is stabilized by functionalizing the ortho position of the diazonium salt with an ethyne group, followed by coarctate cyclization of the triazene-ene-yne product in the presence of CuCl to yield an acid-stable 2H-indazole-3-carbaldehyde

fluorophore/chromophore²² (Figure 48). Consequently, it is compatible with both MS-based proteomics and fluorescence sequencing. We demonstrated that the TCC method is pan-specific and selectively modifies Kme1 in various histone peptides regardless of the sequence, the presence of nearby PTMs, and the presence of multiple Kme1 on a single peptide with different tags, including affinity tags and fluorophores. **Triazene Coarctate Cyclization (TCC)** process was used to enrich Kme1 proteins from the nuclear extract. Notably, we identified Kme sites on a peptide with high efficiency using single-molecule fluorosequencing, and we identified unknown Kme1 proteins and sites using chemoproteomics. Our approach gives both quantitative and residue-specific information regarding total protein lysine monomethylation. In the scientific literature, there are no other pan-selective chemical techniques for the detection of Kme1 sites by fluorosequencing and MS proteomics.

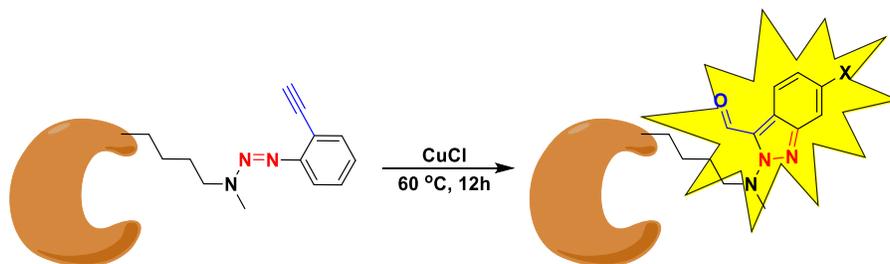


Figure 48. Triazene cyclization for selective tagging monomethyl lysine with chromophore

2.2.1 Development of Triazene Coarctate Cyclization (TCC)

We inserted azide and alkyne at the ortho position to stabilize triazene product generated by selective reactivity of secondary amine with phenyl diazonium ion and generate stable indazole cyclic product at the secondary amine. Starting with a small model chemical, proline methylester, we performed reactions with 2-azido aniline and 2-ethyne aniline. In the first step, aniline was transformed in situ to diazonium ion using NaNO_2 under acidic conditions, which subsequently reacted selectively with proline methyl ester at pH 7.5 in 1 hour to produce triazene. The Triazene-ene-azide or Triazene-ene-yne coarctate cyclization of triazene-ene-azide

or triazene-ene-yne produced benzotriazole²³ (43%) and 2H-indazole-3-carbaldehyde (52%) in reasonable yields (Figure 49, Supplementary Figure. 1), respectively. ¹H and ¹³C NMR were used to completely describe the products.

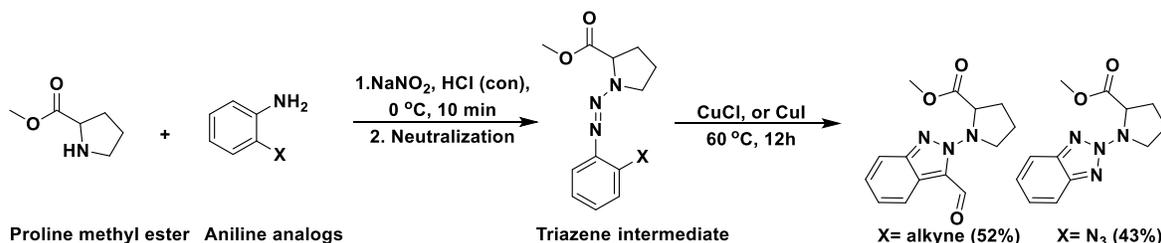
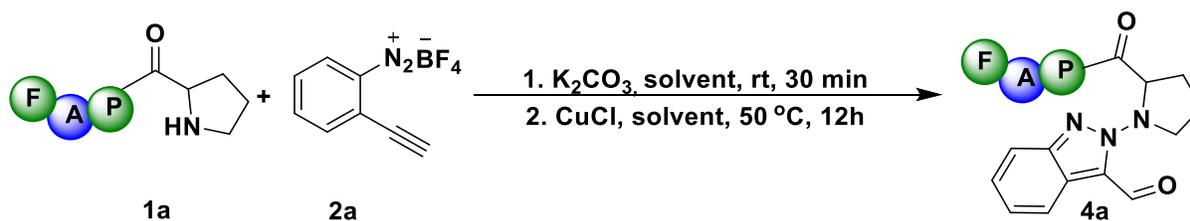


Figure 49. Synthesis of benzotriazole and 2H-indazole-3-carbaldehyde by TCC reaction.

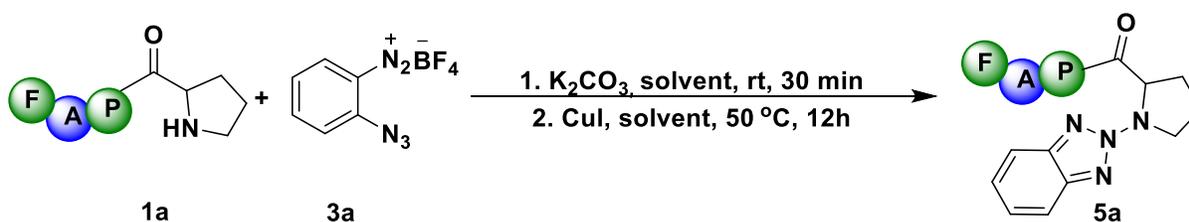
We further optimized the reaction conditions on a model peptide PAF **1a** using 2-ethyne phenyldiazonium ion **2a** and 2-azido phenyldiazonium ion **3a** under varied pH (7.5 to 9.5), temperatures (RT to 60 °C), catalysts (CuCl, CuI, IPrAuCl and AgSbF₆) and additives (DPSO). 2-ethyne phenyldiazonium ion **2a** (10 equiv.) and CuCl (15 equiv.) resulted in the formation of a stable 2H-indazole-3-carbaldehyde **4a** with a peptide PAF **2a** at a high conversion (76%) at 50 °C in ACN:sodium phosphate buffer (9:1) (10 mM, pH 7, Figure 50, Supplementary Figure. 1). Due to the poor solubility of the 2-ethyne phenyldiazonium ion **2a** in aqueous solution, a lower conversion to 2H-indazole-3-carbaldehyde **4a** was observed in ACN:sodium phosphate buffer (1:9) solvent.



Entry	Catalyst	Eq of 2a	Temperature °C	Solvent	Conversion (%)
1	CuCl	3	50 °C	ACN/Nap (1:1)	25 %
2	CuCl	3	50 °C	ACN/Nap (9:1)	46 %
3	CuCl	10	50 °C	ACN/Nap (9:1)	76 %
4	CuI	3	50 °C	ACN/Nap (9:1)	0 %
5	CuCl	3	50 °C	ACN/Nap (1:9)	3 %
6	CuCl	3	50 °C	DMF/Nap (1:9)	14.2 %
7	CuCl	3	rt	ACN/Nap (1:1)	0 %

Figure 50. Optimization of triazene-ene-yne coarctate cyclization with PAF. The best conversion (76 %) we observed was 10 eq of 2a in ACN/Nap (9:1) at 50 °C for 12 h.

The reaction with azido phenyldiazonium ion **3a** produced moderate yields (57%) of benzotriazole product **5a** with peptide PAF **2a** (Figure 51, Supplementary Figure. 2). We continued our investigation with 2-ethyne phenyldiazonium ion **2a** since it produced aldehyde that can be directly utilized for enrichment and functionalization with fluorophores and affinity tags.

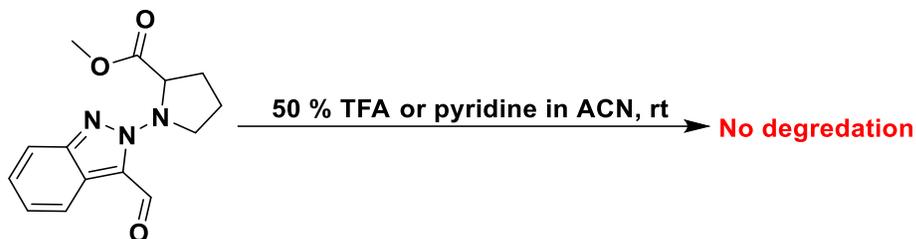


Entry	Catalyst	Eq of 3a	Temperature °C	Solvent	Conversion (%)
1	CuI	3	50 °C	ACN/Nap (1:1)	35 %
2	CuCl	3	50 °C	ACN/Nap (1:1)	15 %
3	CuI	3	50 °C	ACN/Nap (9:1)	17 %
4	CuI	10	50 °C	ACN/Nap (9:1)	57 %
5	CuI	3	50 °C	ACN/Nap (1:9)	30 %
6	CuI	3	50 °C	DMF/Nap (1:9)	12.5 %
7	CuI	3	rt	ACN/Nap (1:1)	10 %

Figure 51. Optimization of Triazene-ene-azide cyclization PAF. The best conversion (57 %) we observed was 10 eq of 3a in ACN/Nap (9:1) at 50 °C for 12 h.

2.2.2 Stability Studies

Because of the high sensitivity to moderate acidic conditions in our previously discovered selective triazene reaction STaR, one of the possible difficulties in employing it to identify the locations of Kme, we investigated the stability of the 2H-indazole-3-carbaldehyde product under acidic conditions. Pro-OMe-2H-indazole-3-carbaldehyde was incubated at room temperature in 50% TFA in ACN, and the product's stability was determined by injecting samples into HPLC at regular intervals. 6 h of acidic conditions resulted in no degradation of Pro-OMe-2H-indazole-3-carbaldehyde, in contrast to the total destruction of the triazene product in 0.1% TFA solution within 5 min. Critical to fluorosequencing is the stability of the resultant product towards pyridine, which is essential for Edman's degradation. We determined that the Pro-OMe-2H-indazole-3-carbaldehyde product is highly stable under basic conditions (50% pyridine in ACN, 6h), therefore the TCC method is entirely compatible for identifying Kme sites by both MS proteomics and fluorescence sequencing (Figure. 52, Supplementary Figure. 3).

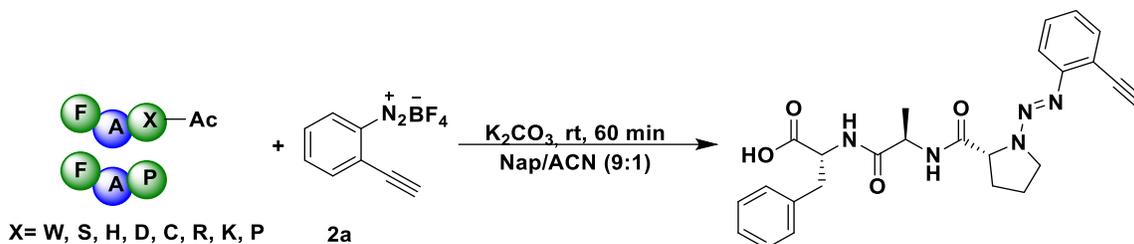


Pro-OMe-2H-indazole-3-carbaldehyde

Figure 52. Stability study of 2H-indazole-3-carbaldehyde group in 50 % TFA and 50 % pyridine. The forming product was incubating in 50 % TFA and 50 % of pyridine for 12 h to confirm the high robustness of indazole product.

2.2.3 Chemoselectivity studies for the formation of 2H-indazole-3-carbaldehyde

The chemoselective studies with 2-ethyne phenyldiazonium ion **2a** under optimized conditions with varying peptides OAc-XAF 1a-1l containing reactive amino acids (X = P, H, R, D, S, C, K, W, and Y) and varying lysine methylation states (Kme1, Kme2, and Kme3) demonstrated that the TCC reaction is highly chemoselective (i.e. N-terminal proline 4a and Kme 4b). Under the reaction conditions, we noticed the formation of a diazo complex with Tyr diazo-OAc-YAF, but this byproduct did not interfere with the analysis of the Kme1-2H-indazole-3-carbaldehyde product (Figure 53, Supplementary Figure. 4).



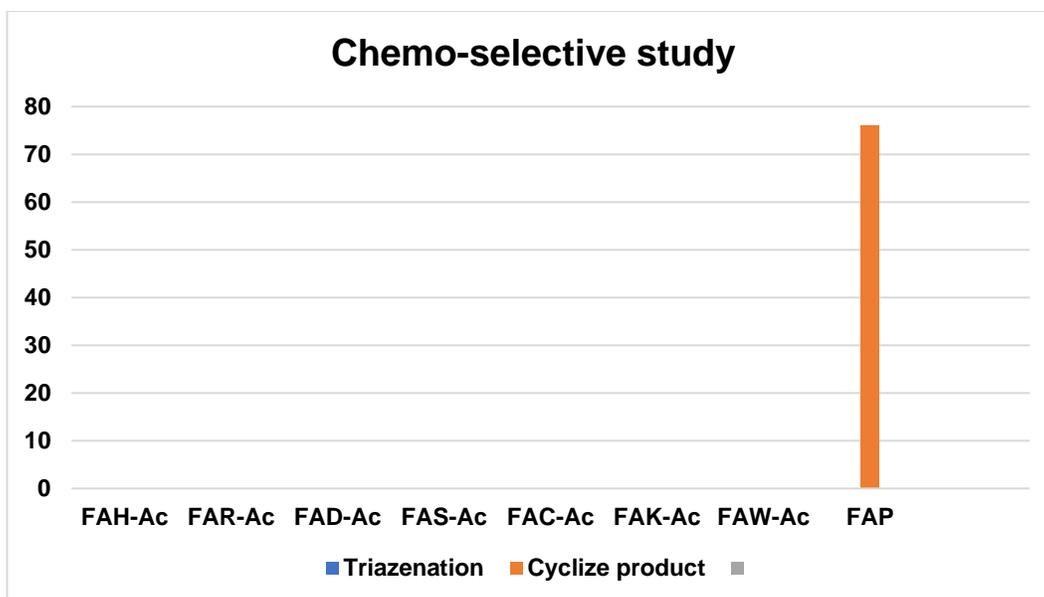
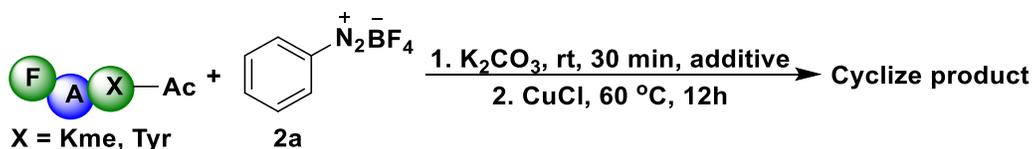


Figure 53. Chemo-selective study of triazene cyclization. Different tripeptides were used to test the chemo-selectivity of the TCC reaction. The result indicated the high specificity of the TCC reaction.

2.2.4 Tyrosine Vs Kme modification

To make the TCC reaction highly selective for tagging Kme1, we modified the reaction conditions by lowering the pH to 7.5, decreasing the reaction time (5 min), quenching the unreacted probe **2a** with potassium iodide KI, and using lower equivalents (3 equiv.) of probe **2a**. Despite the fact that we observed significant selectivity (high conversion) for Kme1 over Tyr, none of these reaction conditions were able to completely stop the modification at Tyr (Figure 54, Supplementary Figure. 5).



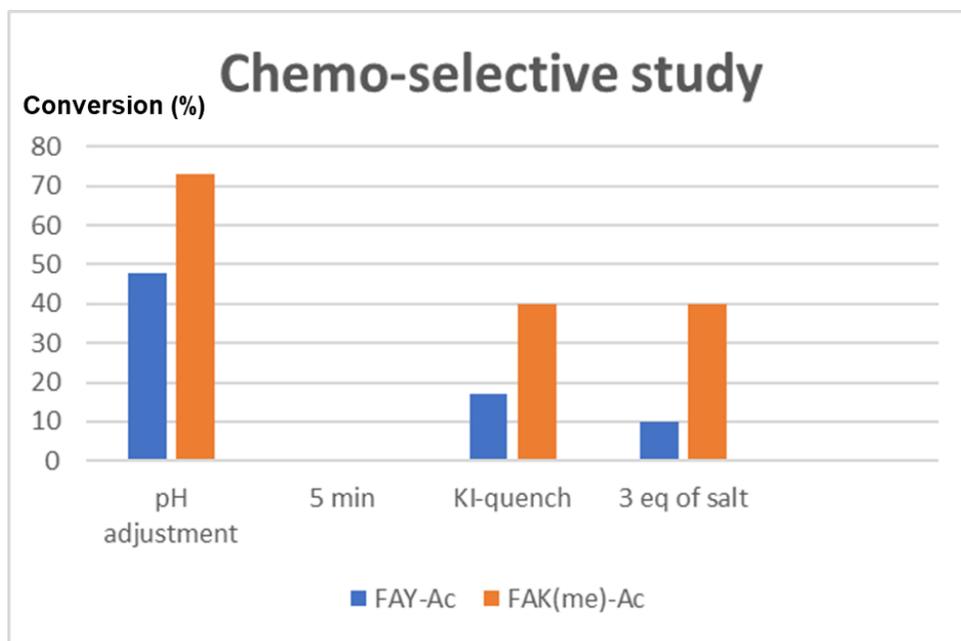


Figure 54. Optimization of tyrosine blocking reagents. (i) The pH adjustment: the pH of the reaction mixture was adjusted to 6.5 by 1.0 M HCl. (ii) The reaction mixture was incubated at 25 °C in the incubator for 5 min (Triazination). (iii) The 10 eq of KI was added to quench the excess diazonium salt. (iv) 3 eq of 2a was used for the first step.

Due to the low abundance of monomethyl lysine Kme relative to tyrosine, high equivalents of the probe **2a** are required for Kme labeling; therefore, we intended to completely block tyrosine before modifying Kme1 with TCC. We proceeded with a peptide containing both Pro Kme1 and Tyr, **1m** PY, and conducted a reaction with 1,3 diphenyl propynone, a recently discovered method for selectively labeling Tyr in our lab. We observed the modification of lysine, cysteine, Kme1, and Tyr with 1,3 diphenyl propynone under the reaction conditions, but the treatment with acidic solution 1:3 (2 M HCl/co-solvent) for 4 h at room temperature reversed the modification on lysine, cysteine, and Kme1, leaving only Tyr modified **1m'** (100% conversion determined by HPLC and LC-MS) under the reaction conditions. (as opposed to the well-known Suttex and Suffex probes that modify other amino acids, such as lysine, and have lower

conversions with Tyr. Figure 55). We chose 1,3 diphenyl propynone over the well-known Suttex and Suffex probes due to their non-selective modification of other amino acids, including lysine, and lower Tyr. With the modified Tyr peptide 1m' in hand, we modified Pro under optimized TCC reaction conditions and observed the formation of a **2m** peptide-2H-indazole-3-carbaldehyde product with a very high conversion (80%) as determined by HPLC and MS. This method resulted in the highly efficient dual functionalization of the peptide (Supplementary Figure. 6).

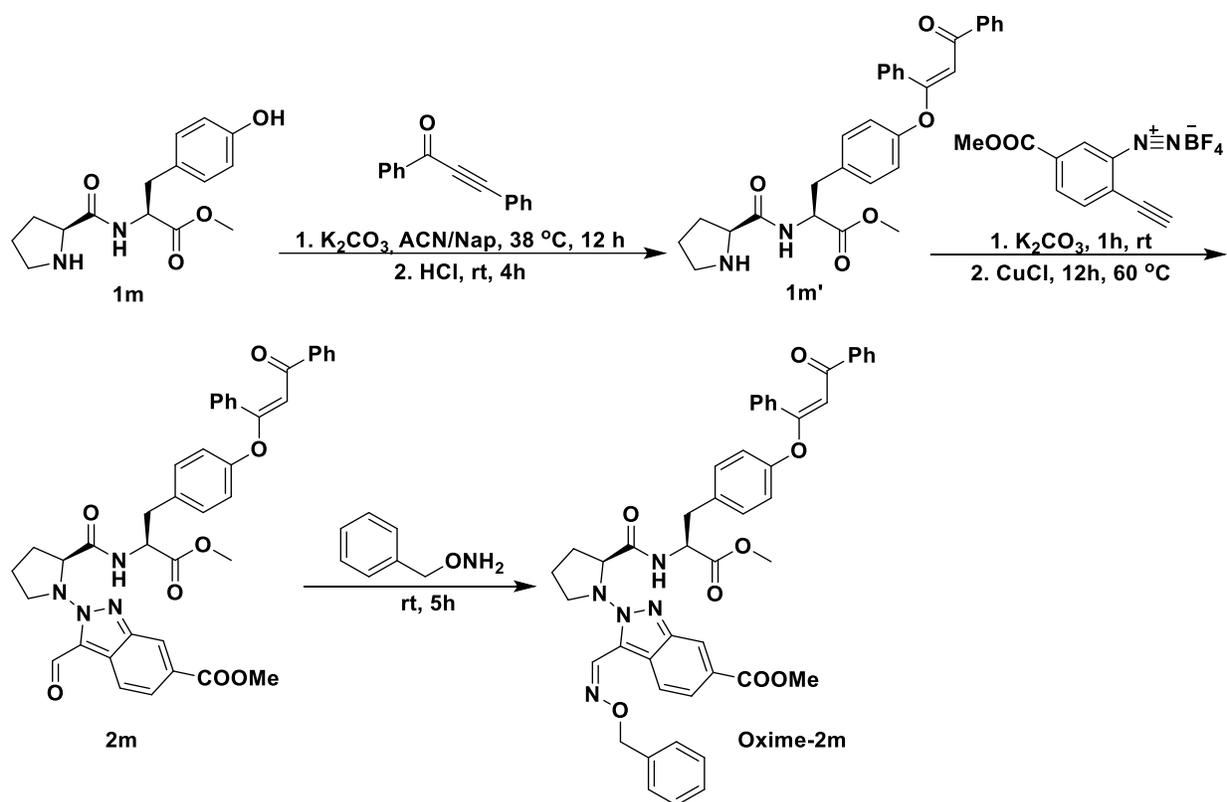


Figure 55. Selective modification of proline by TCC reaction with tyrosine blocking reagent 1,3 diphenyl propynone. Compound 1m was treated with 1,3 diphenyl propynone at 38 °C for 12 h, then HCl was added to reverse the proline modified product for the triazene cyclization. The benzyl hydroxylamine was applied to enrich 2m from the reaction mixture.

2.2.5 Substrate scope with varying 2-ethyne phenyldiazonium ions.

To determine the scope of the TCC reaction on peptides of different lengths, we modified a long peptide, OAc-GKmeGKAKF, with 2-ethyne phenyldiazonium ion **2a** in ACN:sodium phosphate buffer (1:9) and observed the selective modification of Kme1 but with a lower conversion (30%). We hypothesized that the lower conversion was due to the second cyclization step and that adding EWG at the para position to the ethyne group would increase the Kme-indazole product's reactivity and yields. We synthesized and investigated various ethyne phenyldiazonium ions

with EWG at the 5th position (Figure. 56), including F (**2b**), CF₃ (**2c**), and CO₂Me (**2d**), for selective modification of Kme1 on a long peptide OAc-GKmeGKAKF, and obtained high conversions to corresponding Kme-2H-indazole-3-carbaldehyde (**4c**; 71%, **4d**; 53%, **4e**; 83%. Figure 57, Supplementary Figure. 7), respectively. the maximum conversion to 2H-indazole-3-carbaldehyde was achieved with ester-substituted ethyne phenyldiazonium ions **2d** under aqueous reaction conditions (ACN:sodium phosphate (1:9) buffer). The diazonium salt analogs **2a**, **2b**, **2c**, and **2d** were synthesized from corresponding aniline starting materials. The installation of trimethylsilylacetylene was achieved by the known Sonogashira coupling^{24, 25, 26} procedures. Intermediate **a9**, **a10**, and **a11** were treated with potassium carbonate for the TMS deprotection, then the deprotected aniline intermediates **a12**, **a13** and **a14** were converted into corresponding diazonium salt analogs by Jacob's protocol²⁷.

2.2.6 Pan-specificity of TCC: Further Diversification

With the optimized conditions for the formation of peptide-2H-indazole-3-carbaldehyde, we next demonstrated the pan specificity of the TCC method by carrying out reactions with various peptides of different sizes and amino acid compositions with Kme at varying positions, including histone H3.3 peptide fragments, which are known to be frequently methylated at K4, K9, K27, and K36 and are involved in the regulation of biological processes and disease. Using solid-phase peptide synthesis, we synthesized H3.3 peptide fragments Kme14K9 (ARTKme1QTARKS) **1n** and Kme19K14 (ARTKme2STGGKA) **1o**. Under the optimized reaction conditions using ester-substituted ethyne phenyldiazonium ions **2d**, all Kme1 containing peptides were converted to peptide-2H-indazole (Figure 58, Supplementary Figure. 8). These results collectively confirmed the high chemoselectivity and broad specificity of TCC towards monomethyllysine Kme1.

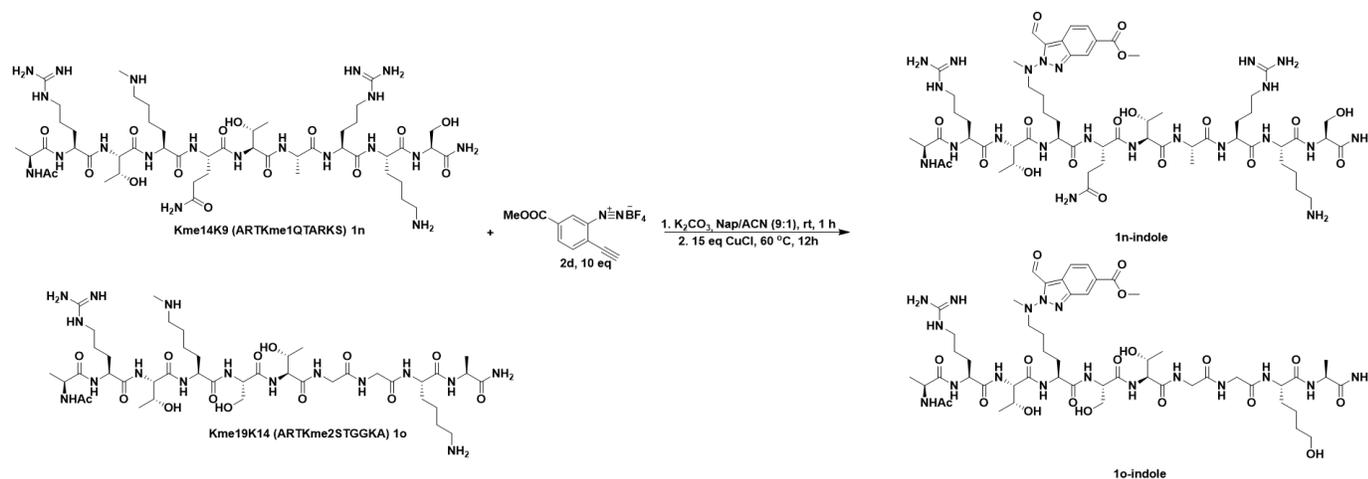


Figure 58. Pan-specificity of triazene coarctate cyclization with truncated histone peptides. The TCC reaction is highly specific to the Kme residue, even in the presence of other reactive amino acids. In 1o example, we observed the lysine side chain was converted into alcohol which was purposed early²¹.

Using aldehyde-specific reactions such as oxime chemistry and thiazolidine chemistry, we modified peptide-2H-indazole-3-carbaldehyde with a variety of functional groups. The modification of the peptide OAc-KmeAF **1j** with benzylhydroxylamine and cysteine methyl ester produced 99 % of the oxime-product **5a** and 99 % of the thiazolidine-product **5b**.

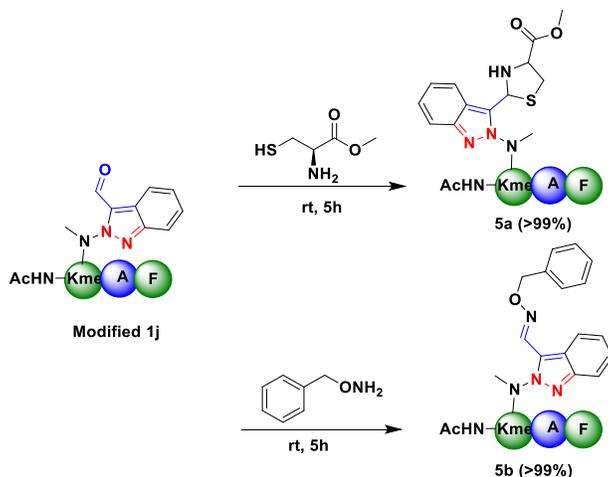


Figure 59. Enrichment of modified OAc-KmeAF by cysteine condensation and oxime chemistry. The indazole product can be easily enrich after modification due to the aldehyde handle.

In addition, 2H-indazole-3-carbaldehyde products **4f-4g** of histone peptides Kme14K9 (ARTKme1QTARKS) and Kme19K14 (ARTKme2STGGKA) were treated with benzylhydroxylamine to produce oxime-products **5c** (99 %) and **5d** (99 %), respectively (Figure 59, Supplementary Figure. 9). Next, the 2H-indazole-3-carbaldehyde products **4f-4g** of histone peptides Kme14K9 (ARTKme1QTARKS) and Kme19K14 (ARTKme2STGGKA) were functionalized with cysteine methyl ester to generate thiazolidine-products **5e** (99 %) and **5f** (99 %), respectively (Figure 60, Supplementary Figure. 10). These findings demonstrate the capacity of TCC to modify and diversify Kme1-containing peptides with distinct functional groups.

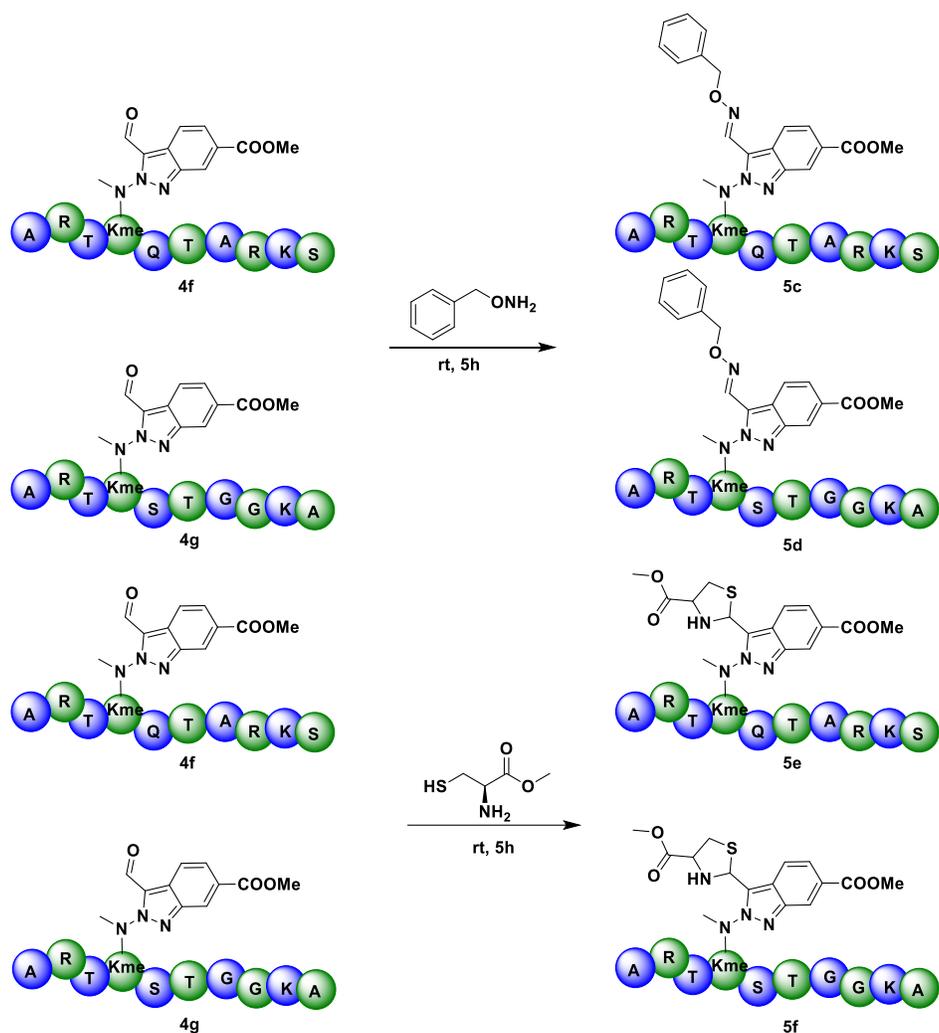


Figure 60. Enrichment of labeling histone peptides by cysteine condensation and oxime chemistry. The indazole product can be easily enrich after modification due to the aldehyde handle.

2.2.7 Selective labeling of Kme1 peptides in a complex cell lysate mixture by TCC

To evaluate the robustness of our TCC method in labeling Kme1 peptides in a complex mixture, the cell lysate spiked with two distinct histone H3 peptides, NH₂-FKme2AGSKmeFS **1p** and Ac-AKTKQTAFKmeS **1q**, was treated with Suttex for one hour to in pull-down experiments, 2H-indazole-3-carbaldehyde products were enriched from the complex mixture by using

hydroxylamine. Under a complex cell lysate mixture, both histone peptides with Kme1 were converted to 2H-indazole-3-carbaldehyde products and enriched via oxime chemistry (Figure 61, Supplementary Figure. 11).

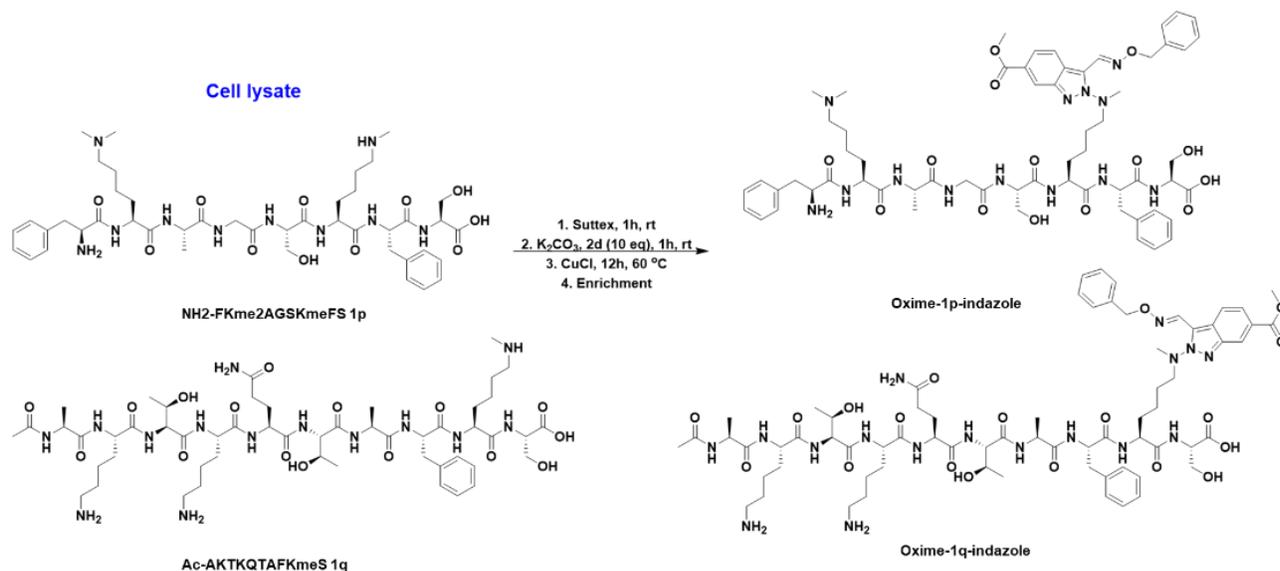
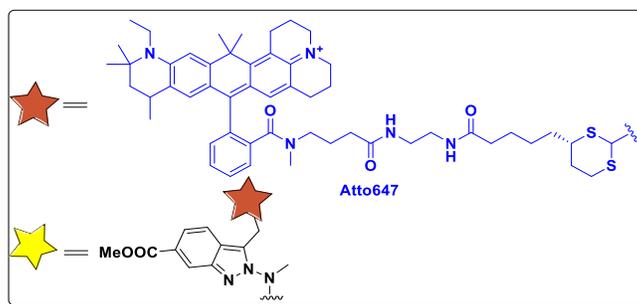


Figure 61. Selective enrichment of Kme1 containing peptides in a complex cell lysate mixture by TCC. The triazene cyclization reaction is able to selectively tag the Kme containing fragment from the complex mixture.

2.2.8 Single-molecule sequencing for identification of Kme1 sites by TCC

We demonstrated the identification of Kme1 sites at the level of a single molecule. We modified Kme1 in a model peptide using the TCC reaction (Figure. 62, Supplementary Figure. 12). The aldehyde group of peptide-2H-indazole-3-carbaldehyde was then functionalized with Atto647N fluorophore using dithiolane chemistry under acidic conditions, followed by HPLC purification and LCMS analysis. The fluorophore-labeled peptide NH₂-AKme1Atto647NGSKAF(PRA)A-CONH₂ was immobilized on an azide-functionalized microscope slide using PRA on a peptide by click chemistry as part of a fluorosequencing workflow. Next, the fluorophore-labeled immobilized peptide was subjected to multiple rounds of Edman's degradation, including two

rounds of mock Edman's degradation (M1-M2) with all the reagents except phenylisothiocyanate, and then analyzed using a total internal reflection fluorescence (TIRF) microscope (Figure. 62). The second position on a peptide was determined to be the site of Kme1 by observing a significant decrease in fluorescence following the second round of Edman's cycle using a TIRF microscope at the single-molecule level. There are no other chemical methods for the identification of Kme1 by any single molecule protein sequencing SMPS techniques.



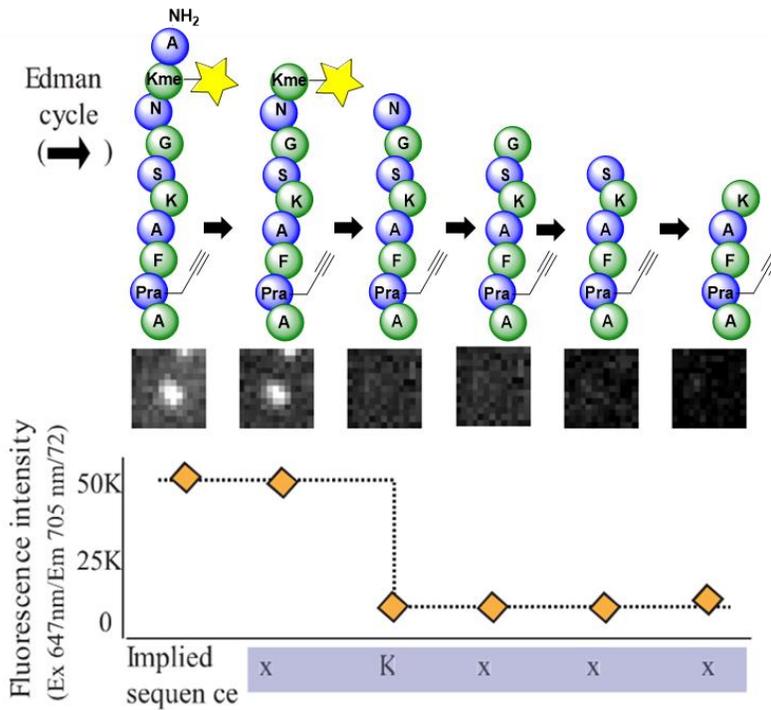


Figure 62. SMPS techniques for identification of Kme1 sites on modified peptide. The Kme residue was selectively modified by TCC reaction, then the N-terminal amino acid was removed by Edman degradation, each molecule's fluorescence intensity was monitoring by fluorescence microscopy to precisely identify the Kme position in the peptide.

Reference:

1. Bedford, M. T.; Richard, S., Arginine Methylation: An Emerging Regulator of Protein Function. *Mol. Cell* **2005**, *18* (3), 263-272.
2. Lu, R.; Zhang, H.; Jiang, Y.-N.; Wang, Z.-Q.; Sun, L.; Zhou, Z.-W., Post-Translational Modification of MRE11: Its Implication in DDR and Diseases. *Genes* **2021**, *12* (8), 1158.
3. Egorova, K. S.; Olenkina, O. M.; Olenina, L. V., Lysine methylation of nonhistone proteins is a way to regulate their stability and function. *Biochemistry (Moscow)* **2010**, *75* (5), 535-548.
4. Yao, X.; Shen, W., Crucial function of histone lysine methylation in plant reproduction. *Sci. Bull.* **2011**, *56* (33), 3493-3499.
5. Blackwell, E.; Ceman, S., Arginine methylation of RNA-binding proteins regulates cell function and differentiation. *Mol. Reprod. Dev.* **2012**, *79* (3), 163-175.
6. Polevoda, B.; Sherman, F., Methylation of proteins involved in translation. *Mol. Microbiol.* **2007**, *65* (3), 590-606.
7. Moore, Kaitlyn E.; Carlson, Scott M.; Camp, Nathan D.; Cheung, P.; James, Richard G.; Chua, Katrin F.; Wolf-Yadlin, A.; Gozani, O., A General Molecular Affinity Strategy for Global Detection and Proteomic Analysis of Lysine Methylation. *Mol. Cell* **2013**, *50* (3), 444-456.
8. Carlson, S. M.; Moore, K. E.; Green, E. M.; Martín, G. M.; Gozani, O., Proteome-wide enrichment of proteins modified by lysine methylation. *Nat. Protoc.* **2014**, *9* (1), 37-50.
9. Kudithipudi, S.; Jeltsch, A., Approaches and Guidelines for the Identification of Novel Substrates of Protein Lysine Methyltransferases. *Cell Chem. Biol.* **2016**, *23* (9), 1049-1055.

10. Nishikori, S.; Hattori, T.; Fuchs, S. M.; Yasui, N.; Wojcik, J.; Koide, A.; Strahl, B. D.; Koide, S., Broad Ranges of Affinity and Specificity of Anti-Histone Antibodies Revealed by a Quantitative Peptide Immunoprecipitation Assay. *J. Mol. Biol.* **2012**, *424* (5), 391-399.
11. Busby, M.; Xue, C.; Li, C.; Farjoun, Y.; Gienger, E.; Yofe, I.; Gladden, A.; Epstein, C. B.; Cornett, E. M.; Rothbart, S. B.; Nusbaum, C.; Goren, A., Systematic comparison of monoclonal versus polyclonal antibodies for mapping histone modifications by ChIP-seq. *Epigenetics & Chromatin* **2016**, *9* (1), 49.
12. Islam, K., The Bump-and-Hole Tactic: Expanding the Scope of Chemical Genetics. *Cell Chem. Biol.* **2018**, *25* (10), 1171-1184.
13. Wang, R.; Islam, K.; Liu, Y.; Zheng, W.; Tang, H.; Lailier, N.; Blum, G.; Deng, H.; Luo, M., Profiling Genome-Wide Chromatin Methylation with Engineered Posttranslation Apparatus within Living Cells. *J. Am. Chem. Soc.* **2013**, *135* (3), 1048-1056.
14. Wang, R.; Luo, M., A journey toward Bioorthogonal Profiling of Protein Methylation inside living cells. *Curr. Opin. Chem. Biol.* **2013**, *17* (5), 729-737.
15. Peters, W.; Willnow, S.; Duisken, M.; Kleine, H.; Macherey, T.; Duncan, K. E.; Litchfield, D. W.; Lüscher, B.; Weinhold, E., Enzymatic Site-Specific Functionalization of Protein Methyltransferase Substrates with Alkynes for Click Labeling. *Angew. Chem., Int. Edit.* **2010**, *49* (30), 5170-5173.
16. Bothwell, I. R.; Islam, K.; Chen, Y.; Zheng, W.; Blum, G.; Deng, H.; Luo, M., Se-Adenosyl-l-selenomethionine Cofactor Analogue as a Reporter of Protein Methylation. *J. Am. Chem. Soc.* **2012**, *134* (36), 14905-14912.
17. Wang, Q.; Wang, K.; Ye, M., Strategies for large-scale analysis of non-histone protein methylation by LC-MS/MS. *Analyst* **2017**, *142* (19), 3536-3548.

18. Young, N. L.; DiMaggio, P. A.; Plazas-Mayorca, M. D.; Baliban, R. C.; Floudas, C. A.; Garcia, B. A., High Throughput Characterization of Combinatorial Histone Codes. *Mol. Cell. Proteom.* **2009**, *8* (10), 2266-2284.
19. Emenike, B.; Nwajiobi, O.; Raj, M., Covalent Chemical Tools for Profiling Post-Translational Modifications. *Front. Chem.* **2022**, *10*.
20. Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G., Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* **2007**, *4* (10), 798-806.
21. Nwajiobi, O.; Mahesh, S.; Streeby, X.; Raj, M., Selective Triazene Reaction (STaR) of Secondary Amines for Tagging Monomethyl Lysine Post-Translational Modifications. *Angew. Chem., Int. Edit.* **2021**, *60* (13), 7344-7352.
22. Kimball, D. B.; Herges, R.; Haley, M. M., Two Unusual, Competitive Mechanisms for (2-Ethynylphenyl)triazene Cyclization: Pseudocoarctate versus Pericyclic Reactivity. *J. Am. Chem. Soc.* **2002**, *124* (8), 1572-1573.
23. Shang, X.; Zhao, S.; Chen, W.; Chen, C.; Qiu, H., Copper-Catalyzed Cascade Cyclization Reaction of 2-Haloaryltriazenes and Sodium Azide: Selective Synthesis of 2 H-Benzotriazoles in Water. *Eur. J. Chem.* **2014**, *20* (7), 1825-1828.
24. He, M.; Chen, N.; Zhou, T.; Li, Q.; Li, H.; Lang, M.; Wang, J.; Peng, S., Copper-Catalyzed Tandem Cross-Coupling/[2 + 2] Cycloaddition of 1,6-Allenynes with Diazo Compounds to 3-Azabicyclo[5.2.0] Ring Systems. *Org. Lett.* **2019**, *21* (23), 9559-9563.
25. Wang, Y.; Zhou, Y.; Ma, X.; Song, Q., Solvent-Dependent Cyclization of 2-Alkynylanilines and ClCF₂COONa for the Divergent Assembly of N-(Quinolin-2-yl)amides and Quinolin-2(1H)-ones. *Org. Lett.* **2021**, *23* (15), 5599-5604.

26. Lasányi, D.; Tolnai, G. L., Copper-Catalyzed Ring Opening of [1.1.1]Propellane with Alkynes: Synthesis of Exocyclic Allenic Cyclobutanes. *Org. Lett.* **2019**, *21* (24), 10057-10062.
27. Jacob, N.; Guillemard, L.; Wencel-Delord, J., Highly Efficient Synthesis of Hindered 3-Azaindoles via Metal-Free C–H Functionalization of Indoles. *Synthesis* **2020**, *52* (04), 574-580.

3.1 Design of Chemical Probes for Kme-Directed Modification (Kme-DM)

In this project, we report a method that enables site-selective modification of a high-frequency Lys residue in the mono-methyl lysine containing (Kme) protein for studying, monitoring, and tracking the PPI between the K-me protein and its reader and eraser proteins. The probes we have designed have a diazonium salt warhead with mask group for selective labeling of mono-methyl lysine, flexible linker connects with various electrophiles for intramolecular modification and the light activated group for capturing the reader and eraser proteins (Figure 63). In our purpose, once the mask group is removed, the diazonium salt will trap the K-me residue selectively. The nearby nucleophilic amino acids will react with the electrophile to achieve the site selective intramolecular labeling of the K-me protein. After the modification, we can decouple the triazene-conjugate to regenerate the unmodified K-me residue. Once the K-me residue is recognized and bound with its reader or eraser, we will utilize the photo-reactive group to capture the enzymes by the covalent bond formation. The captured proteins will be analyzed by SDS-PAGE and LC-MS-MS (Figure. 63).

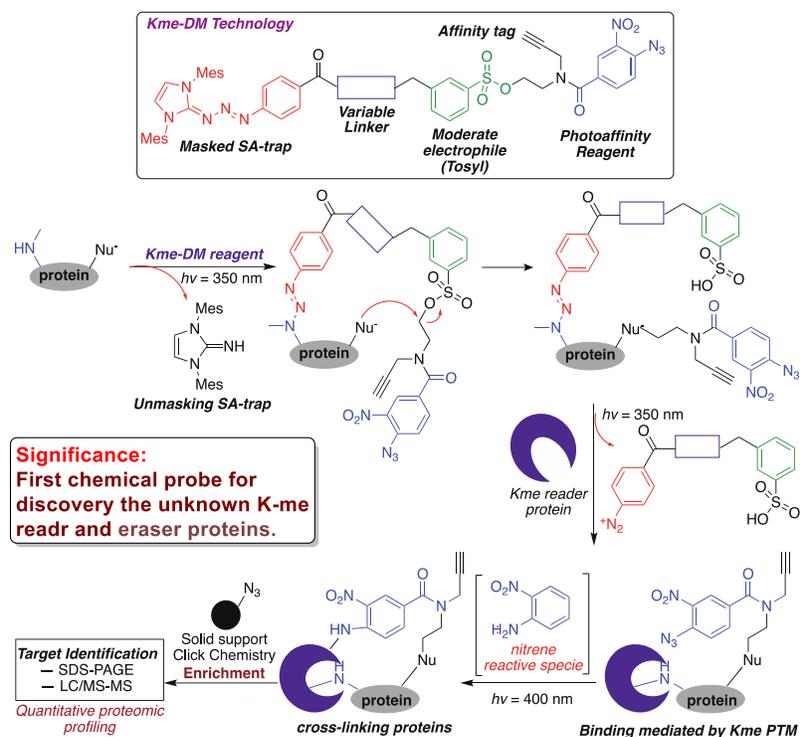


Figure 63. Site-selective K(me)-directed lysine modification for Kme-directed modification (Kme-DM). The Kme protein is selective modified by the Kme-directed intramolecular modification, then the reader and eraser will come to bind with Kme residue once the triazene will be selectively removed. The reader and eraser will be captured by the photo reactive group, the whole complex will be analyzed by SDS-page and LC-MS/MS.

3.2 Design and synthesis of various probes for Kme-DM

In order to optimize the intramolecular reaction, we made several different probes with various electrophiles and different affinity tag such as acyl imidazole (AI), N-acyl-N-alkyl sulfonamide (NASA), and N-sulfonyl pyridone (SP) (Figure 64, Supplementary Figure. 1).

Synthesis of AI-I:

Fragment **b3** was synthesized from 2-(2-aminoethoxy)ethan-1-ol **b1** by tosylation of free amine and activation of the alcohol side for the amide bond formation. Intermediate **b9** was synthesized by following Addy's procedure¹. Histamine was coupled with activated ester **b9** to afford the histamine intermediate **b11**. Next the pre-activated intermediate **b3** was treated with **b11** to obtain the final probe **AI-I**.

Synthesis of AI-III:

The linker **b11** was coupled with activated ester **b9** to obtain the precursor **b12**. The activated ester **b14** was synthesized by hydrolysis of ester intermediate **b12** following by activating of the carboxylic acid with NHS-ester group. Intermediate **b14** was coupled with histamine and **b3** to afford the final product **AI-III**.

Synthesis of SP-I and SP-II:

Intermediate **b17** was synthesized by following Matsuo's procedure². The reaction condition of coupling reaction between **b18** and **b19** was modified from Matsuo's procedure. The N-sulfonyl pyridine was synthesized by coupling **b20** with **b21** under basic condition. The **SP-II** was synthesized by coupling **b20** with Ts-Cl.

Synthesis of NASA-I and NASA-II:

The synthetic protocols for **NASA-I** and **NASA-II** were modified and optimized from Tamura's procedures³.

Synthesis of AI-Coumarin:

The intermediate **31** was synthesized by following Tamura's protocols³. The intermediate **b11** was coupled with intermediate **b31** under basic condition to obtain **AI-Coumarin**.

Synthesis of AI-Azide:

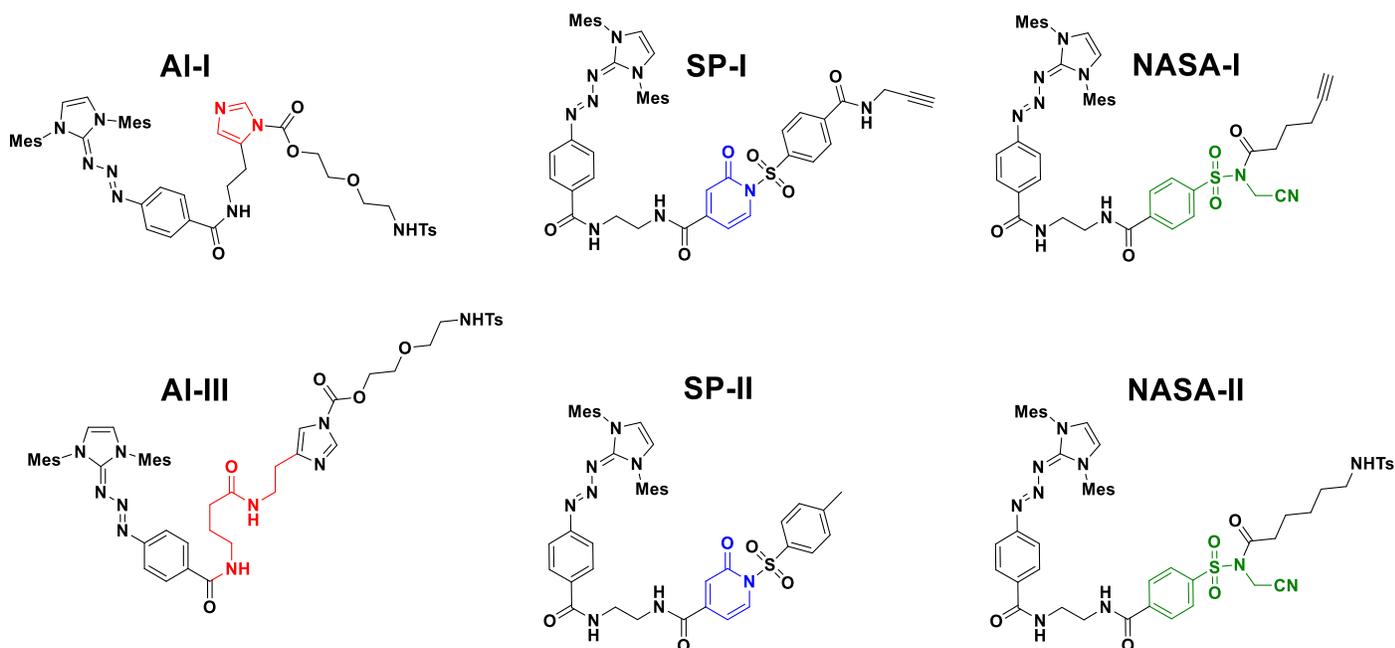
The azide intermediate **b34** was synthesized by the known protocol⁴. Intermediate **b34** was treated with DSC to obtain the activated intermediate **b35** for amide bond formation. The **AI-Azide** was generated by coupling between **b11** and **b35** with pyridine.

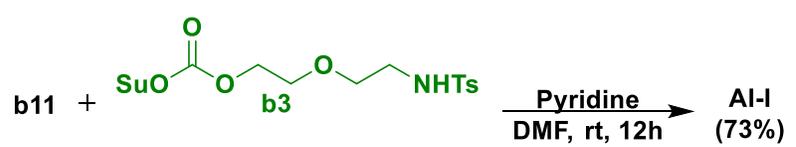
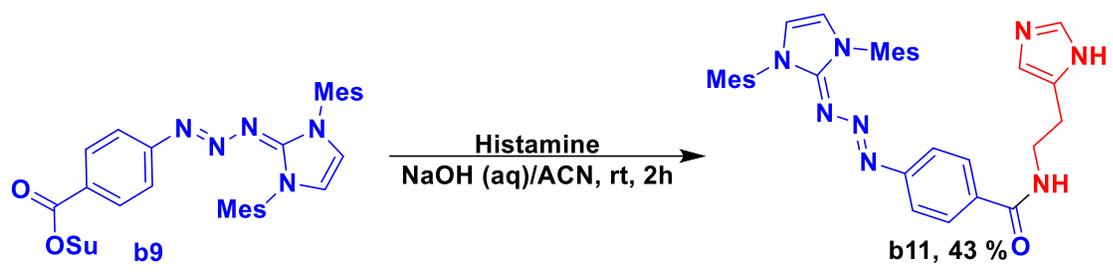
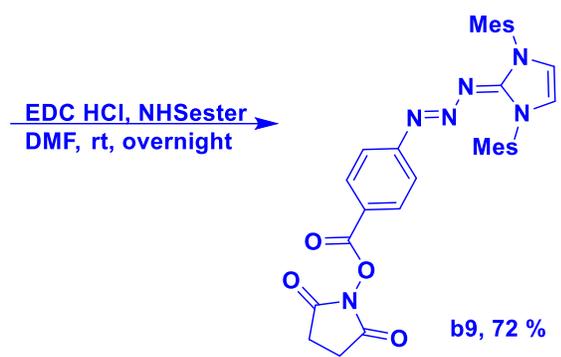
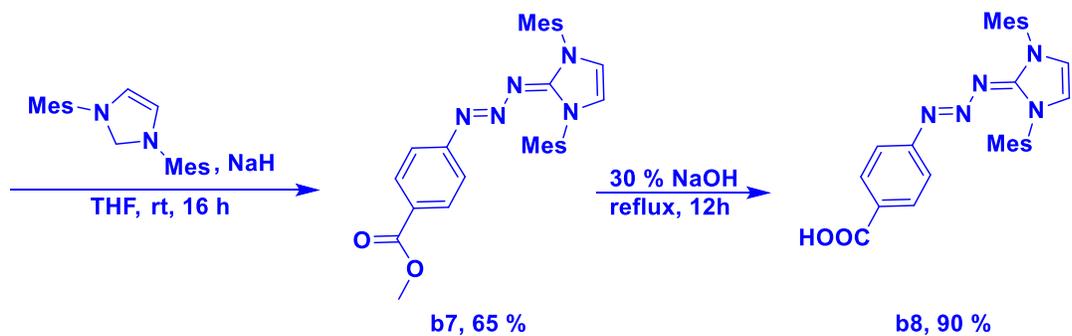
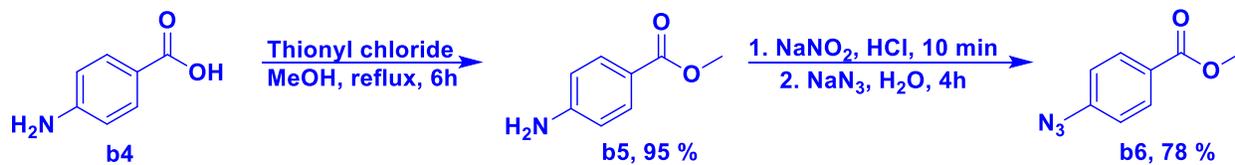
Synthesis of AI-Alkyne:

Intermediate **b39** was synthesized by Fujishima's protocol⁵. Intermediate **b11** was treated with **b39** under basic condition to afford **AI-Alkyne**.

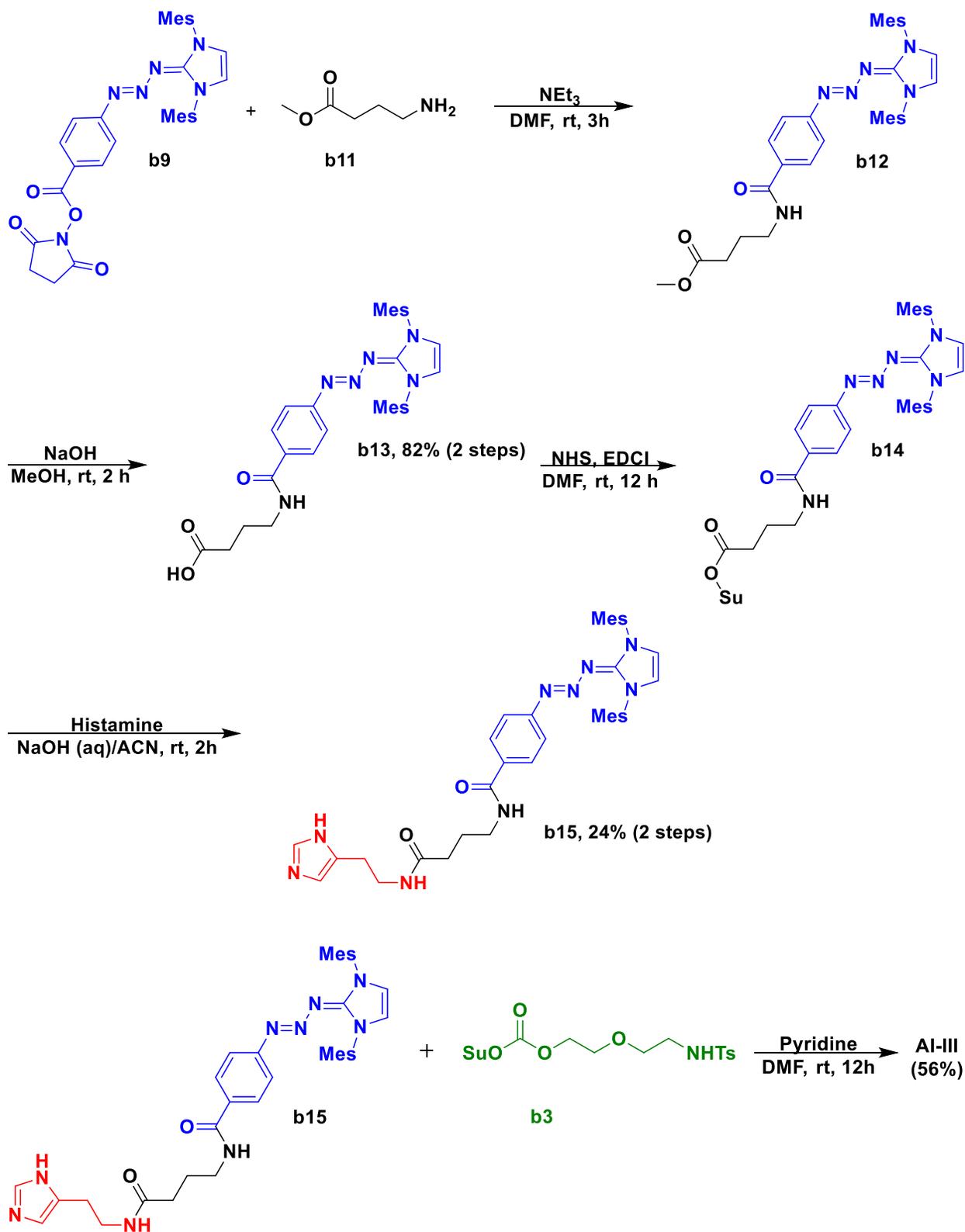
Synthesis of unmasked-AI-I:

Intermediate **b42** was synthesized by known protocol⁶. Next, the intermediate **b42** was coupled with histamine and intermediate **b3** to obtain intermediate **b44**. The Boc group was deprotected under the acid condition, then the aniline intermediate was treated with sodium nitrite and tetrafluoroboric acid solution to afford **unmasked-AI-I probe**.

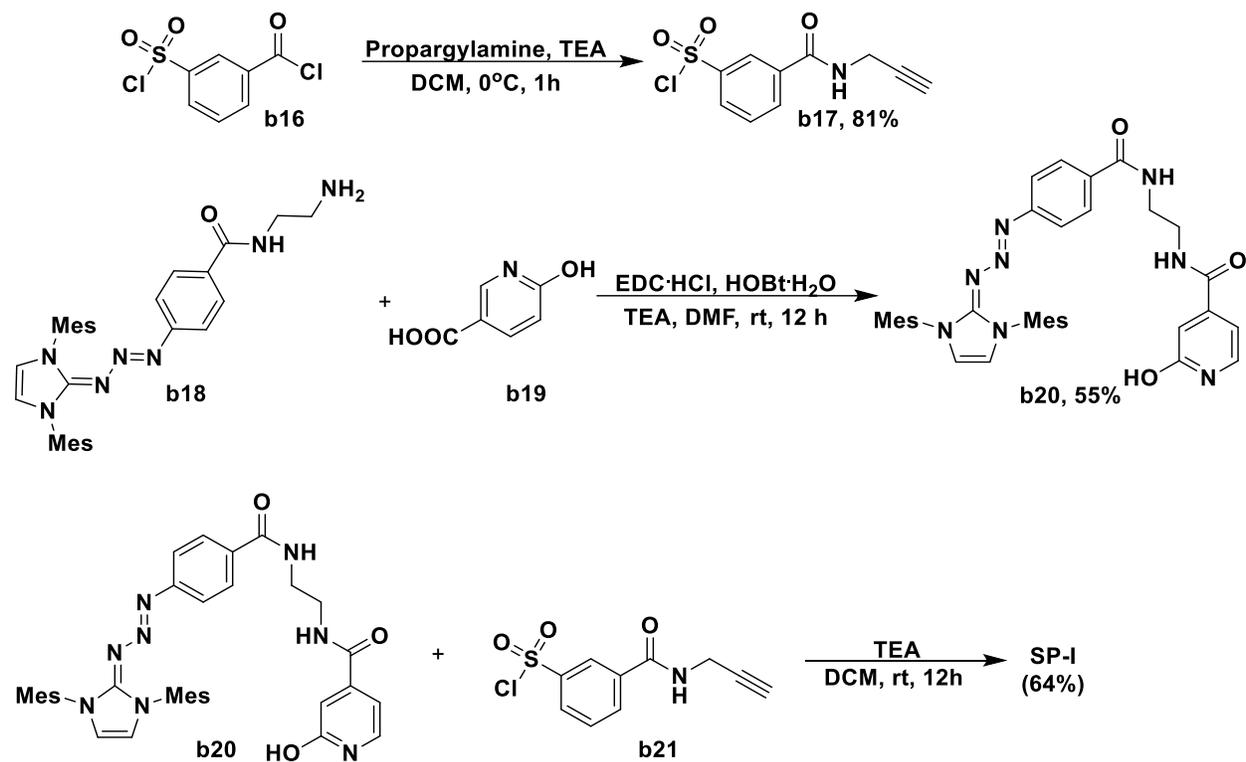




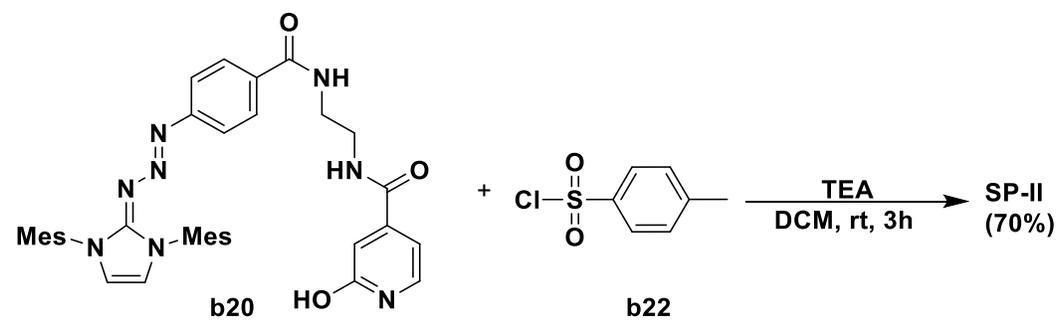
Synthesis of AI-III



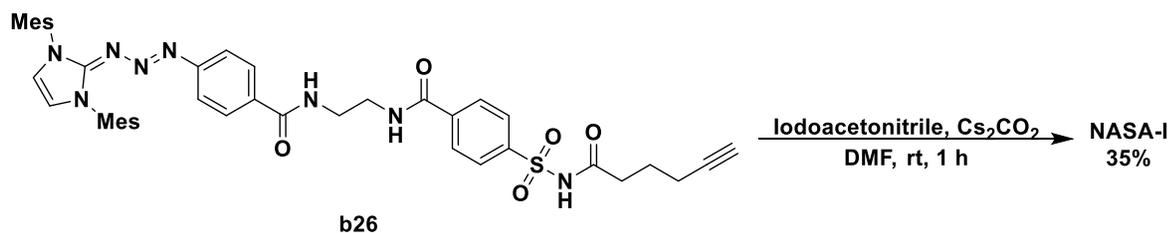
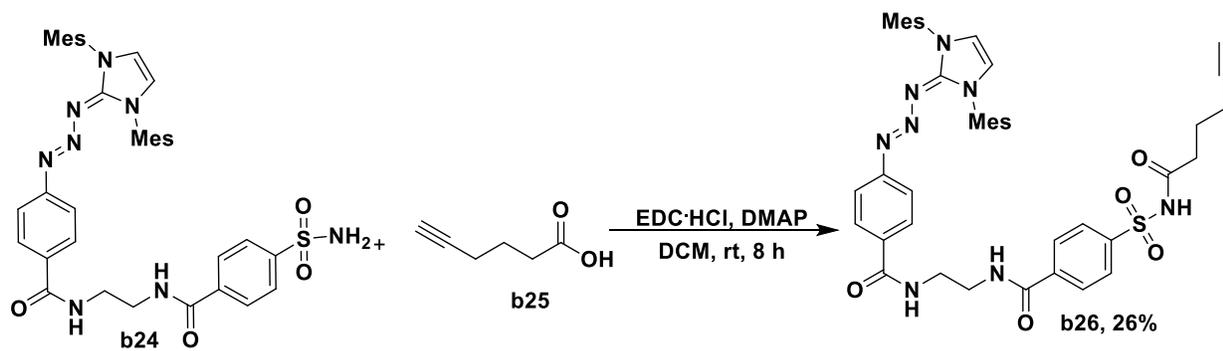
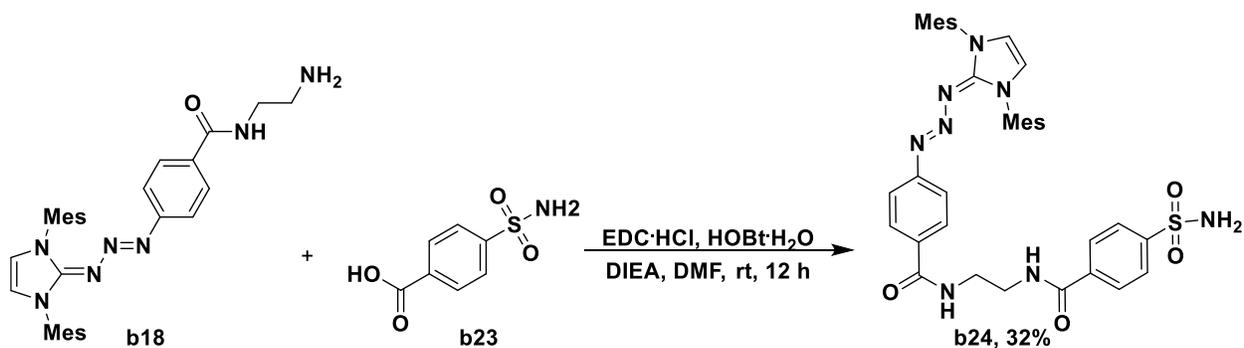
Synthesis of SP-I

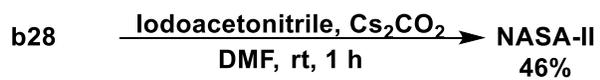
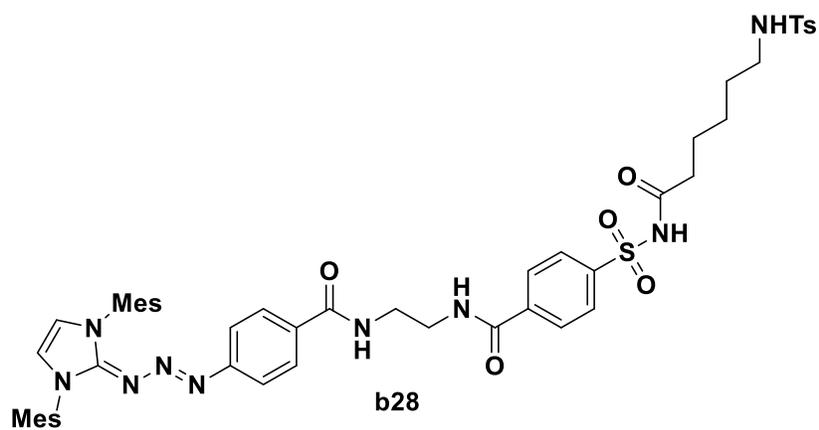
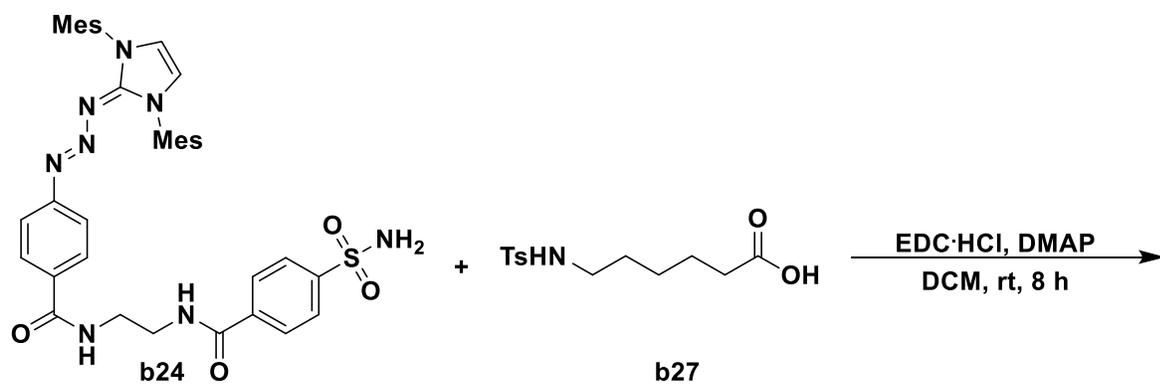


Synthesis of SP-II

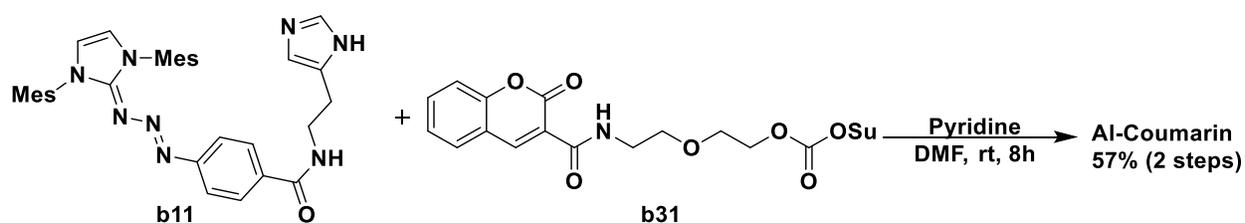
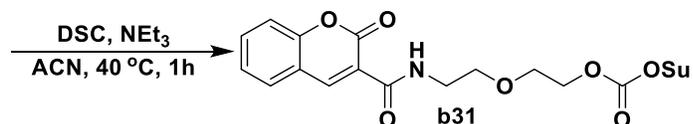
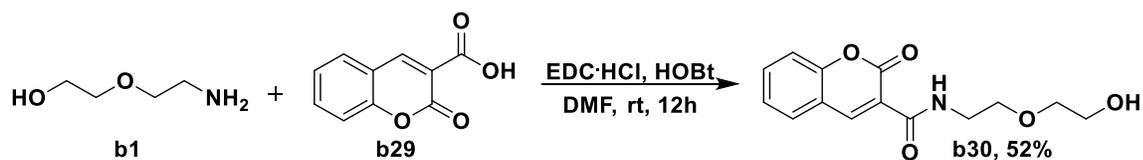


Synthesis of NASA-I

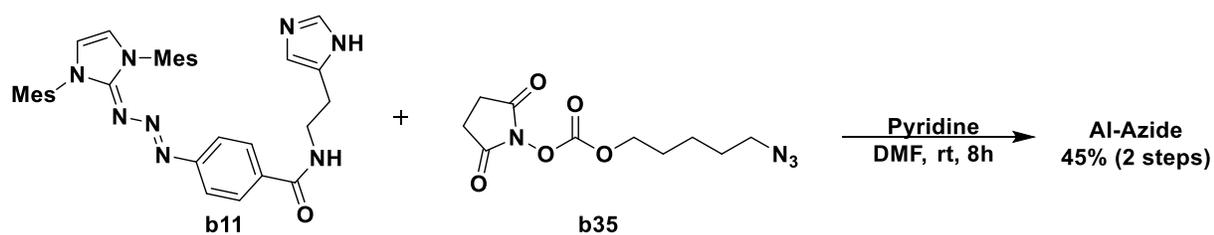
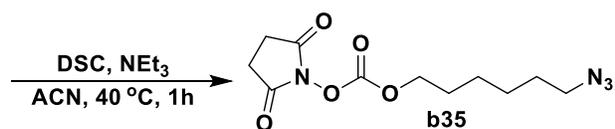
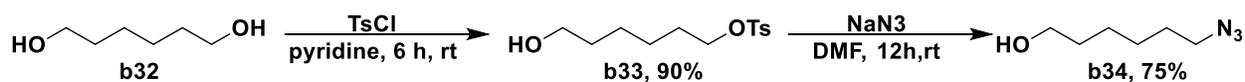




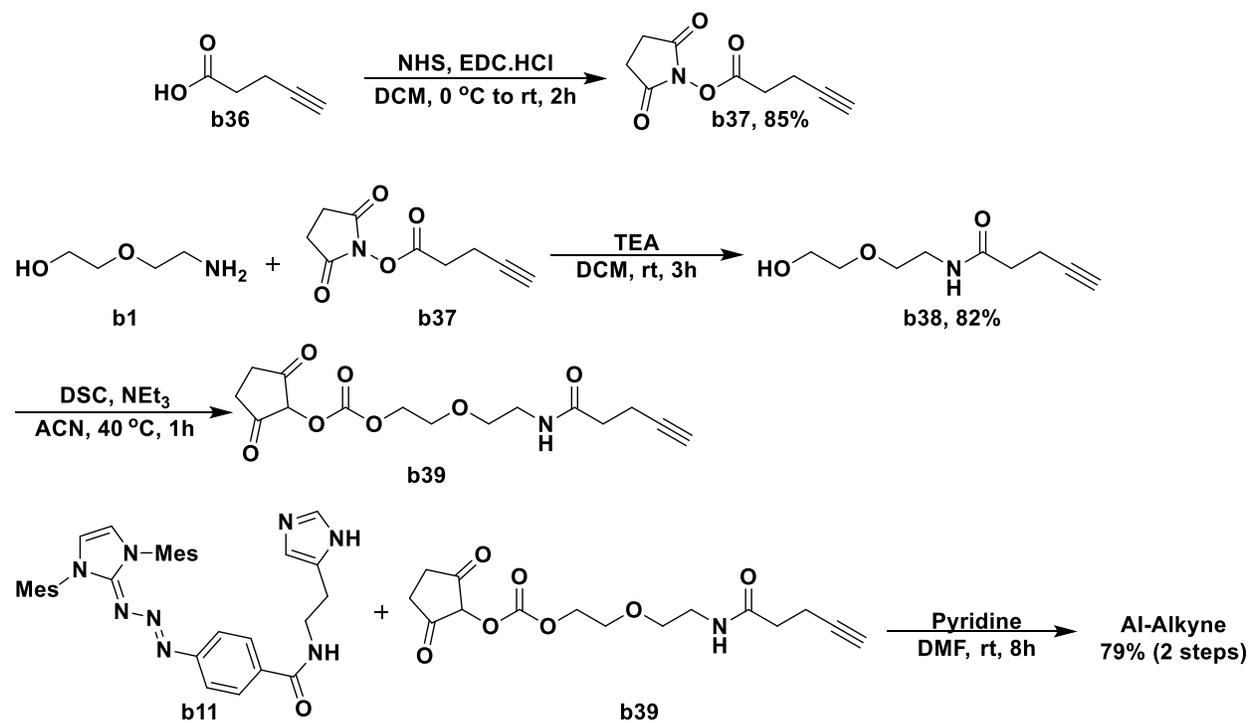
Synthesis of Al-Coumarin



Synthesis of Al-Azide



Synthesis of Al-Alkyne



Synthesis of unmasked-AI-I

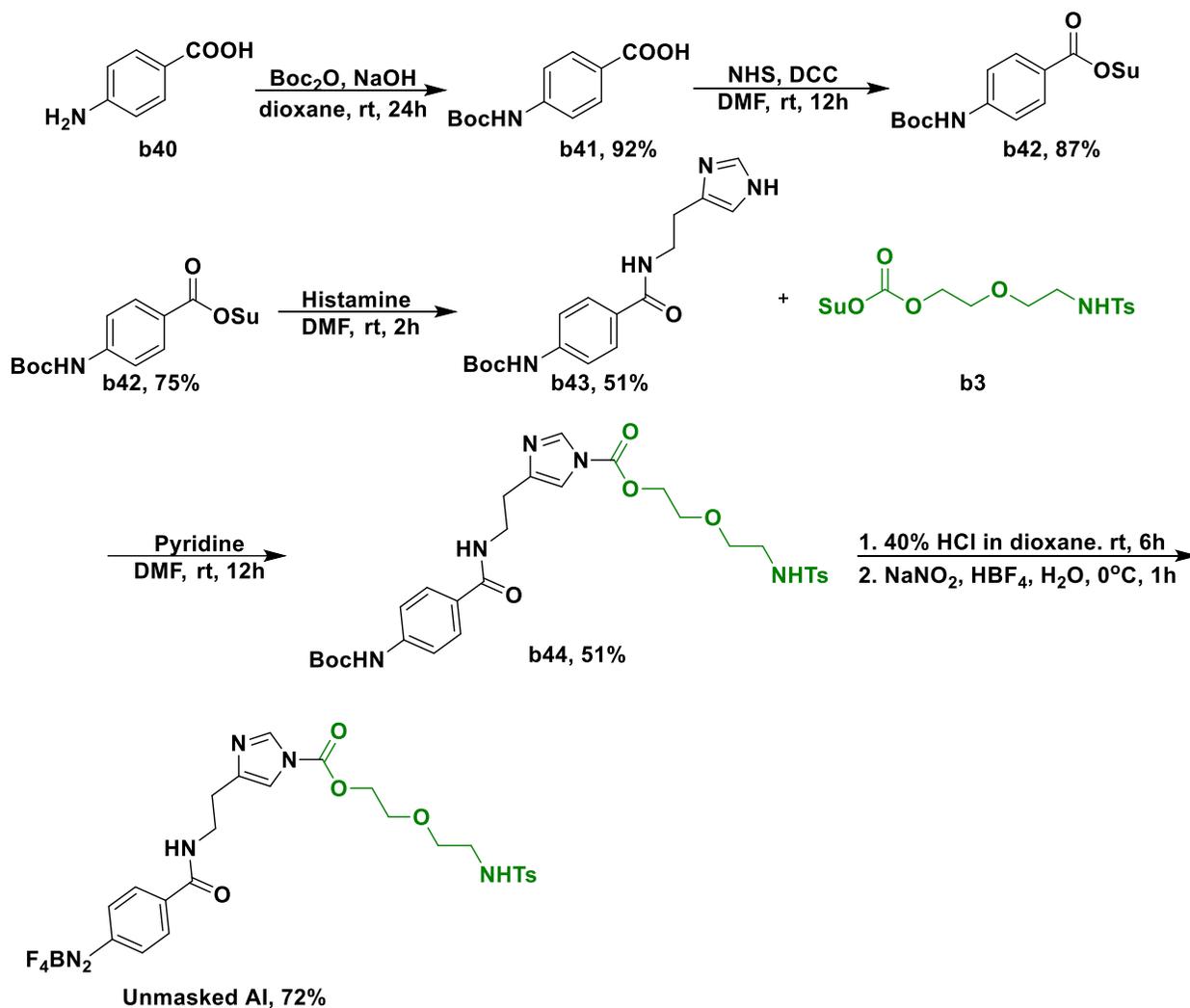


Figure 64. Synthesis of AI, SP, and NASA probes for investigation of Kme-DM reaction

3.3 Optimization of intramolecular reaction under UV lamp with different proline containing peptides and chemo selectivity study.

To optimize the intramolecular reaction with various types of probes, we incubated the probes with proline-containing peptides which have the lysine residue at position 2 or 3 or 4 (K-2, K-3, K-4) to determine the efficiency of intramolecular reaction with various position of lysine residue. After one-hour triaznation at room temperature, the reaction mixture was incubated at

40 °C for 6 h. After intramolecular reaction the reaction mixture was treated in 10 % TFA to remove the triazene group from proline in order to determine the conversion of intramolecular modified product (Figure 65).

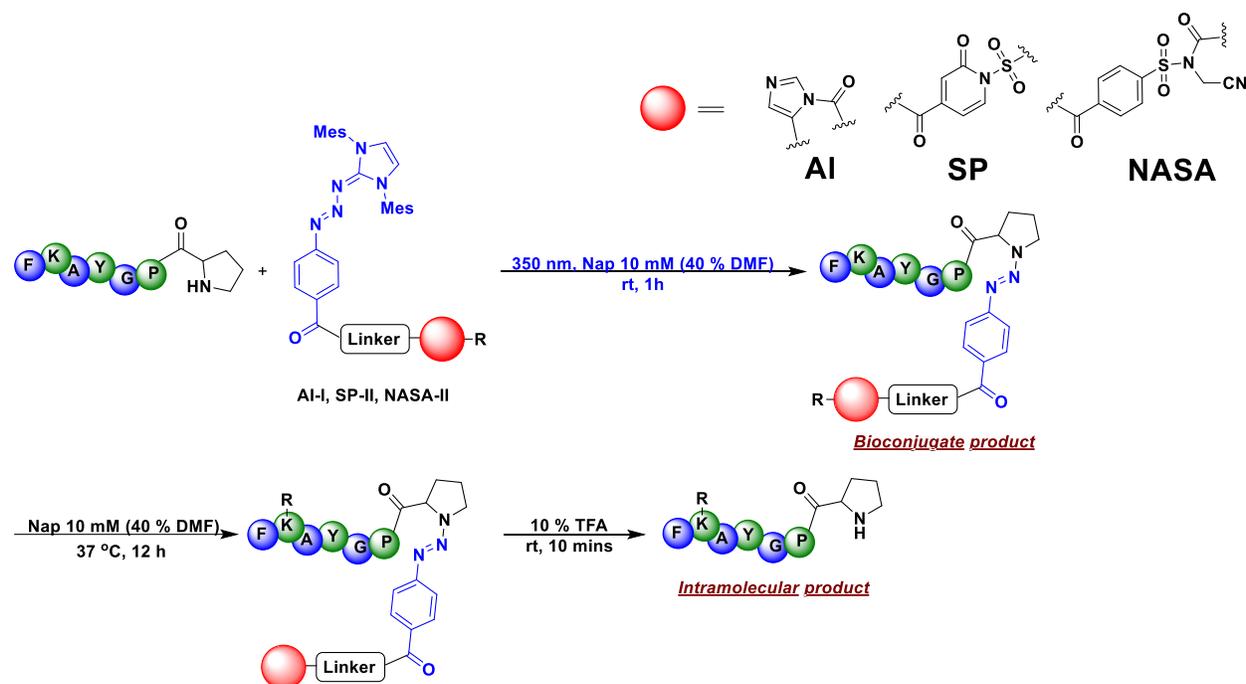


Figure 65. Site-selective proline-directed lysine modification with different probes. The Acyl imidazole (AI), N-acyl-N-alkyl sulfonamide (NASA), and N-sulfonyl pyridone (SP) probes were used to test the conversion of proline-directed intramolecular lysine modification.

We observed high conversion of triazene intermediate (100 %) with probe **AL-I** and **SP-II**. However, a significant amount of intermolecular (85 %) side product was observed in **NASA-II** example after 1h at room temperature. After another 6 h incubation, the **AL-I** resulting bioconjugates with **K-2** and **K-3** have been fully converted to the intramolecular product at 40 °C and 67 % (**K-2**) and 32 % (**K-3**) at 25 °C (Table 2, Supplementary Figure. 2) , but we only observed moderate intramolecular conversion with **k-4** peptide due to the short distance between

the lysine residue and acyl imidazole. The distance effect seems negligible in **SP-II** examples because we observed similar intramolecular reactivity (62 % - 72 %, Table 2, Supplementary Figure. 2) with different peptides. Moreover, to further confirm, the lysine was modified through intramolecular reaction, not intermolecular reaction. We incubated the **AI-I** probe with **K-2** peptide at 40 °C under dark conditions. After 12 hours, no intermolecular product was formed (Figure 66). This result shows that our labeling strategy with the **AI-I** probe is a site-selective reaction that only occurs via the intramolecular pathway, and the reaction cascade is only triggered by the UV lamp. After the intramolecular reactivity study, we decide to move forward with our **AI-I** probe due to its high intramolecular reactivity and preference for various peptides.

Table 2. Optimization investigation with AI-I, SP-II, NASA-II with K-2, K-3, k-4 peptides

Probe	Peptide	Bioconjugate product	Intramolecular product
AI-I	k-2	100 %	100 %, 67 % (25 °C)
AI-I	K-3	100 %	100 %, 32 % (25 °C)
AI-I	k-4	100 %	58 %
SP-II	k-2	100 %	62.5 %
SP-II	K-3	100 %	63 %
SP-II	k-4	100 %	72 %
NASA-II	k-2	13 % (85 % intermolecular reaction)	ND

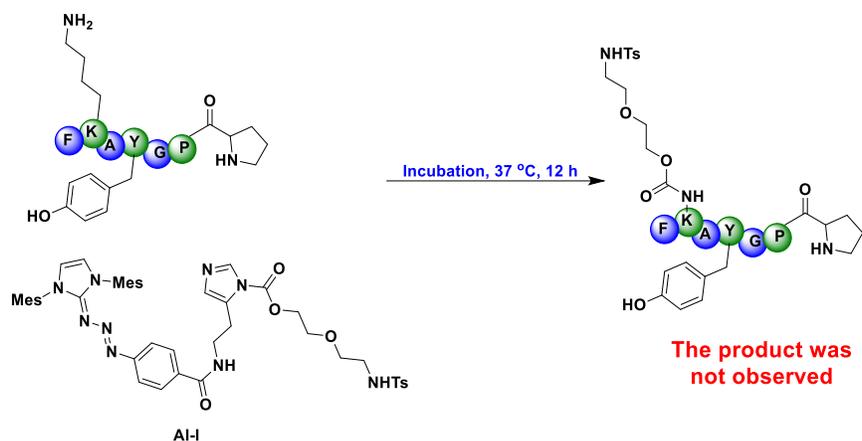


Figure 66. Intermolecular reactivity study of AI-I. The peptides which have proline at different position were used to test the compatibility of different probes.

3.4 Chemoselectivity and control study of AI-I probe

In order to test the chemoselectivity of the intramolecular reaction with acyl imidazole, we have made cysteine and histidine containing peptides (C2, H2) and FRQDW-NH-Ac for chemoselective investigation. The intramolecular cysteine product was not observed after the TFA cleavage; however, we observed the 50 % intramolecular conversion with the **H2** peptide. The bioconjugate product and intramolecular product were not observed with FRQDW-NH-Ac (Table 3, Supplementary Figure. 3).

Table 3. Chemo-selective study of AI-I probe

Probe	Peptide	Bioconjugate product	Intramolecular product
AI-I	C-2	100 %	0 %
AI-I	H-2	100 %	50 %
AI-I	FRQDW-NHAc	No bioconjugate product	No intramolecular product

3.5 Screening of AI-III probe with K-2, K-3 and K-4 peptides under UV lamp

For the next study, we want to know if the probes' linker length will affect the efficiency of intramolecular modification. So, we made **AI-III** analog which has five extra atoms between acyl imidazole and masked group compared to **AI-I**. **AI-III** probes only showed 72 % (48 %, rt)

and 42 % (18 %, rt) of intramolecular conversion with **K-2** and **K-3** at 40 °C, respectively. However, we observed the complete transformation from conjugate intermediate to the intermolecular product with **K-4**. This difference could be due to the flexibility of the longer linker (Table 4, Supplementary Figure. 4).

Table 4. Intramolecular reactivity investigation of AI-III probe

Probe	Peptide	Intramolecular product
AI-III	K-2	72% (40 °C), 48 (rt)
AI-III	K-3	42% (40 °C), 18 (rt)
AI-III	K-4	100% (40 °C), xx (rt)

3.6 Site-specific study with multiple lysins containing peptide

In order to study the site-selectivity of our **AI** probes, we incubated our **AI-I** and **AI-III** probes with a long peptide PTAPKSTGGKA to determine the site specificity of our intramolecular reaction with different **AI** probes. As a result, we observed completed conversion for the triazeneation with PTAPKSTGGKA in both **AI-I** and **AI-III**; however, after intramolecular reaction, we noticed a significant difference in reactivity between **AI-I** (50 %) and **AI-III**(96 %) due to the difference in length of linker (Figure 67, Supplementary Figure. 5).

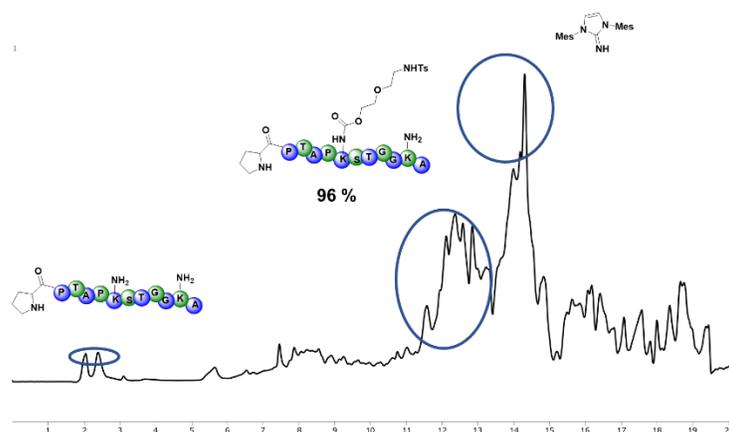


Figure 67. HPLC trace of intermolecular modification of lysine. The high intramolecular conversion of PTAPKSTGGKA most like due to the turn cause by proline.

In order to identify the site of modification, the modified peptide was treated with trypsin for the LC-MS/MS mapping. The LC-MS/MS of cleavage fragments indicated the labeling site is in the K-5 position (Figure 68, Supplementary Figure. 6).

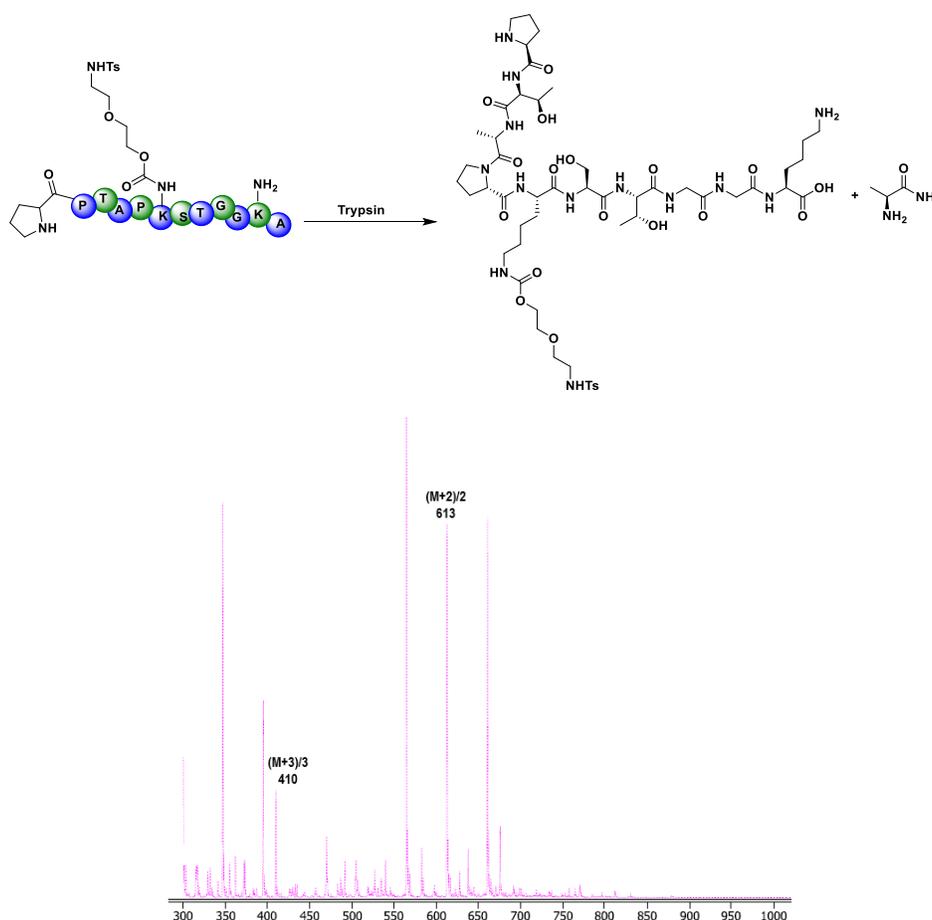


Figure 68. Trypsin digestion of modified peptide. The LC-MS fragment showed the K-5 is the labeling site.

3.7 Intramolecular reaction study of various AI analogs with different affinity tags

After the intramolecular reactivity, chemo, and site selectivity study, we functioned our **AI-I** probe with different affinity tags such as coumarin, alkyne, and azide in order to monitor and enrich the captured enzymes after proteomic profiling. Although we got the promising result

from the intramolecular study with **AI-I** and **AI-III** probes, we noted that the intense amount of intermolecular side product was forming with all the affinity tag analogs. We suspected this inconsistent result was caused by the poor solubility of those hydrophobic affinity tags and masked groups (Figure 69, Supplementary Figure. 7).

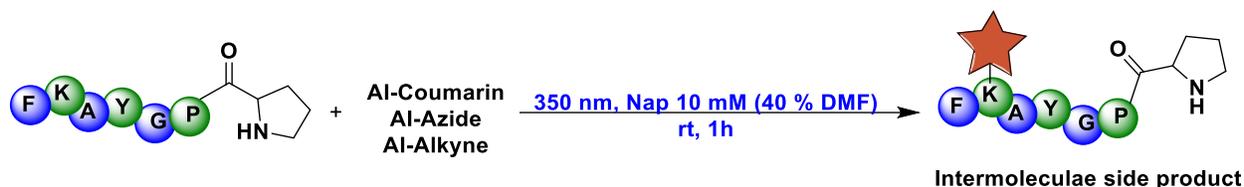


Figure 69. Intermolecular side reaction with various AI affinity tag analogs.

3.8 Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe

To solve the solubility issue with masked probe, we synthesized the more water-soluble unmasked probe for the further investigation. For the first study we treated our unmasked **AI** probe with NH₂-PGKAKF for the intramolecular reactivity and site selectivity study. After one hour, we observed the formation of both triazene intermediate (49 %) and intramolecular intermediate (46 %). The triazene intermediate was fully converted to the intramolecular intermediate after 6 hours incubation at 40 °C based on the HPLC (Figure 70, Supplementary Figure. 8). After the TFA cleavage we found the 76 % (71 %, 25 °C) conversion of intramolecular modification on the **K-3** position based on the LC-MS/MS and HPLC analysis (Supplementary Figure. 8).

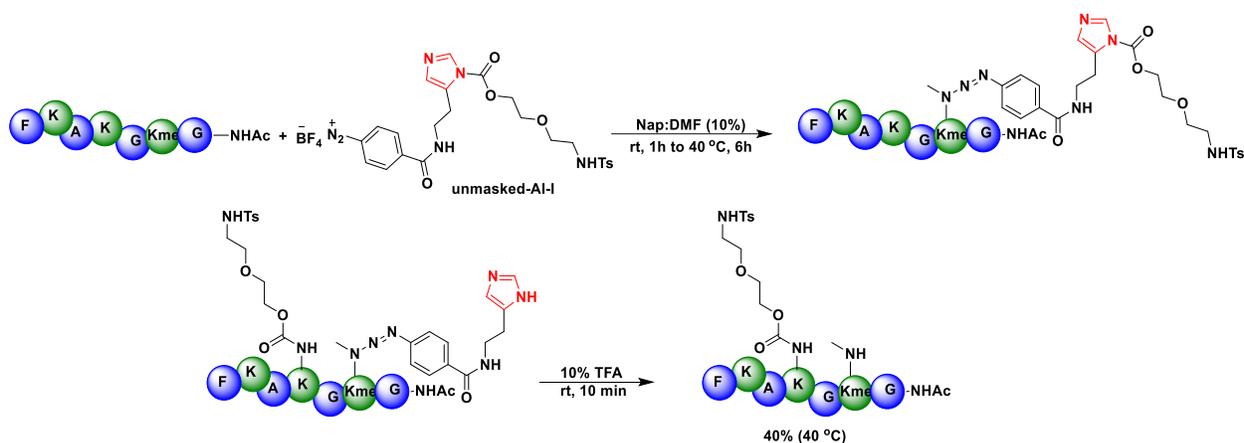


Figure 71. Site-selective Kme-directed lysine modification with unmasked AI probe with GKmeGKAKF.

The unmasked analog showed high chemo selectivity and favored intramolecular reaction compared to the parent compound. No triazene intermediate and intramolecular product was observed with $\text{NH}_2\text{-GCGKAKF}$; moreover, we only observed poor conversion of intermolecular side product with $\text{NH}_2\text{-GHGKAKF}$ (15 %, 1 h; 16 %, 6 h, Figure 72, Supplementary Figure.10), which can further confirm our unmasked probe is specific to the lysine residue, and the modification will only occur through intramolecular mechanism.

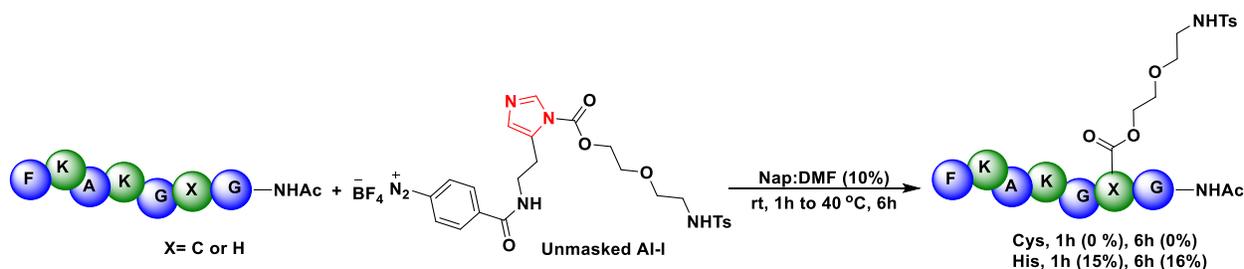


Figure 72. Chemo-selective study of unmasked AI probe. High lysine reactivity and selectivity was confirmed by the chemo-selective study with GCGKAKF and GHGKAKF.

3.9 Kme-Directed modification of truncated histone peptide:

To test our Kme-DM strategy in a more complicated system, we made the truncated version of the histone truncated peptide for the next investigation. We observed the more than 99 % conversion of triazene intermediate after 1 hour at 25 °C without forming the intermolecular side product (Figure 73). The LC-MS/MS indicated the K-2 lysine residue is the favored site for intramolecular modification (Supplementary Figure.11) .

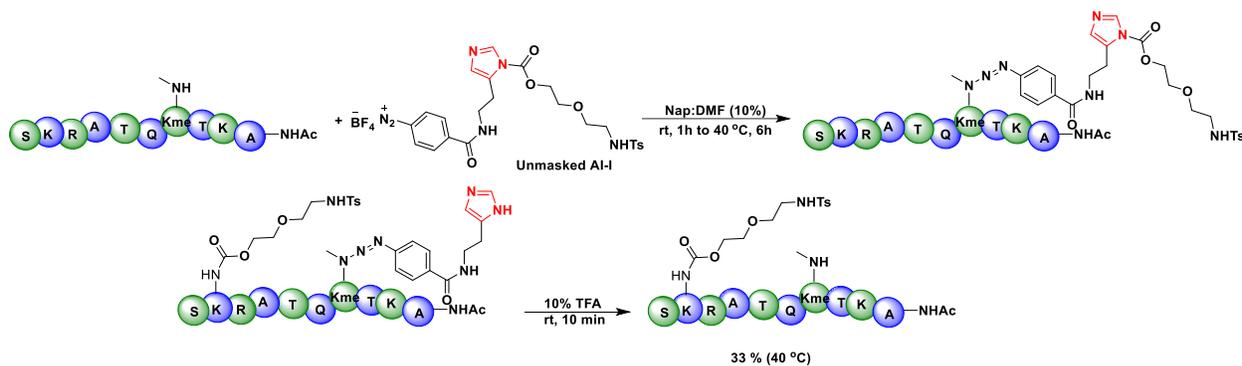


Figure 73. Kme-directed lysine modification of truncated histone peptide. The Kme-DM technique is compatible with the truncated histone peptide.

3.10. Future work

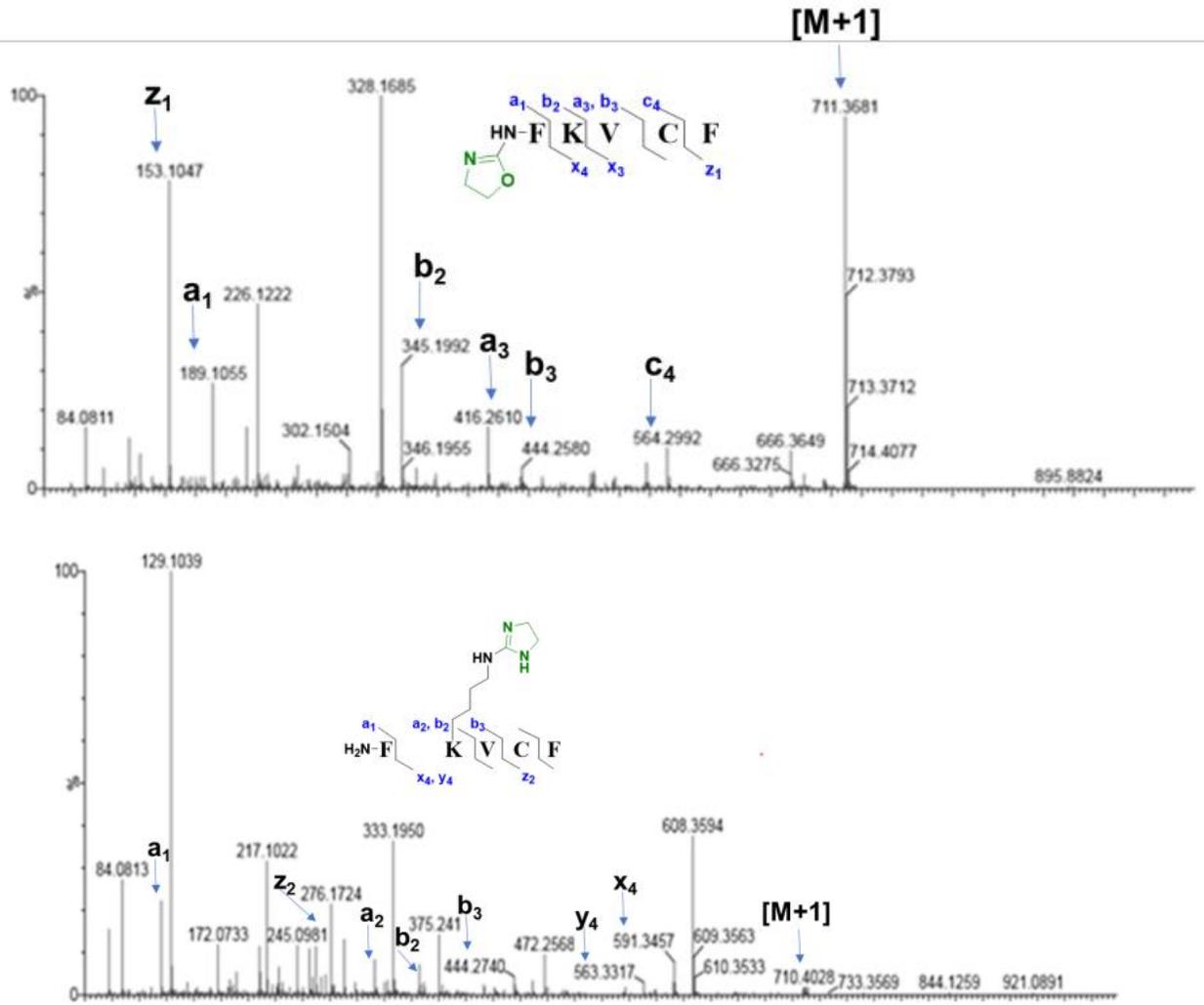
The protein capturing and analysis of the captured erasers and readers are undergoing.

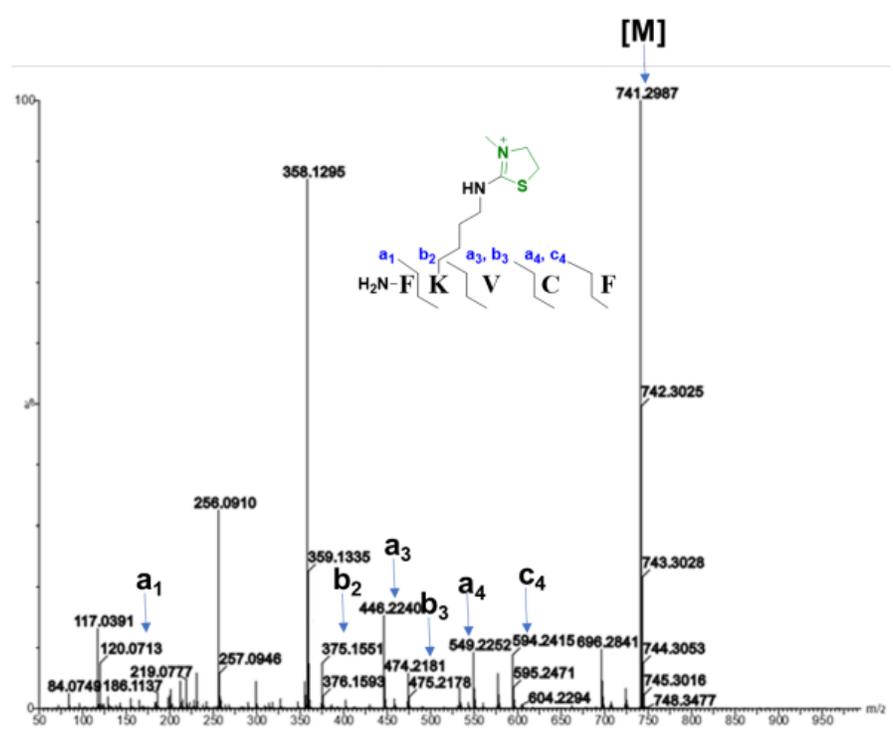
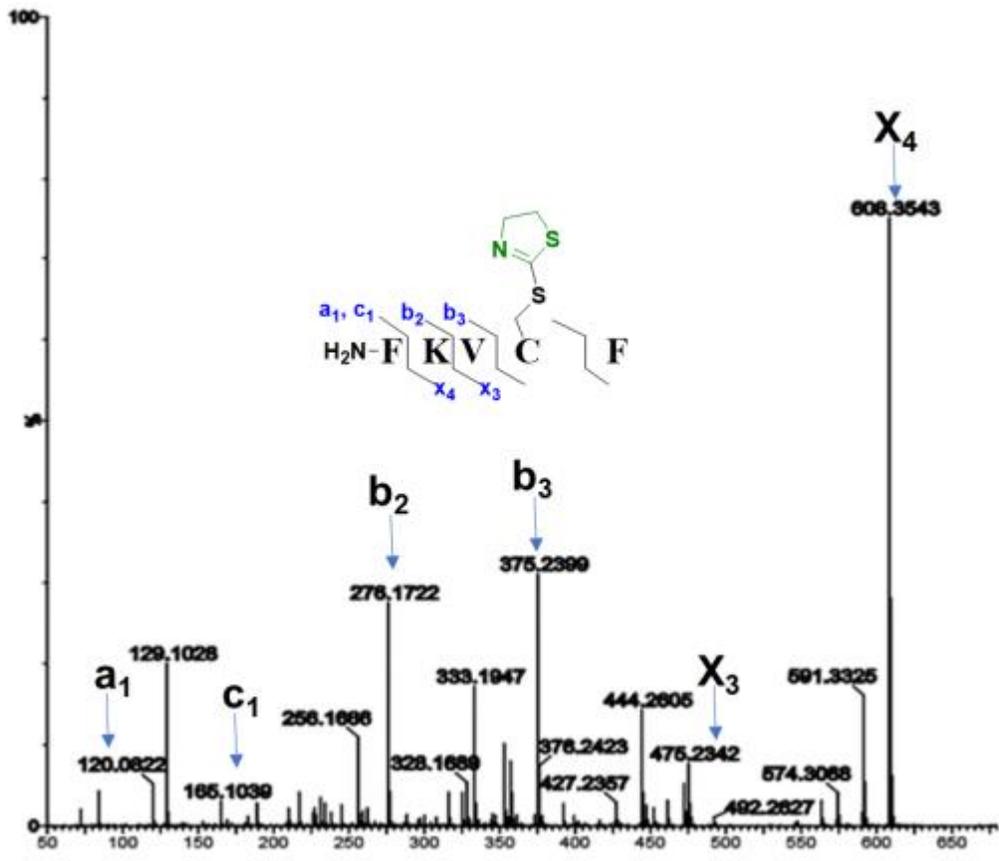
Reference:

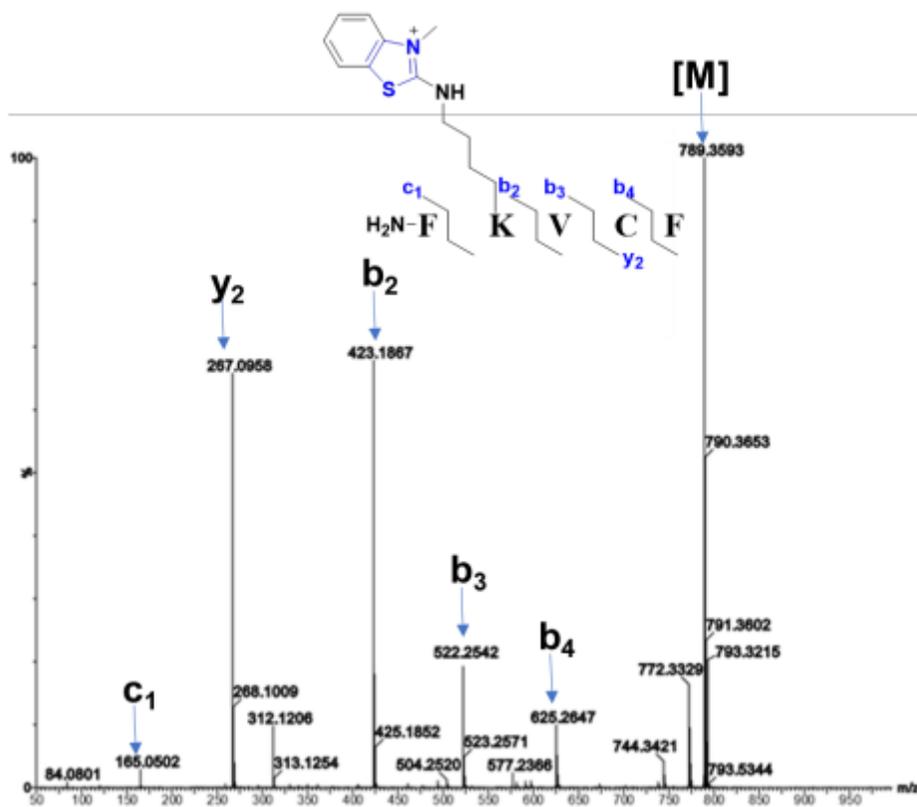
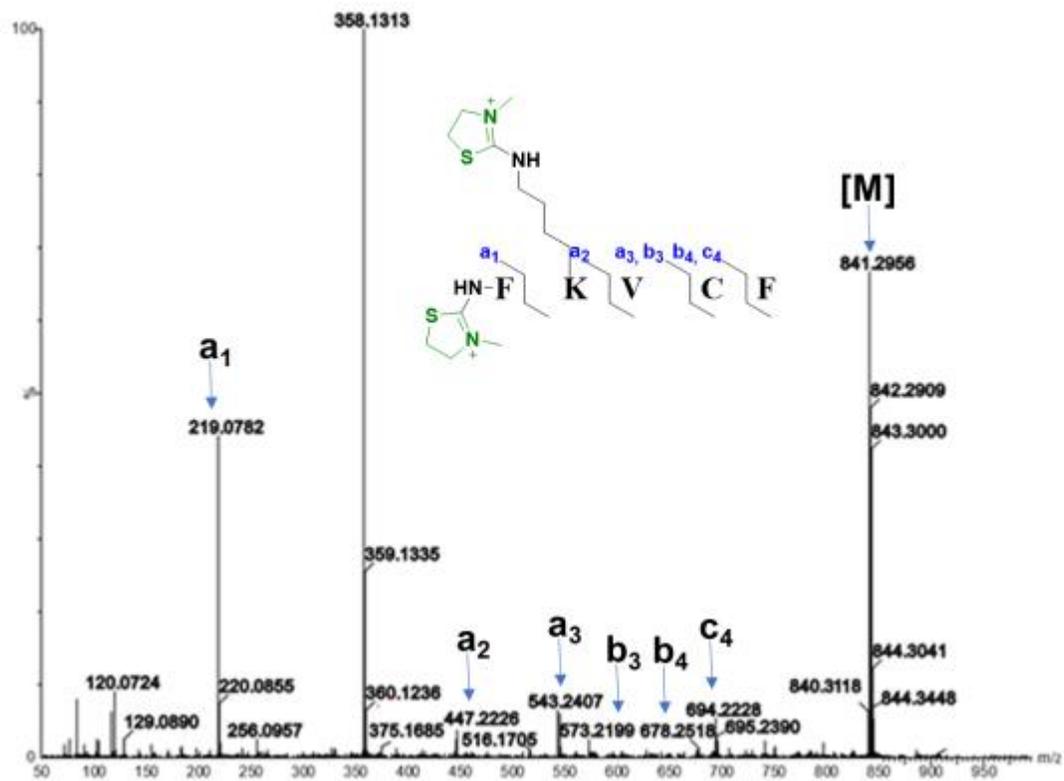
1. Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A., A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation. *J. Am. Chem. Soc.* **2017**, *139* (34), 11670-11673.
2. Matsuo, K.; Nishikawa, Y.; Masuda, M.; Hamachi, I., Live-Cell Protein Sulfonation Based on Proximity-driven N-Sulfonyl Pyridone Chemistry. *Angew. Chem. Int. Ed.* **2018**, *57* (3), 659-662.
3. Tamura, T.; Ueda, T.; Goto, T.; Tsukidate, T.; Shapira, Y.; Nishikawa, Y.; Fujisawa, A.; Hamachi, I., Rapid labelling and covalent inhibition of intracellular native proteins using ligand-directed N-acyl-N-alkyl sulfonamide. *Nat. Commun.* **2018**, *9* (1), 1870.
4. Yao, W.; Zhu, Y.; Zhang, X.; Sha, M.; Meng, X.; Li, Z., Semisynthesis of Chondroitin Sulfate E Tetrasaccharide from Hyaluronic Acid. *J. Org. Chem.* **2018**, *83* (22), 14069-14077.
5. Fujishima, S.-h.; Yasui, R.; Miki, T.; Ojida, A.; Hamachi, I., Ligand-Directed Acyl Imidazole Chemistry for Labeling of Membrane-Bound Proteins on Live Cells. *J. Am. Chem. Soc.* **2012**, *134* (9), 3961-3964.
6. Hofmann, K.; Finn, F. M.; Kiso, Y., Avidin-biotin affinity columns. General methods for attaching biotin to peptides and proteins. *J. Am. Chem. Soc.* **1978**, *100* (11), 3585-3590.

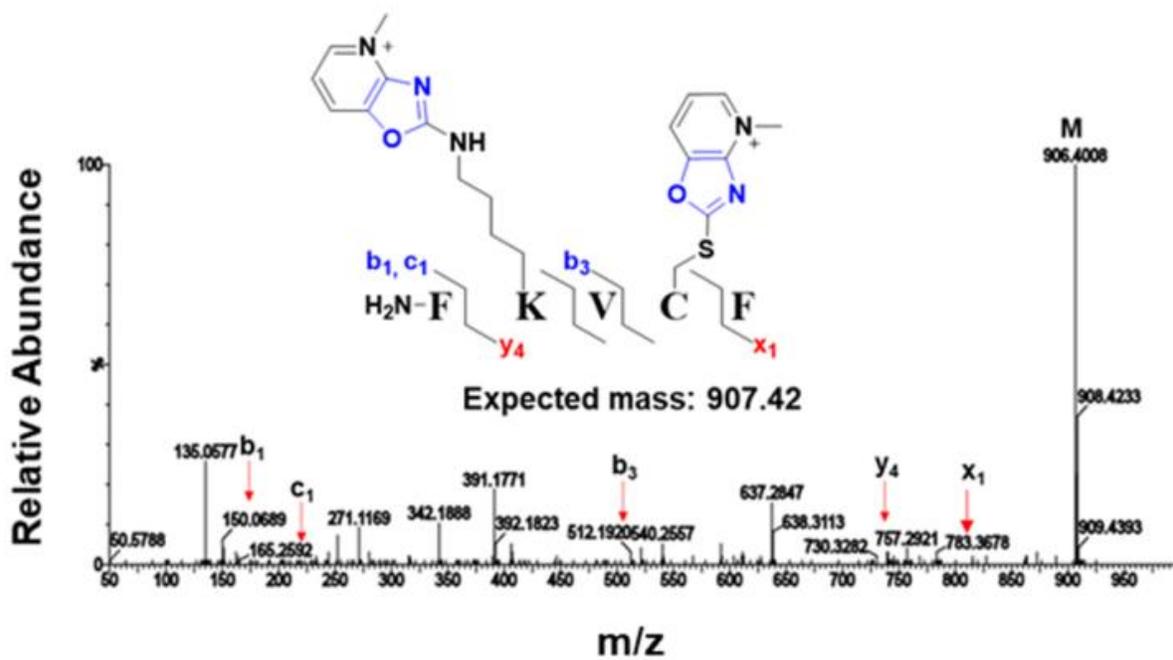
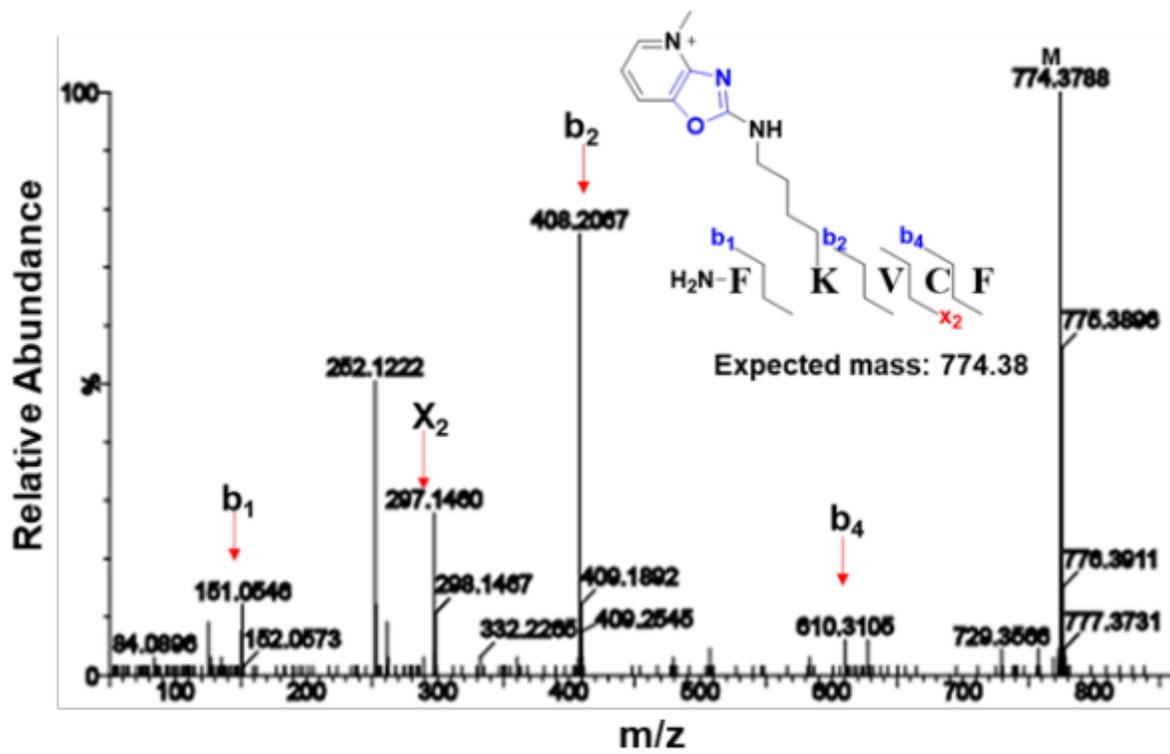
Supporting Information for Chapter One:

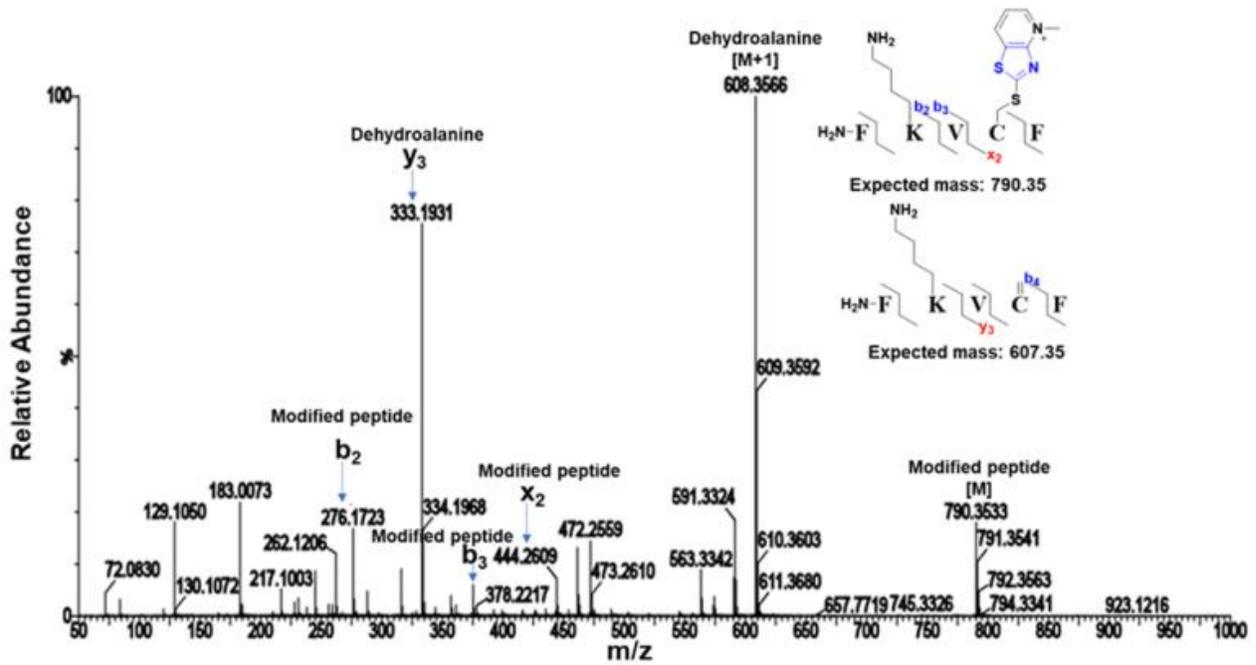
Supplementary Figure 1 : MS/MS of modified peptide







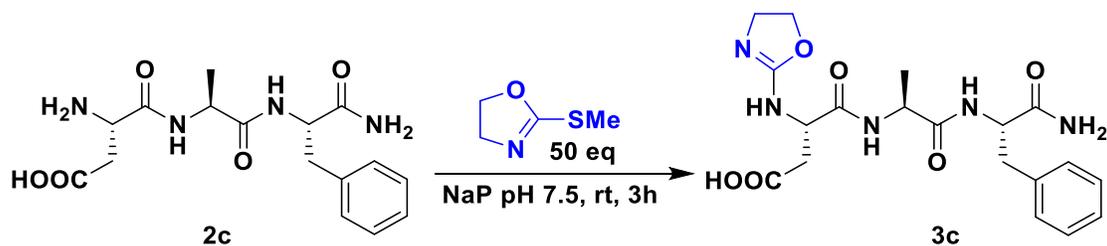




Supplementary Figure 2: General method for the verification of the chemo- and site-selective nature of oxazoline probe with peptides.

Procedure for synthesis of Ox1-XAF (3c-3j)

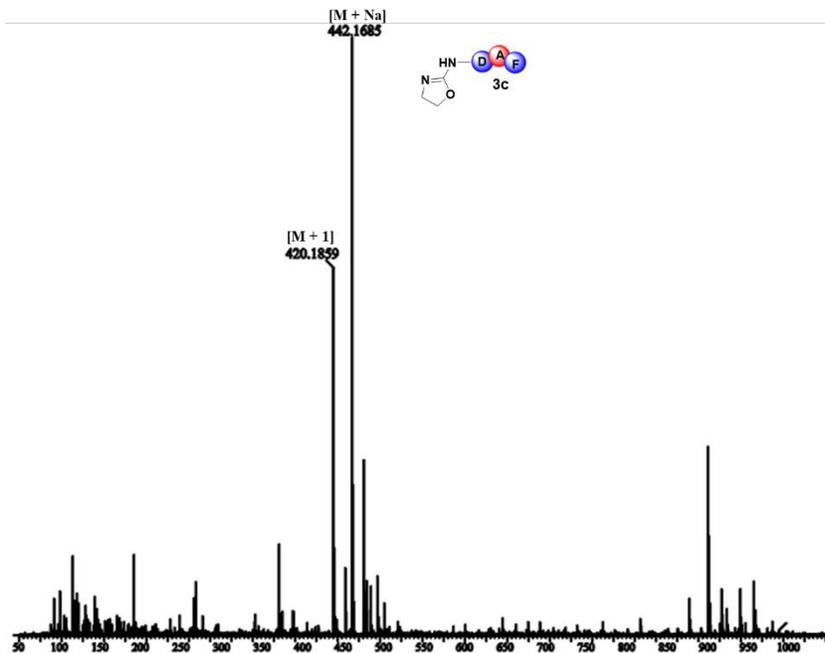
To peptides XAF (6.25 mM) in 0.4 mL of 10 mM phosphate buffer (Nap, pH 7.5) was added 2-methylthio oxazoline **1a** (50 equiv., 312.5 mM). The solution was stirred at room temperature for 3 h. The reaction was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



Ox1-DAF 3c. LCMS: m/z 420.1 (calcd $[M+H]^+ = 420.1$), m/z 442.1 (calcd $[M+Na]^+ = 442.1$),

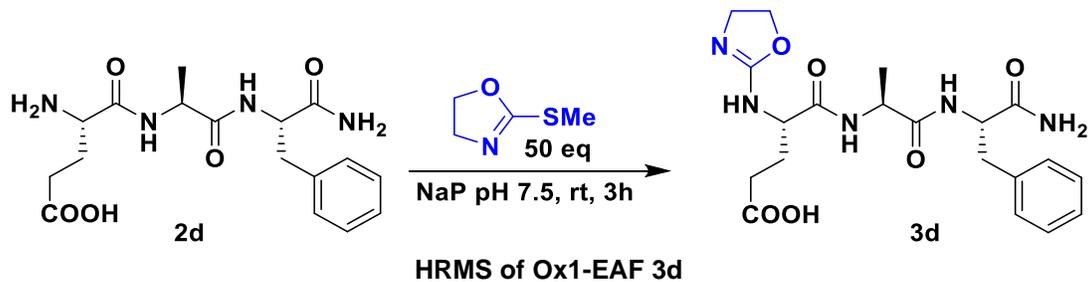
Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 8.50 min.

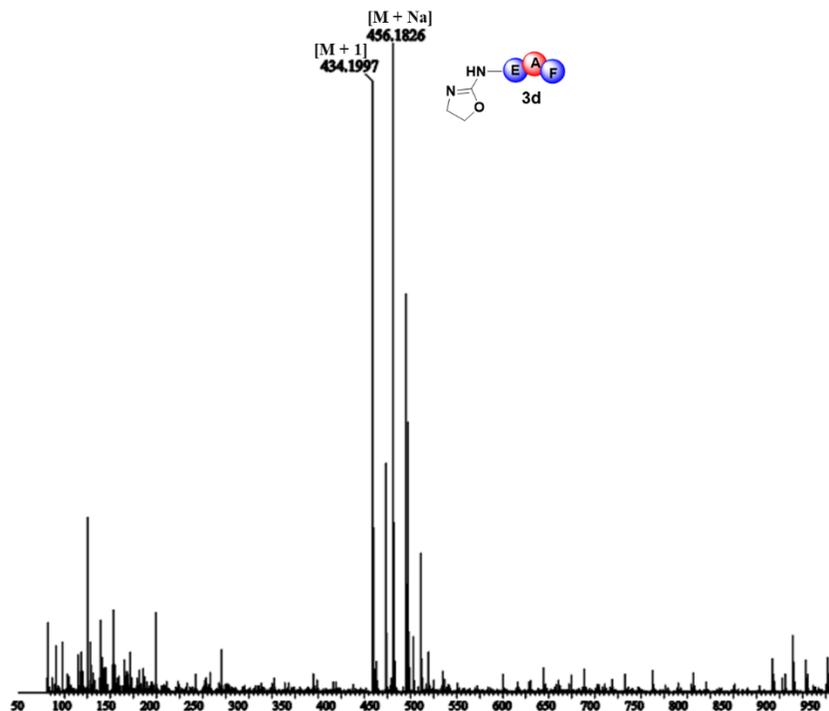
HRMS of OX-DAF 3c



Ox1-EAF 3d. LCMS: m/z 434.1 (calcd $[M+H]^+ = 434.1$), m/z 456.1 (calcd $[M+Na]^+ = 456.1$),

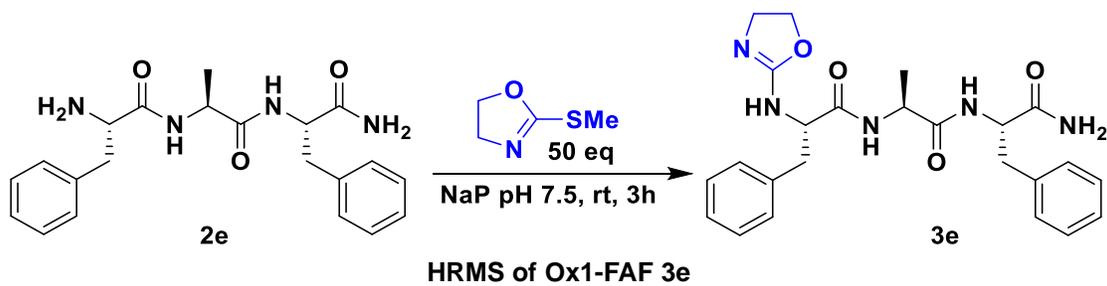
Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 8.11 min.

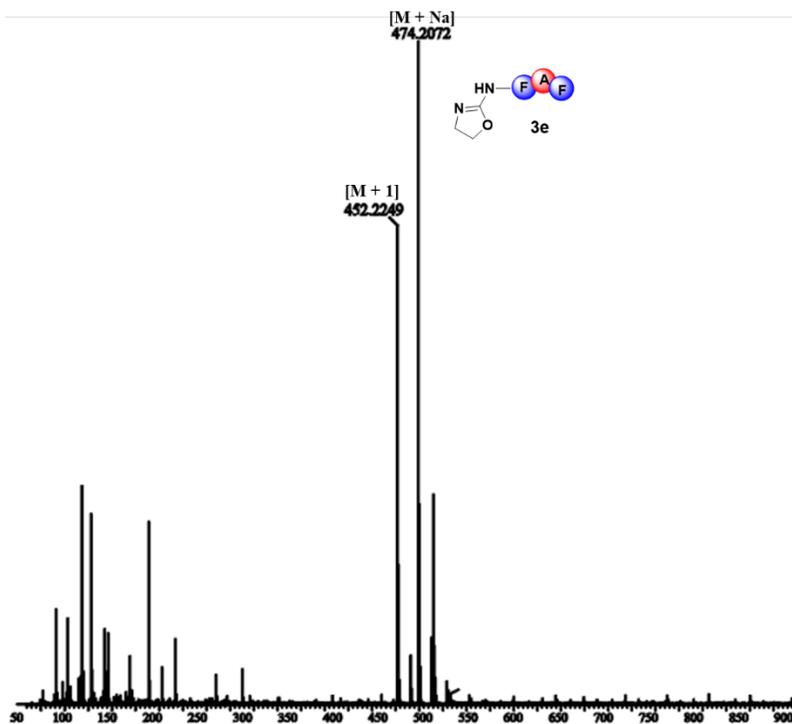




Ox1-FAF 3e. LCMS: m/z 452.2 (calcd $[M+H]^+ = 452.2$), m/z 474.2 (calcd $[M+Na]^+ = 474.2$),

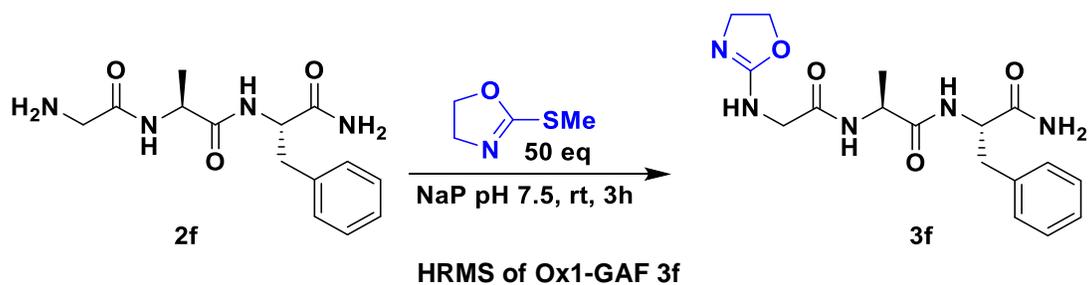
Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 11.96 min

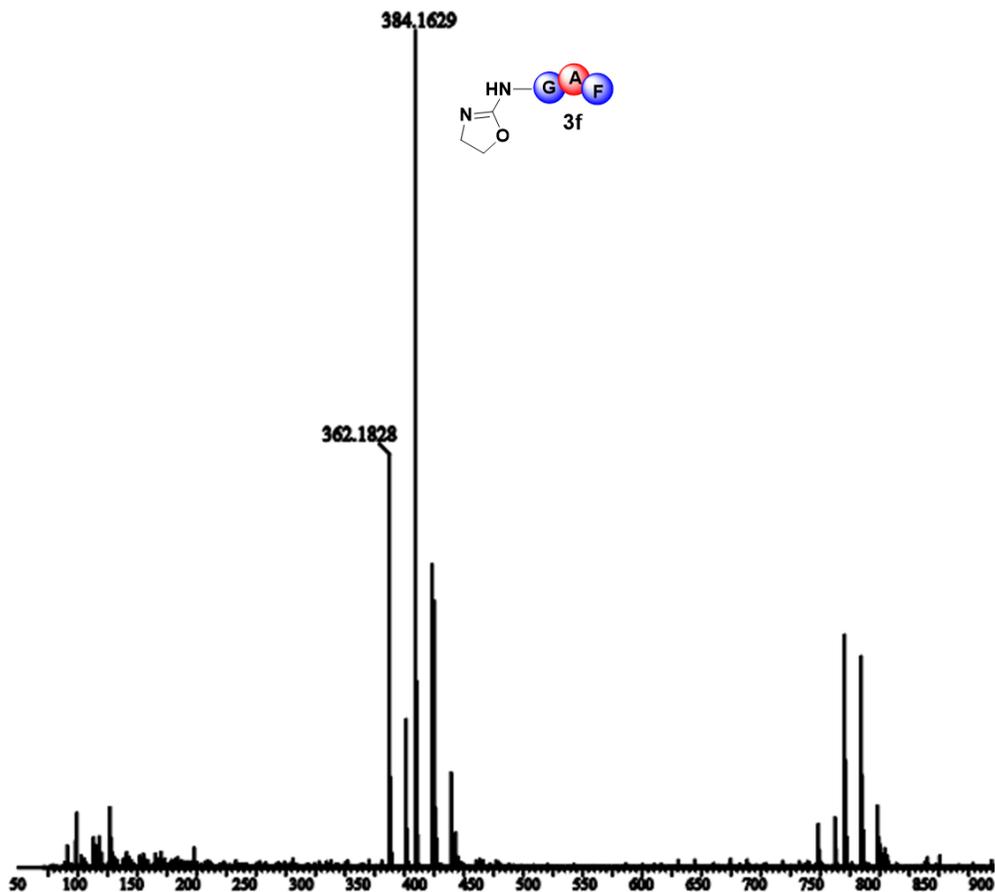




Ox1-GAF 3f. LCMS: m/z 362.18 (calcd $[M+H]^+ = 362.2$), m/z 384.16 (calcd $[M+Na]^+ = 385.2$),

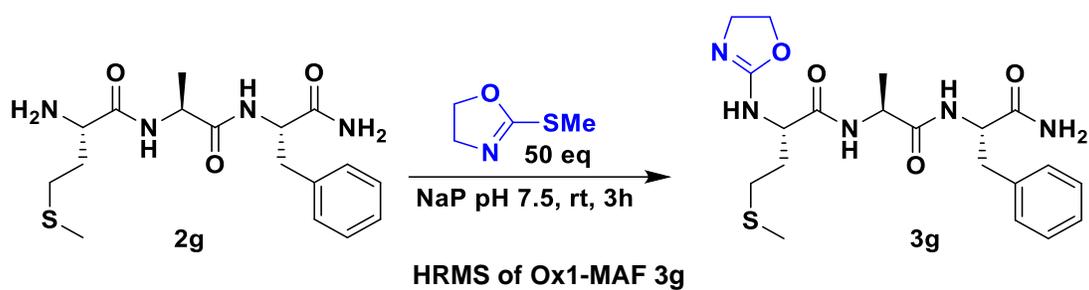
Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 7.21 min

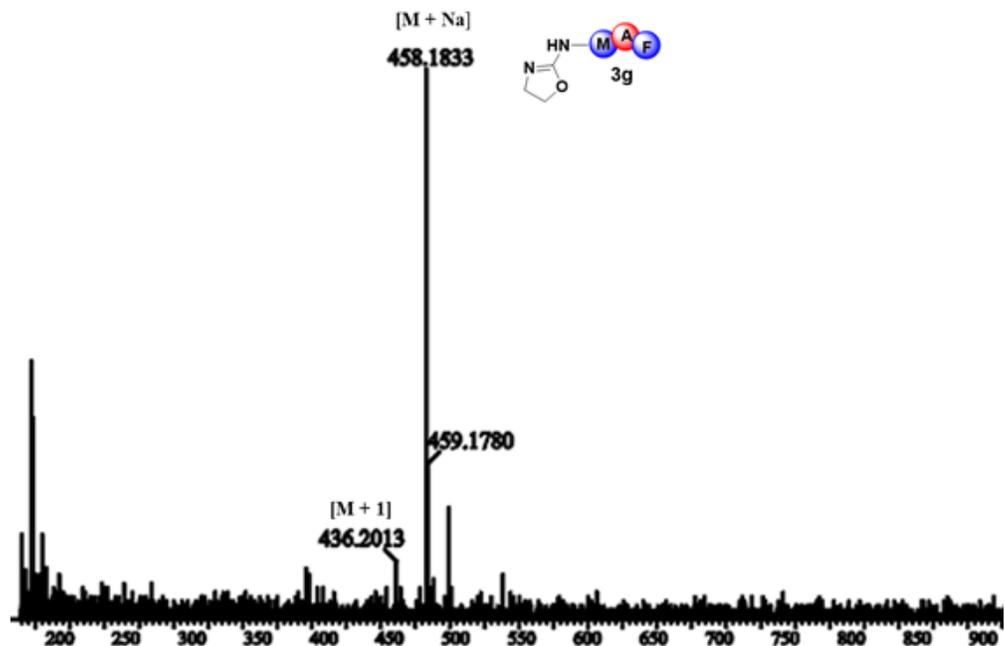




Ox1-MAF 3g. LCMS: m/z 436.2 (calcd $[M+H]^+ = 436.1$), m/z 458.1 (calcd $[M+Na]^+ = 458.1$),

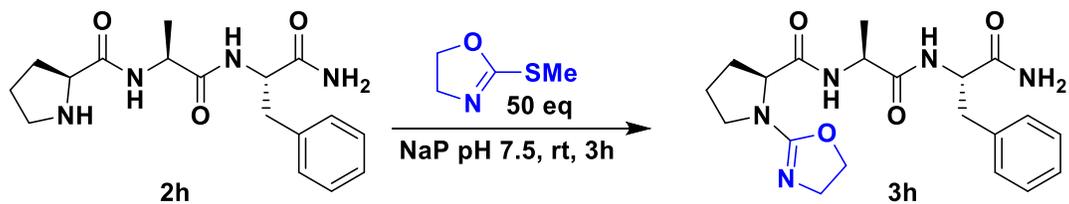
Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 10.20 min



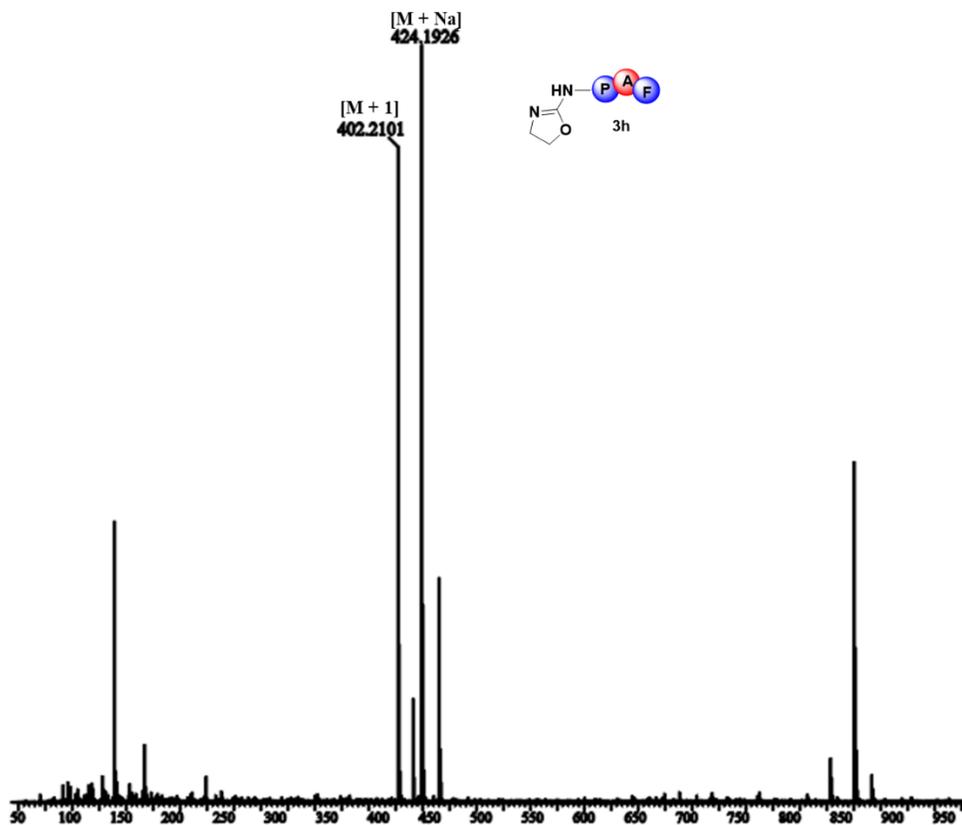


Ox1-PAF 3h. LCMS: m/z 402.2 (calcd $[M+H]^+ = 402.2$), m/z 424.1 (calcd $[M+Na]^+ = 424.2$),

Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 8.47 min

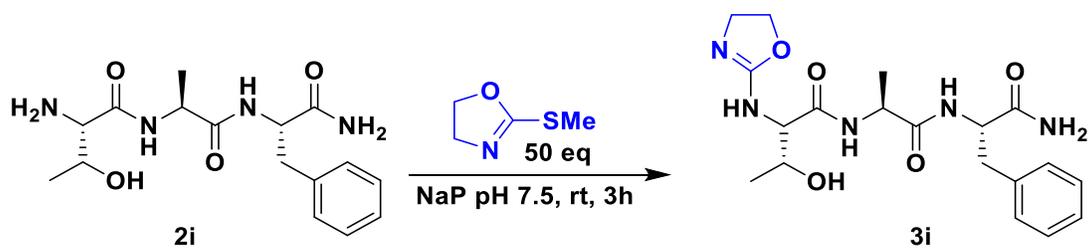


HRMS of Ox1-PAF 3h

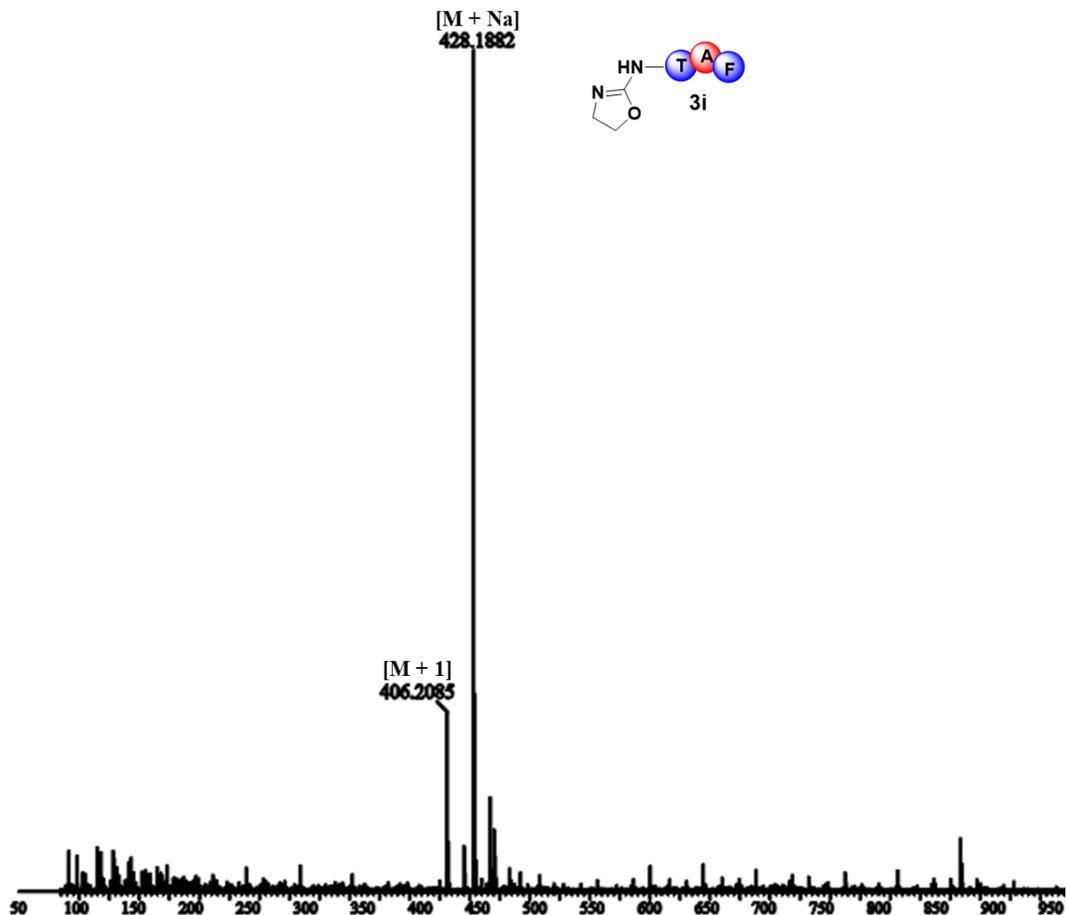


Ox1-TAF 3i. LCMS: m/z 406.2 (calcd $[M+H]^+ = 406.2$), m/z 428.2 (calcd $[M+Na]^+ = 428.2$),

Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 11.82 min

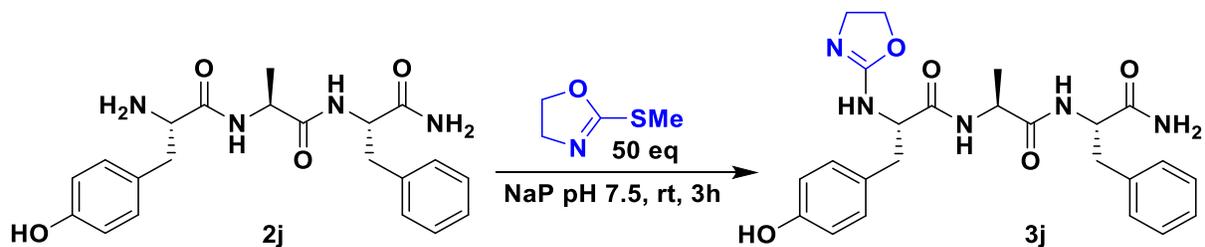


HRMS of Ox1-TAF 3i

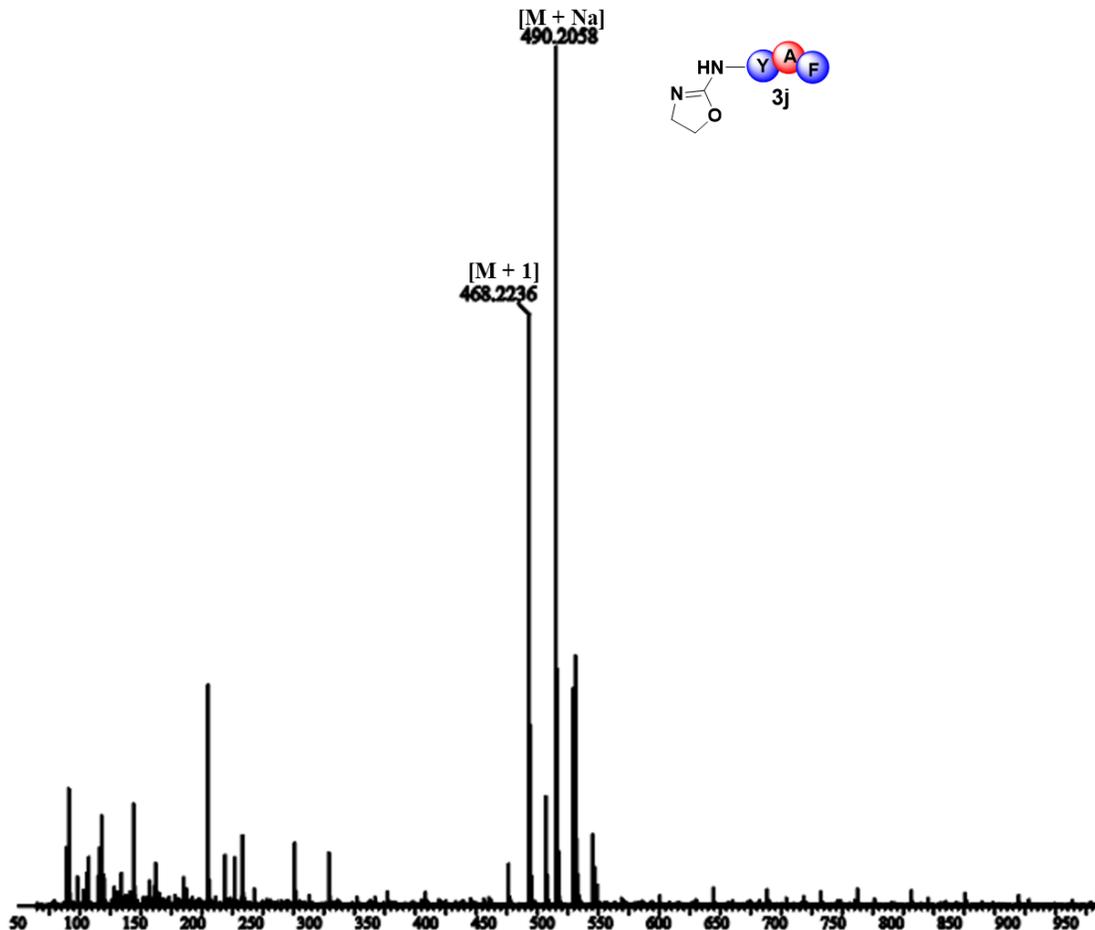


Ox1-YAF 3j. LCMS: m/z 468.2 (calcd $[M+H]^+ = 468.2$), m/z 490.2 (calcd $[M+Na]^+ = 490.2$),

Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 11.50 min



HRMS of Ox1-YAF 3j



Supplementary Figure 3. Azolation vs 2-PCA method

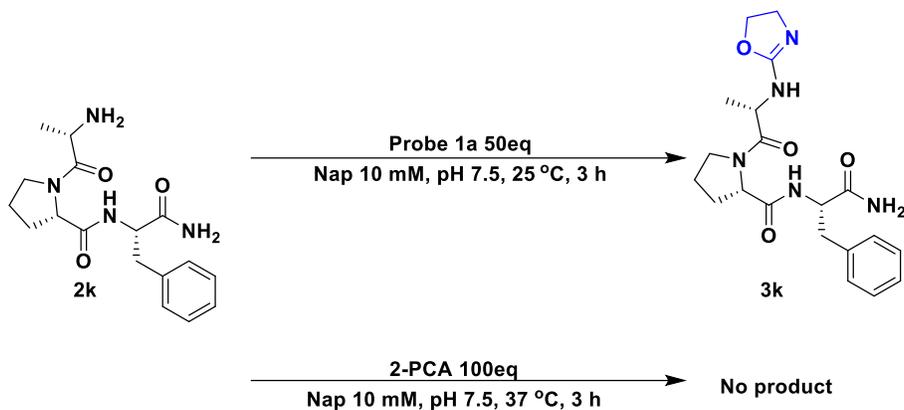
Procedure for synthesis of Ox1-APF 3k.

To peptide APF **2k** (0.6 mM) in 0.4 mL of 10 mM phosphate buffer (Nap, pH 7.5) was added 2-methylthio oxazoline **1a** (50 equiv.). The solution was stirred at room temperature for 3 h. The reaction was analyzed by LC/MS. LC: water with 1 % formic acid (solvent A): acetonitrile with 1 % formic acid (solvent B); gradient 0-80 %, acetonitrile in 30 min, flow rate = 1.0 mL/min, detection wavelength 220.

2-PCA method.

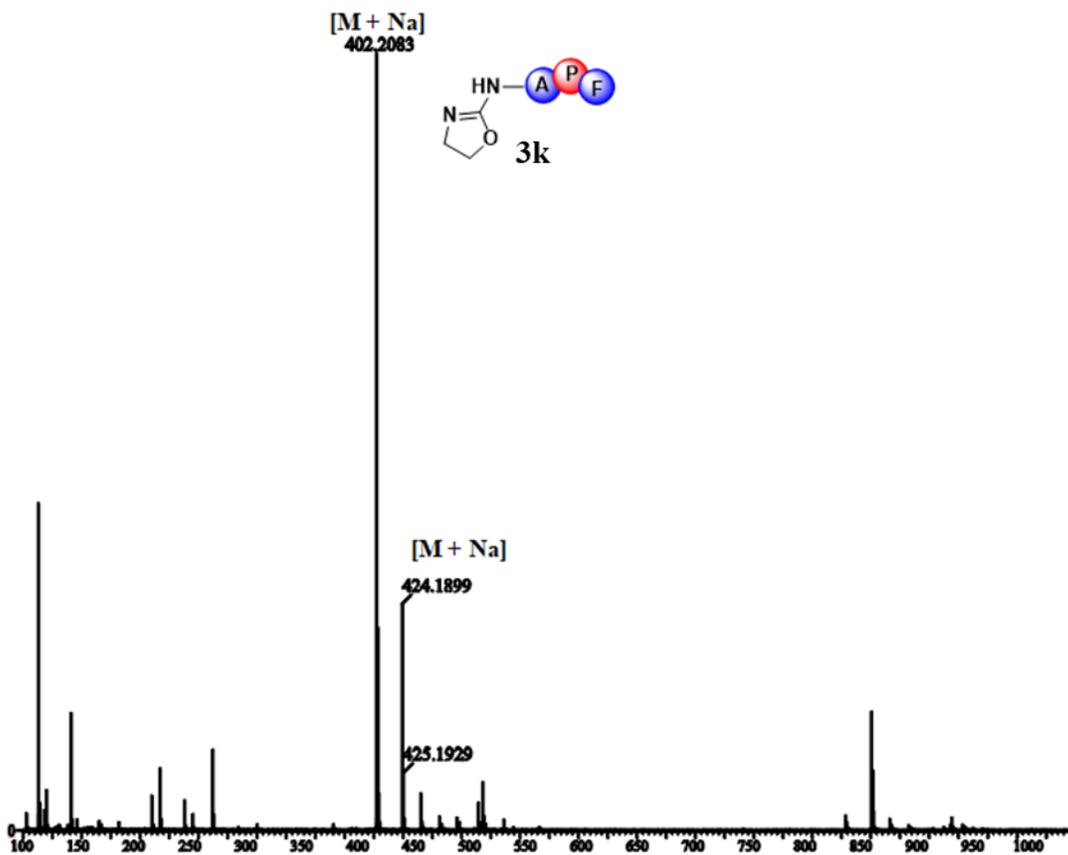
APF **2k** (0.1 mM) was reacted with 10 mM 2PCA in 10 mM phosphate buffer (Nap, pH 7.5) for 4 h at 37 °C. The reaction was analyzed by LC/MS. LC: water with 1 % formic acid (solvent A):

acetonitrile with 1 % formic acid (solvent B); gradient 0-80 %, acetonitrile in 30 min, flow rate = 1.0 mL/min, detection wavelength 220.



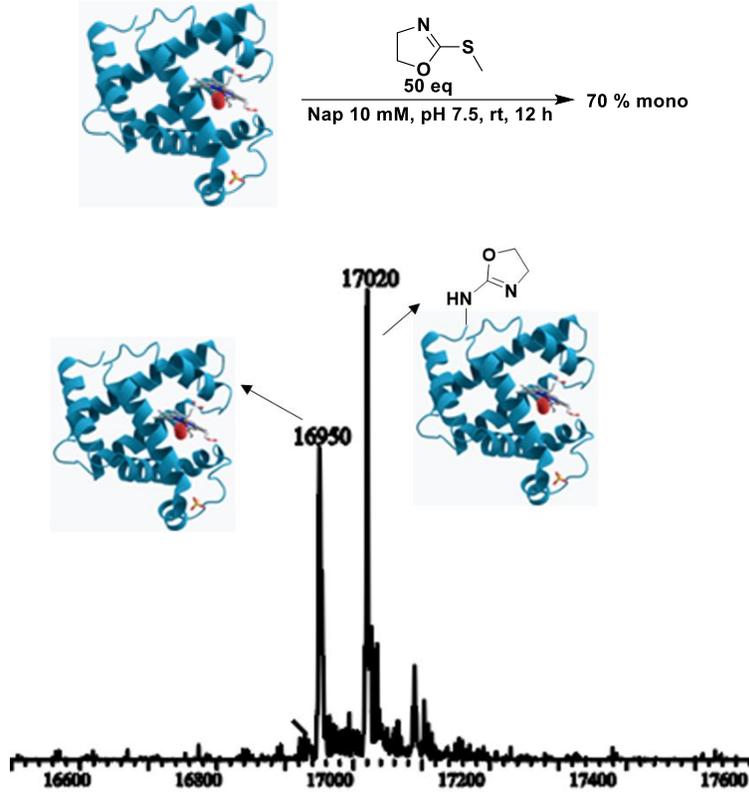
Ox1-APF 3k. LCMS: m/z 402.1 (calcd $[M+H]^+ = 402.2$), m/z 424.1 (calcd $[M+Na]^+ = 424.2$),

Purity: >95% (HPLC analysis at 220 nm). Retention time in HPLC: 8.50 min

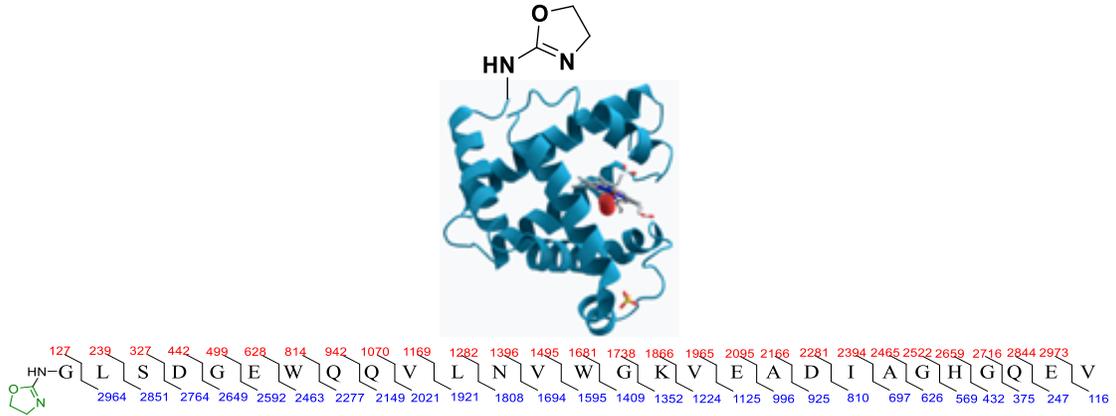


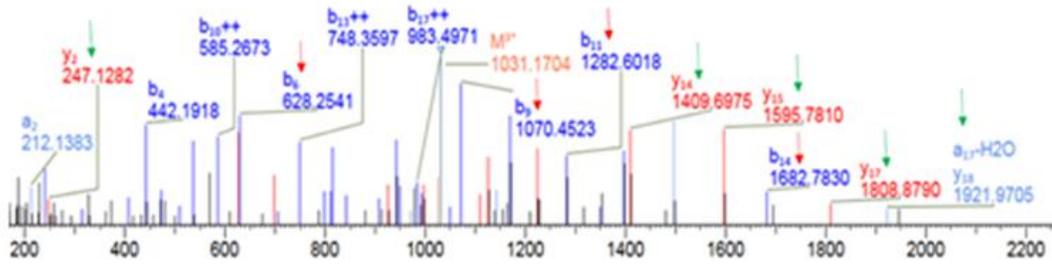
Supplementary Figure 4. Optimization of the oxazolization on protein Myoglobin Mb.

MS spectrum of Ox1-Mb at pH 7.5

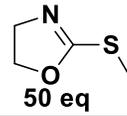
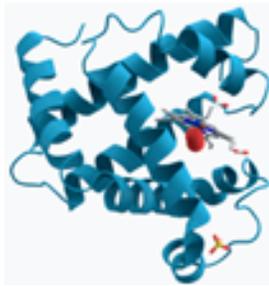


ESI MS-MS spectrum of Ox1-Mb.

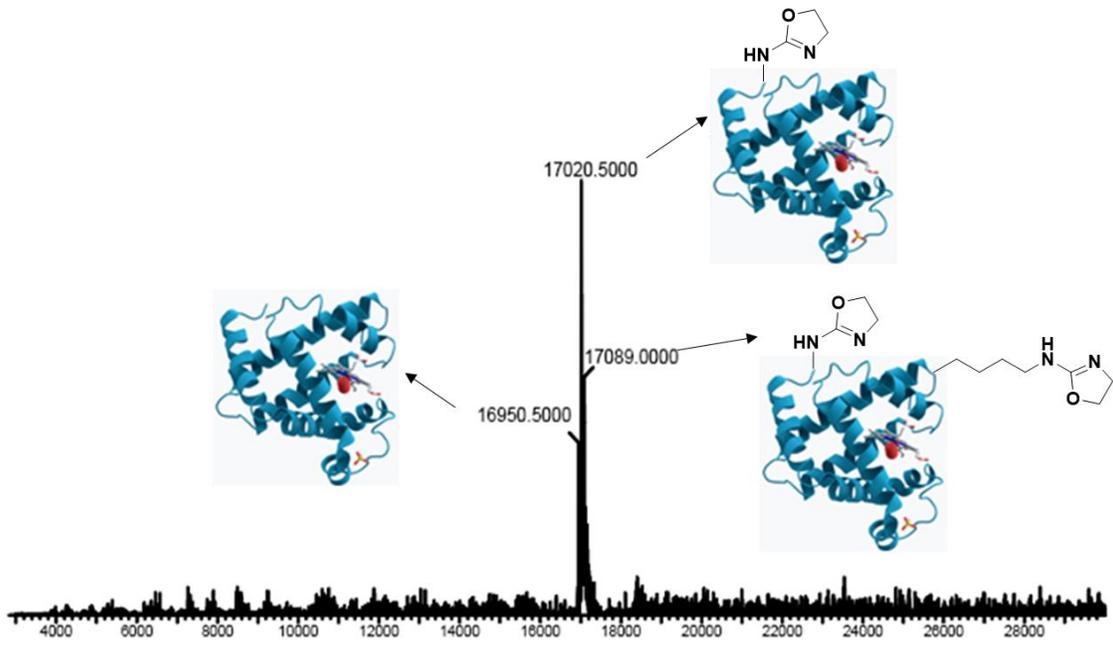


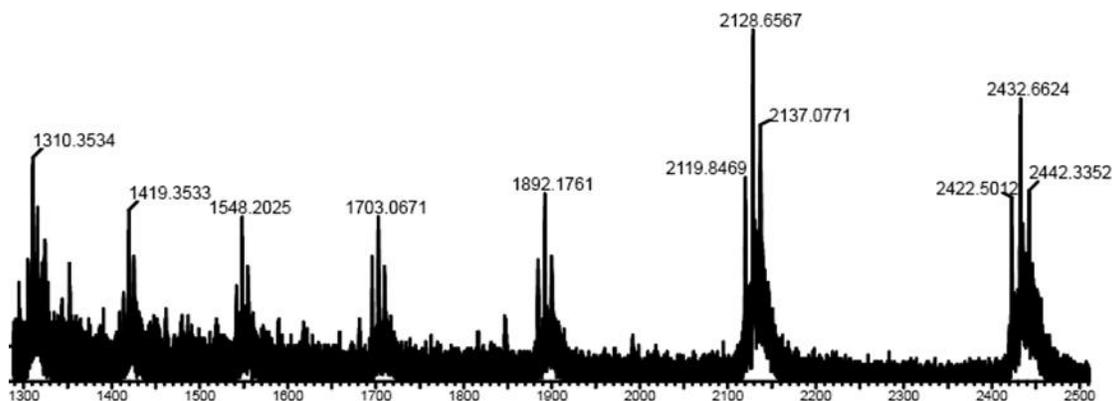


Oxazolation of Myoglobin at pH 8.5



50 eq
 Nap 10 mM, pH 8.5, rt, 12 h → 52 % mono, 30 % double





Supplementary Figure 5. Modification of myoglobin Mb by different oxazoline derivatives.

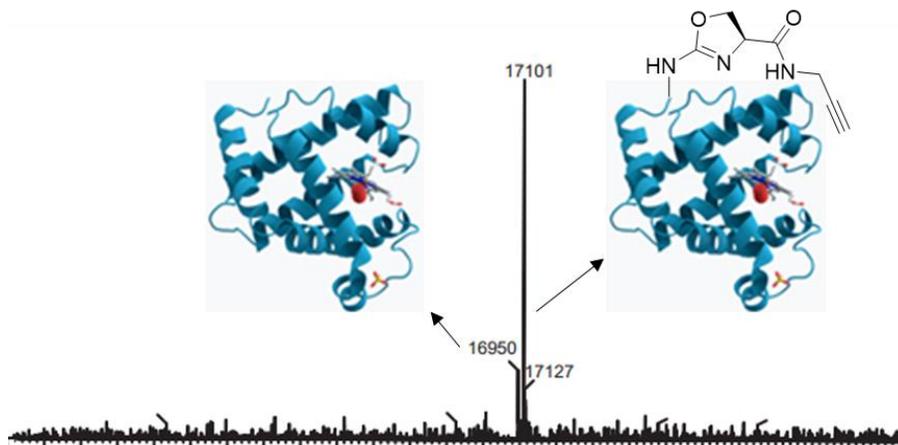
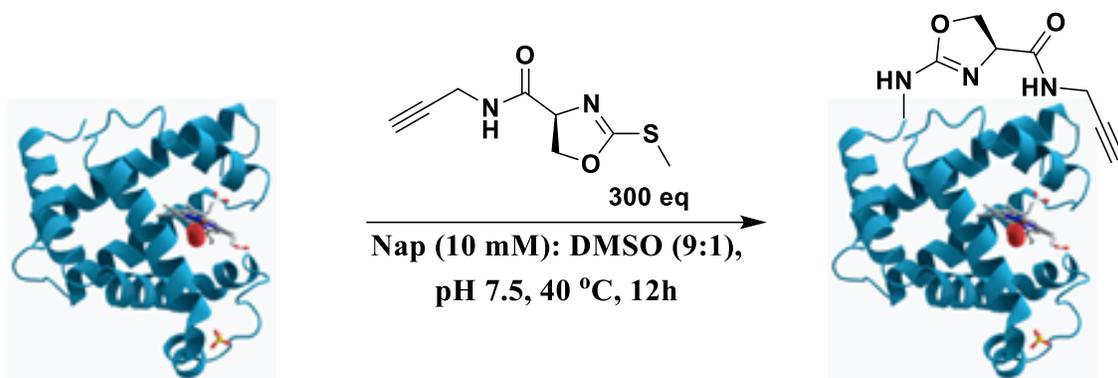
Functionalized-oxazoline myoglobin bioconjugation.

To a 1 mg of myoglobin (3 mM) in 0.2 mL 10 mM phosphate buffer (pH 7.5) with 10% DMSO, the oxazoline derivatives (300 equiv.) were added. The reaction was stirred at 40 °C for overnight. The oxazoline-protein bioconjugates were purified by molecular weight cut off and characterized by LCMS.

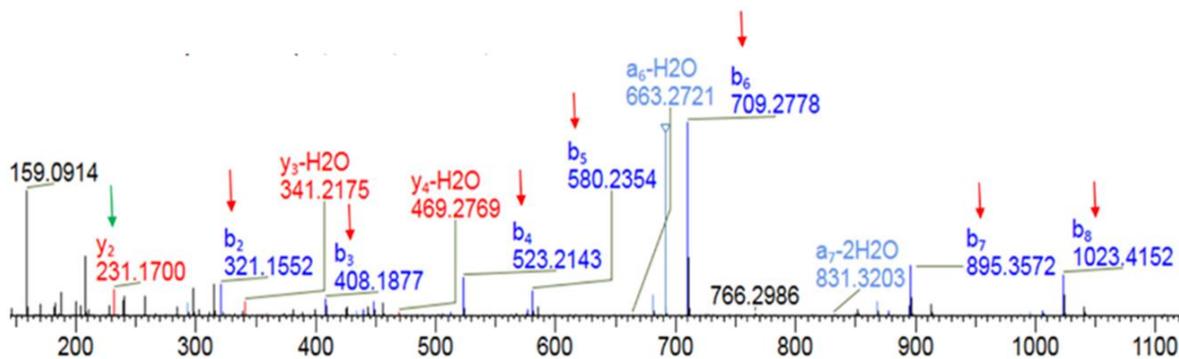
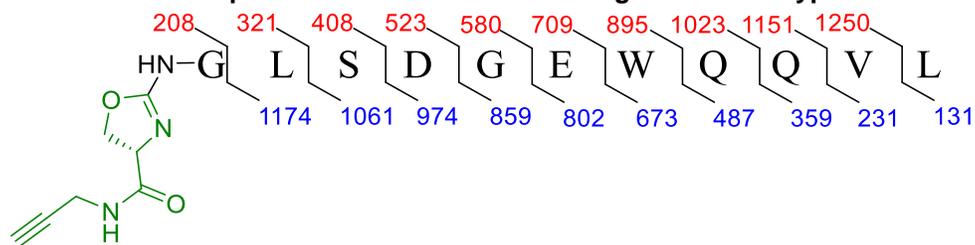
Digestion of functionalized-oxazoline myoglobin.

Protein (1 mg) in 100 mM tris (100 μ L, pH 7.5) with urea (6 M) was incubated for 30 min at 37 °C. To reduce the urea concentration to 0.6 M, the sample was diluted with grade I water. To this solution, 100 μ L of enzyme (α -chymotrypsin/trypsin) solution [0.1 mg, enzyme/protein (1:20); enzyme in 1 mM HCl was dissolved in 0.1 M tris and 0.01 M CaCl_2] was added and the mixture was incubated at 37 °C for 18 h. The sample was used for peptide mapping by MS and sequencing by MS/MS.

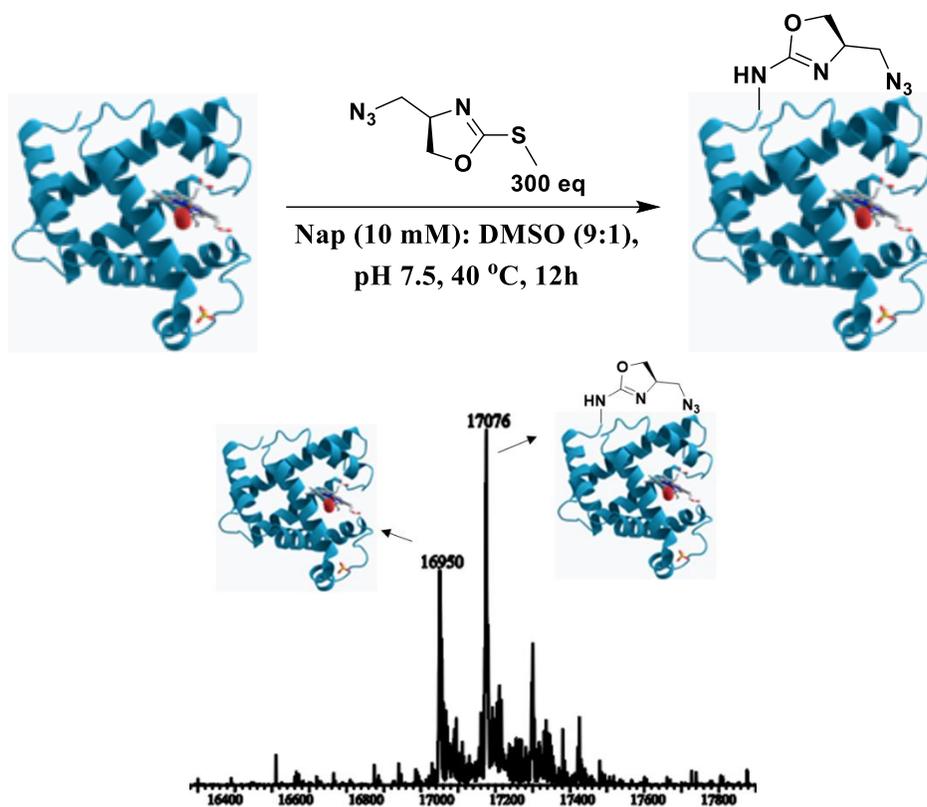
Modification of Mb with Ox2:



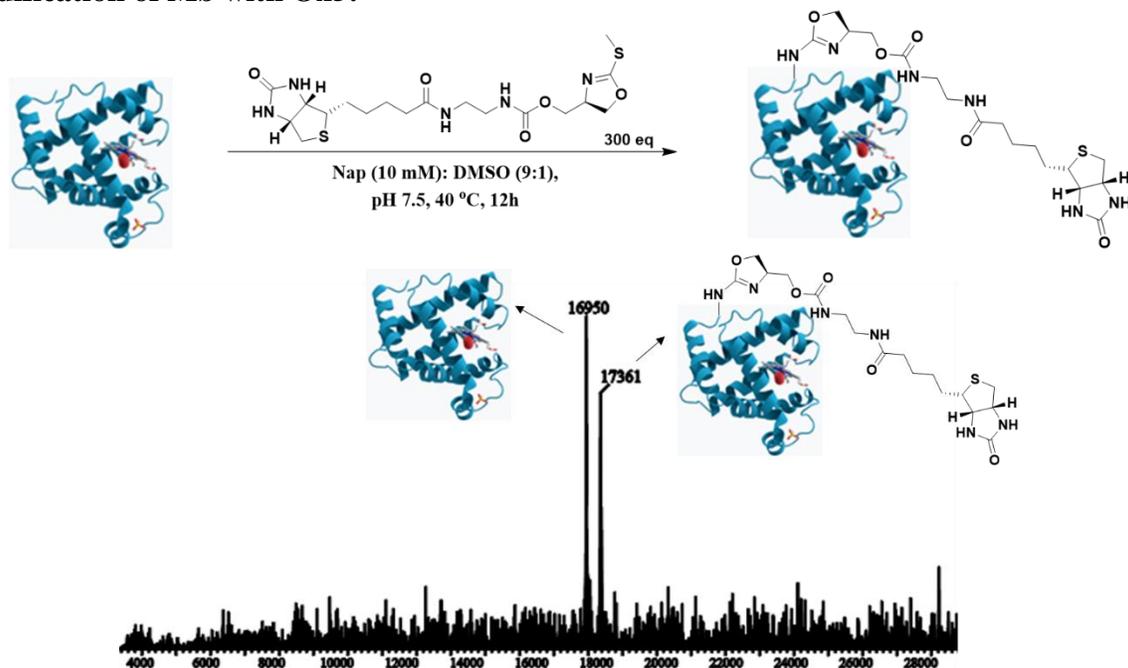
MS/MS spectrum of Ox2-Mb after the digestion with trypsin



Modification of Mb with Ox3:



Modification of Mb with Ox5:

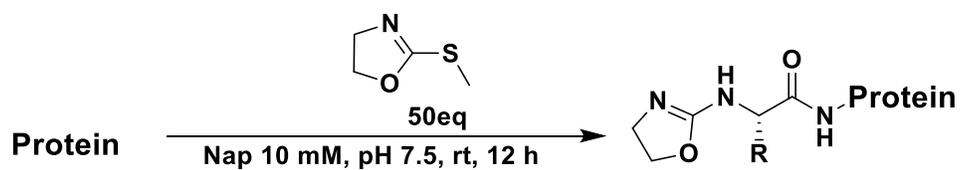


Procedure for bioconjugation of functionalized-oxazoline with different proteins.

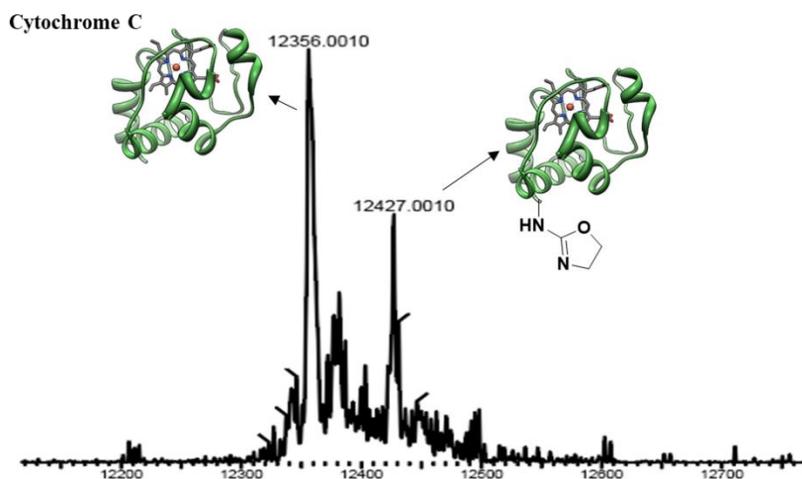
To proteins (3 mM) in 0.2 mL 10 mM phosphate buffer (pH 7.5), oxazoline **1a** (50 equiv., 150

mM) was added. The reactions were incubated at room temperature for overnight. The oxazoline-protein bioconjugates were purified by molecular weight cut off and characterized by LCMS.

Modification of a variety of proteins with compound 1a

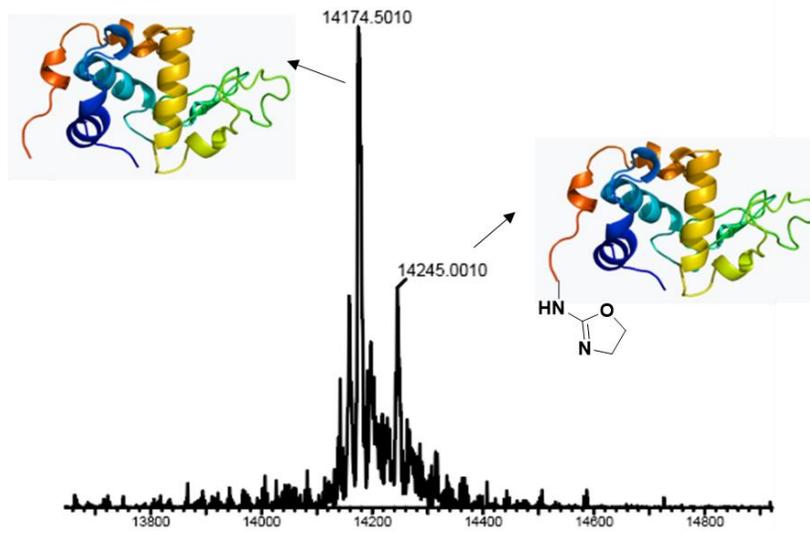


HRMS of Ox1-Cytochrome C

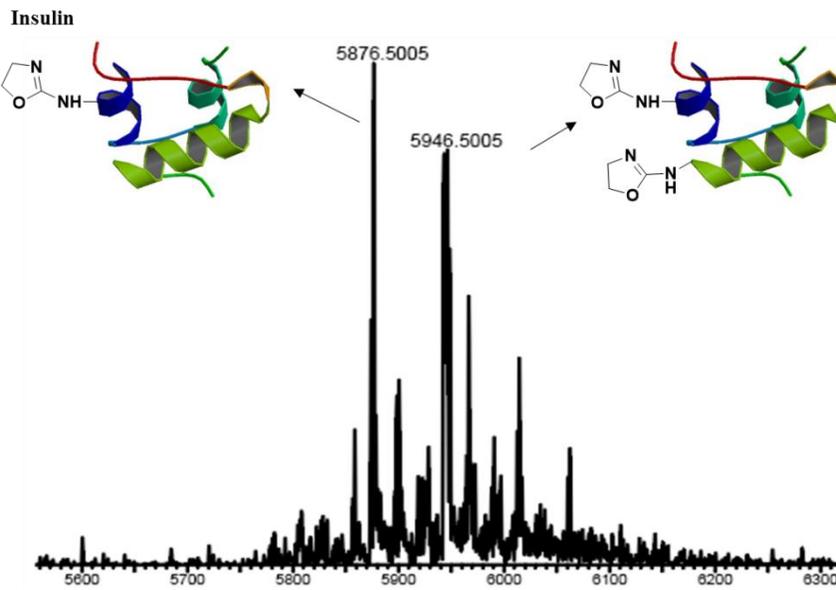


HRMS of Ox1- α -Lactalbumin

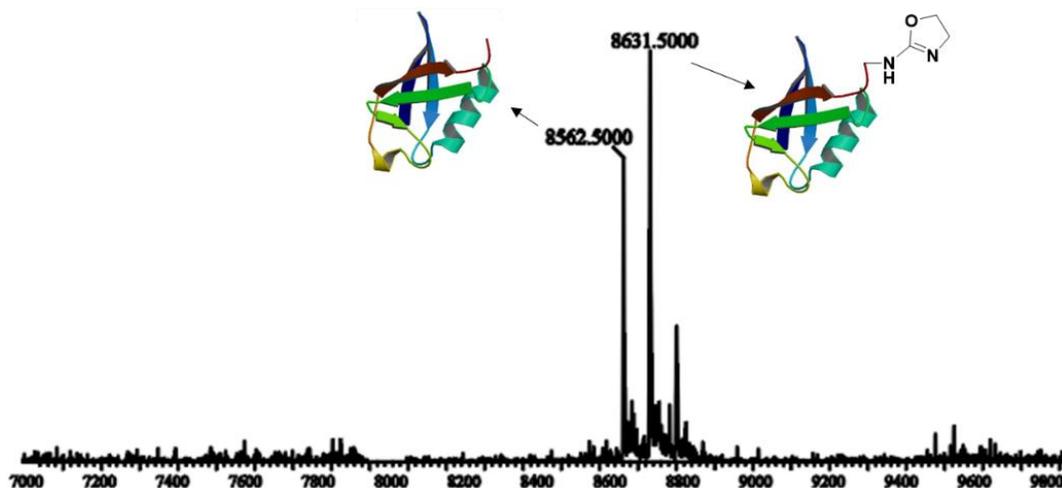
α -Lactalbumin



HRMS of Ox1-Insulin (two N-terminus)



HRMS of Ox1-Ubiquitin



Procedure of oxazolation of various proteins substrates by Ox2, Ox3, Ox4, Ox5

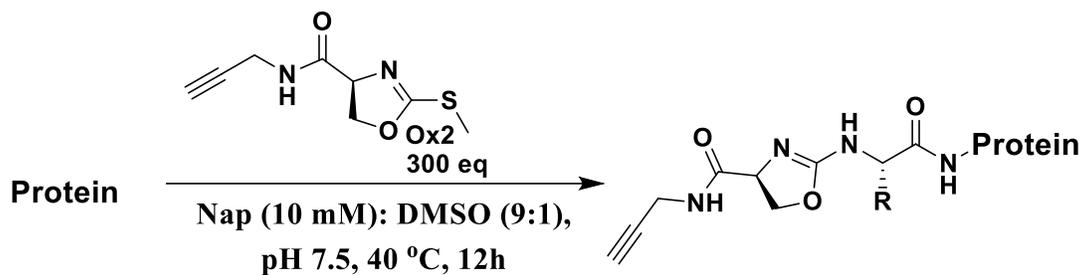
To proteins (3 mM) in 0.2 mL 10 mM phosphate buffer (pH 7.5) with 10% DMSO, the oxazoline derivatives (100-300 equiv.) were added. The reaction was incubated at 40 °C for overnight. The functionalized-oxazoline-protein bioconjugates were purified by molecular weight cut off and characterized by LCMS.

Procedure for digestion of Ox2-aldolase:

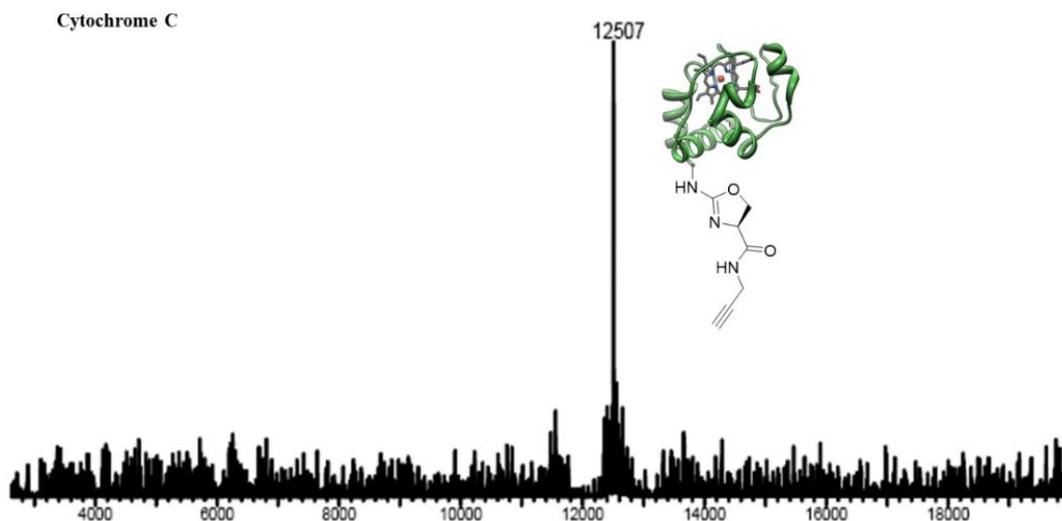
Modified protein (1 mg) in 100 mM tris (100 μ L, pH 7.5) with urea (6 M) was incubated for 30 min at 37 °C. To this solution, reducing agent (1 μ L, 0.2 M DTT in 0.1 M tris) was added and sample was incubated for 1 h at 37 °C. Alkylating agent (4 μ L, 0.2 M iodoacetamide in 0.1 M tris) was added to the solution and incubated (in dark) for 1 h at 25 °C for blocking the free sulfhydryl groups. The unreacted iodoacetamide was quenched with reducing agent (4 μ L, 0.2 M DTT in 0.1 M tris) for 1 h at 25 °C. To reduce the urea concentration to 0.6 M, the sample was diluted with grade I water. To this solution, 100 μ L of enzyme (α -chymotrypsin/trypsin) solution [0.1 mg, enzyme/protein (1:20), enzyme in 1 mM HCl was dissolved in 0.1 M tris and 0.01 M CaCl₂] was added and the mixture was incubated at 37 °C for 18 h. The pH of digested mixture was adjusted to < 6 (confirmed by pH paper) with trifluoroacetic acid (0.5%). Subsequently, the

sample was used for peptide mapping by MS and sequencing by MS/MS.

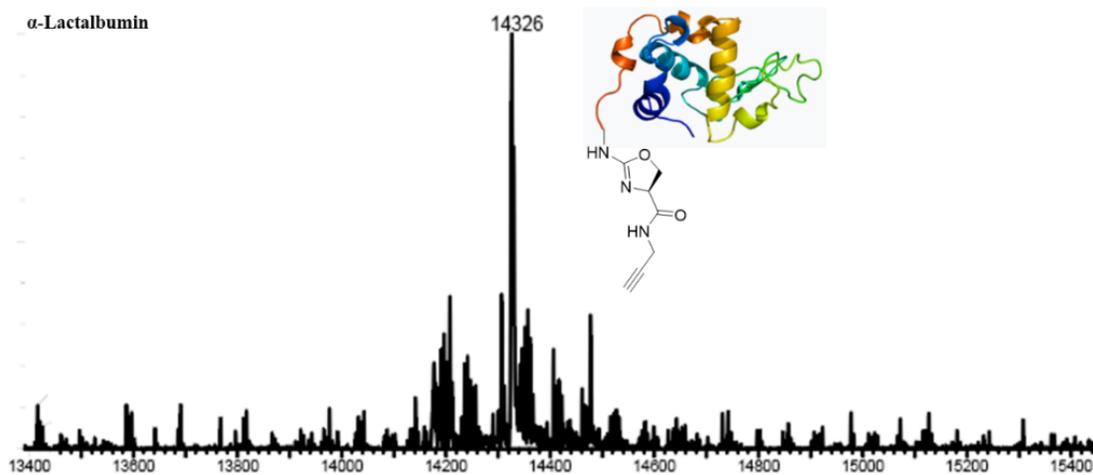
Modification of a variety of proteins with compound Ox2



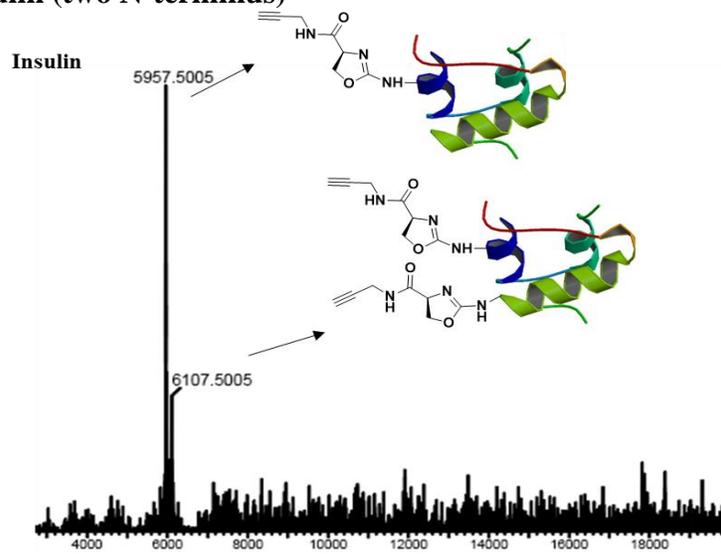
HRMS of Ox2-Cytochrome C



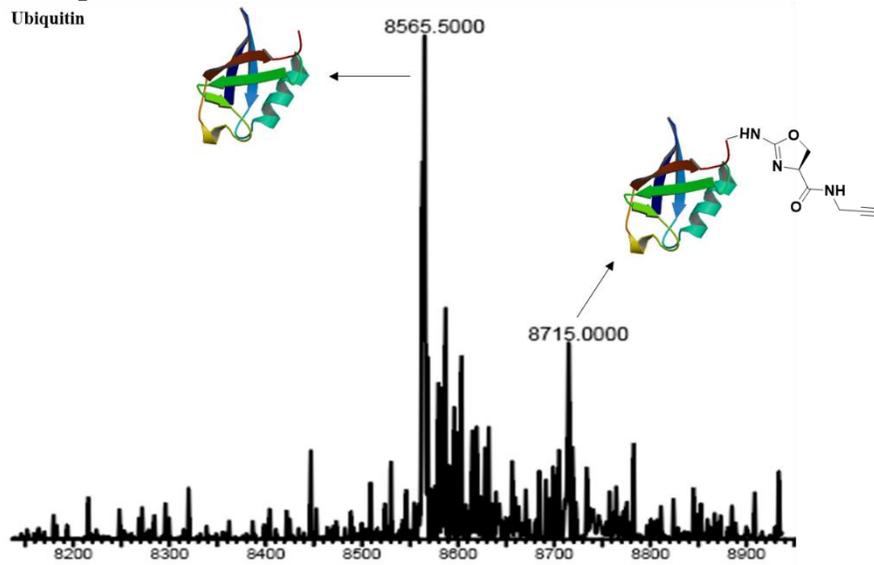
HRMS of Ox2- α -Lactalbumin



HRMS of Ox2-Insulin (two N-terminus)

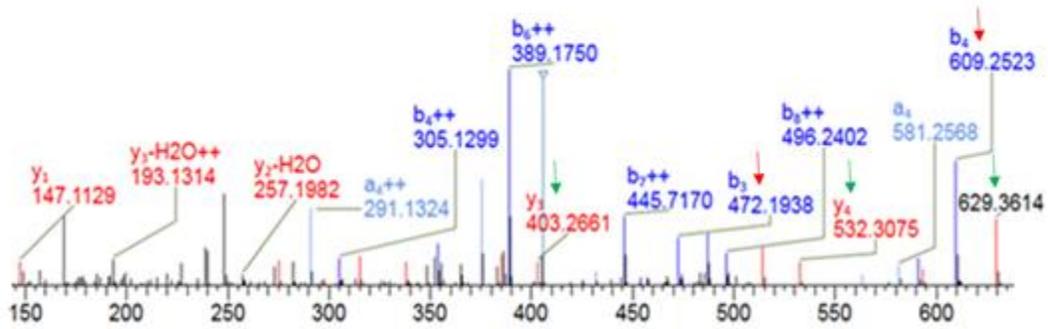


HRMS of Ox2-Ubiquitin

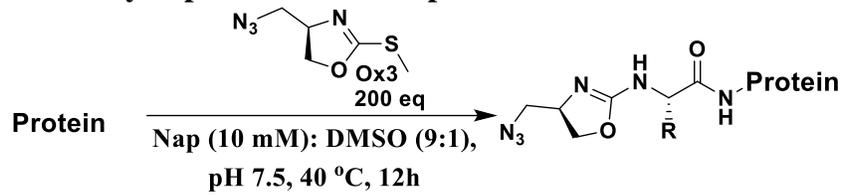


MS/MS spectrum of Ox2-Aldolase after the digestion with trypsin

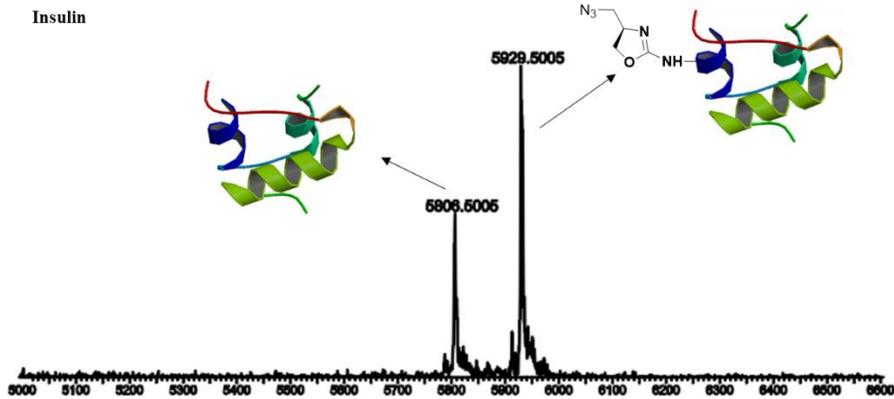




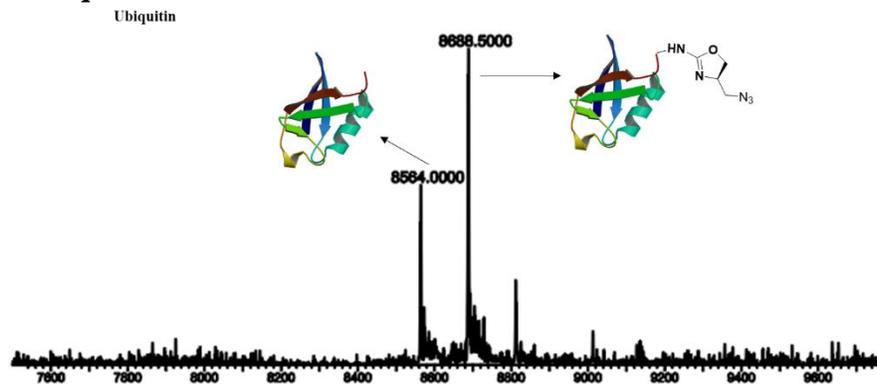
Modification of a variety of proteins with compound Ox3



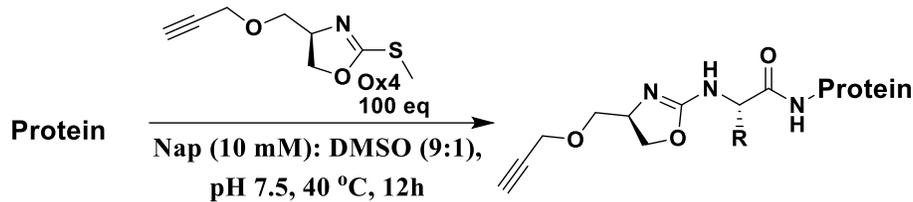
HRMS of Ox3-Insulin



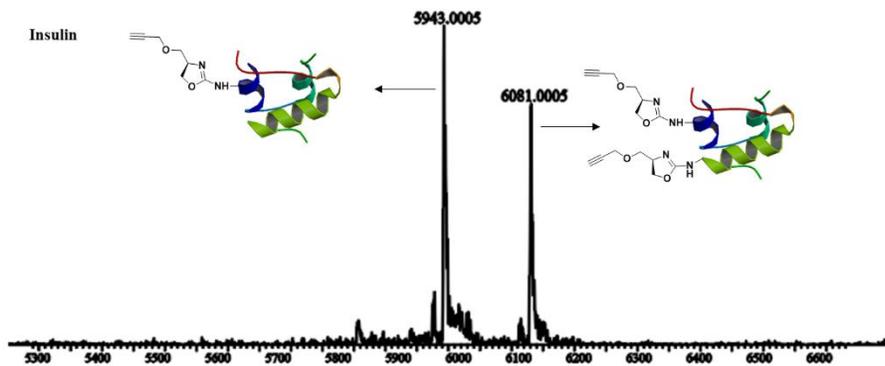
HRMS of Ox3-Ubiquitin



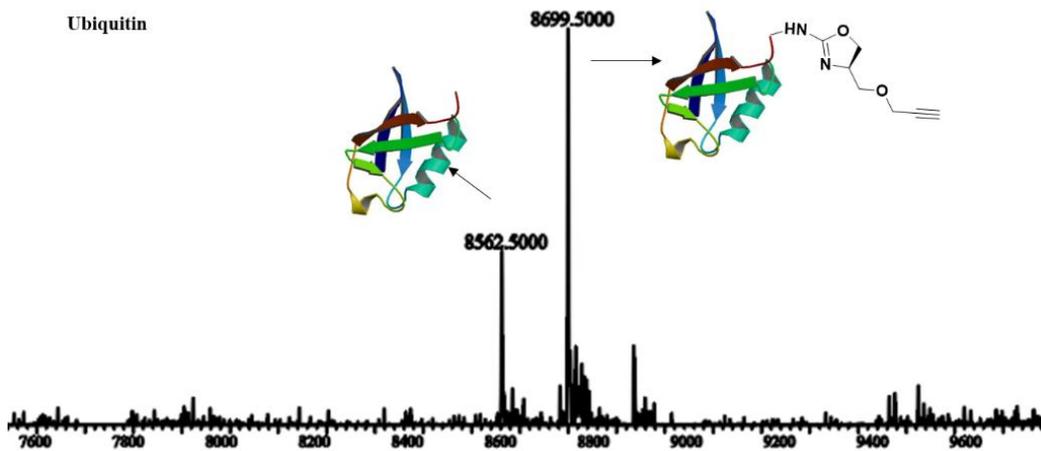
Modification of a variety of proteins with compound Ox4



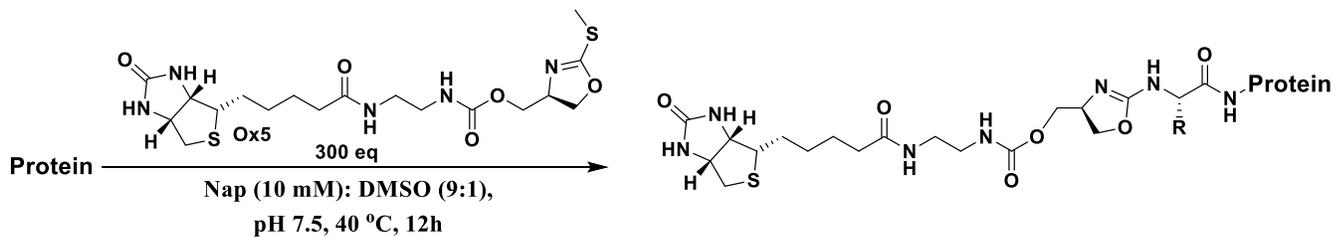
HRMS of Ox4-Insulin (two N-terminus)



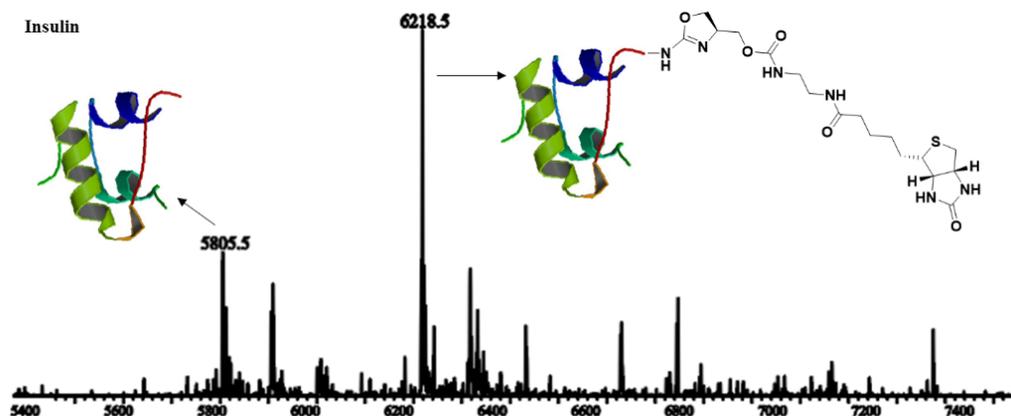
HRMS of Ox4-Ubiquitin



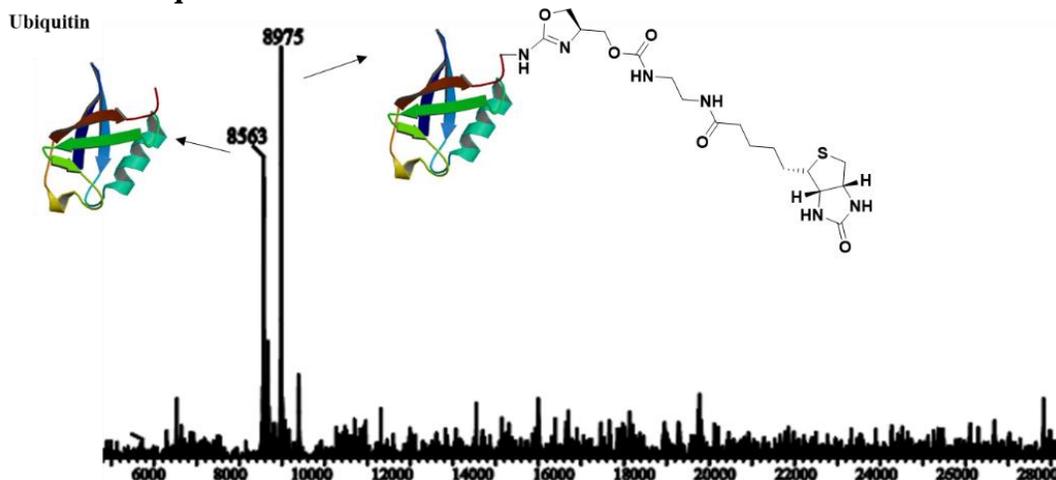
Modification of a variety of proteins with compound Ox5



HRMS of Ox5-Insulin



HRMS of Ox5-Ubiquitin

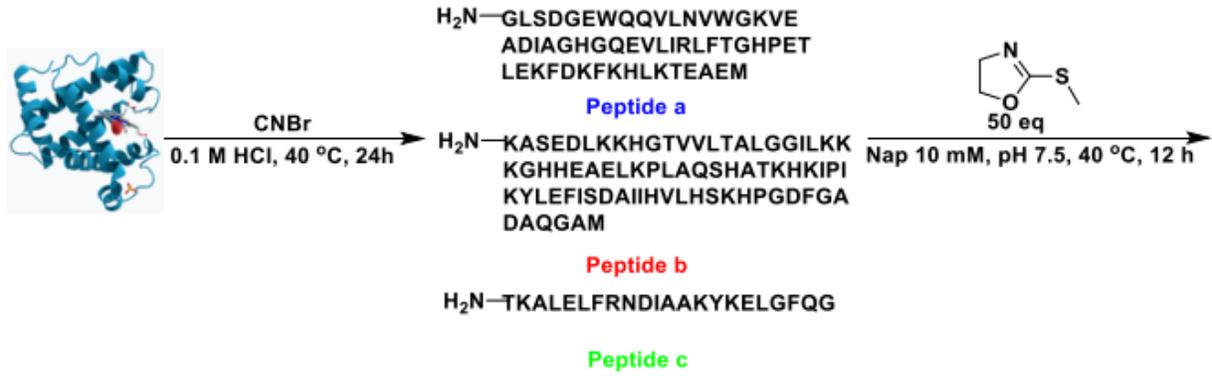


Supplementary Figure 6. Mass intensity enhancement of digested myoglobin by oxazolone.

Procedure for the digestion of myoglobin by CNBr.

1 mg (0.25 mM) of myoglobin and 0.3 mg of CNBr (0.1875 mM) were mixed in 324 μ L 0.1 M HCl, the reaction mixture was incubated at 40 °C for 24 h. The reaction mixture was quenched by freezing the sample at -80 °C. The frozen samples were then lyophilized to afford the dry peptide fragments. Procedure for oxazolone of digested myoglobin. To the 1 mg mixture of proteolytic fragments in 0.4 mL of 10 mM sodium phosphate buffer (Nap, pH 7.5), 1a Ox1 (50

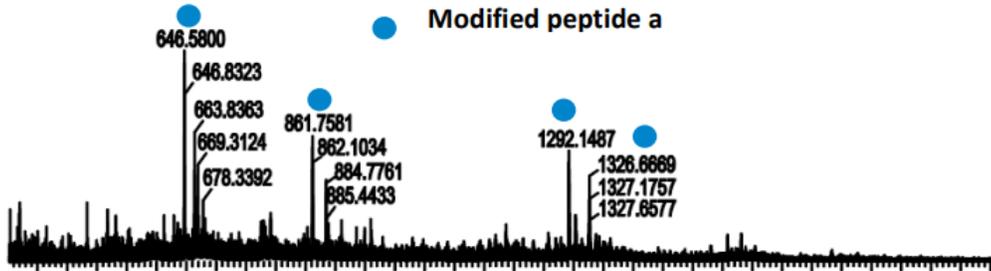
equiv.) was added. The reaction was allowed to react at 40 °C in the incubator for 12 h. The tagged proteolytic fragments were analyzed by LCMS without purification.

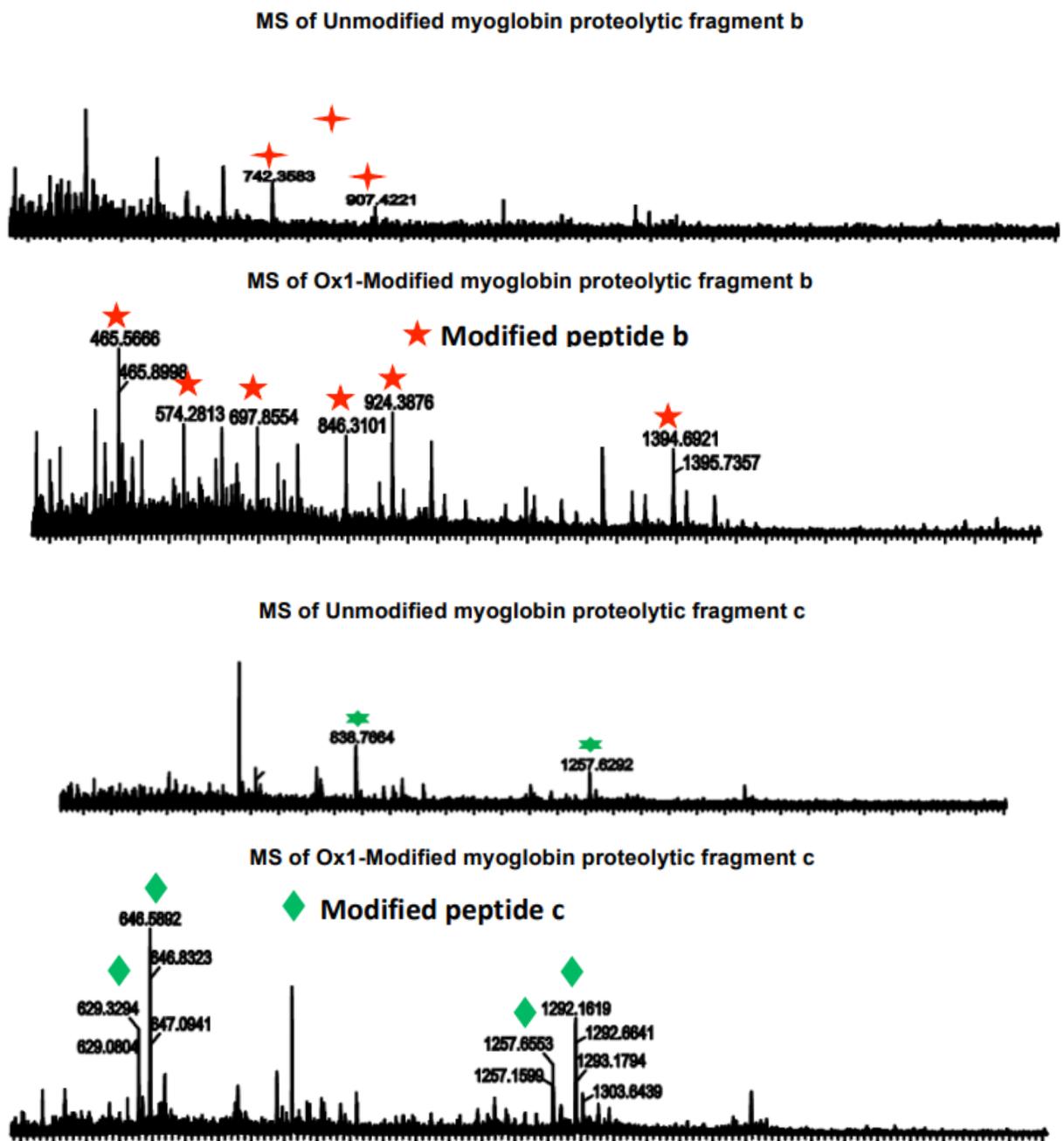


MS of Unmodified myoglobin proteolytic fragment a



MS of Ox1-Modified myoglobin proteolytic fragment a

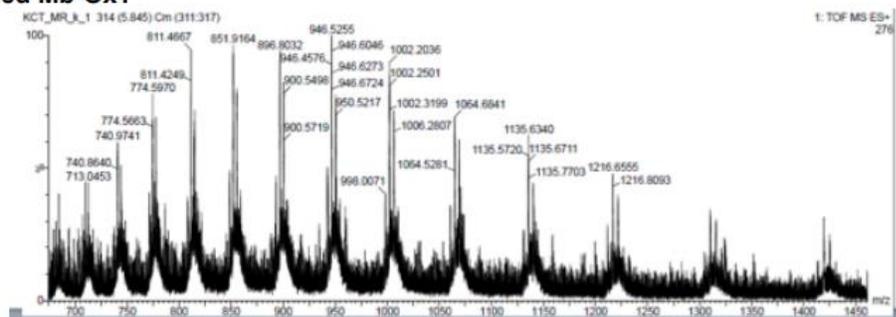




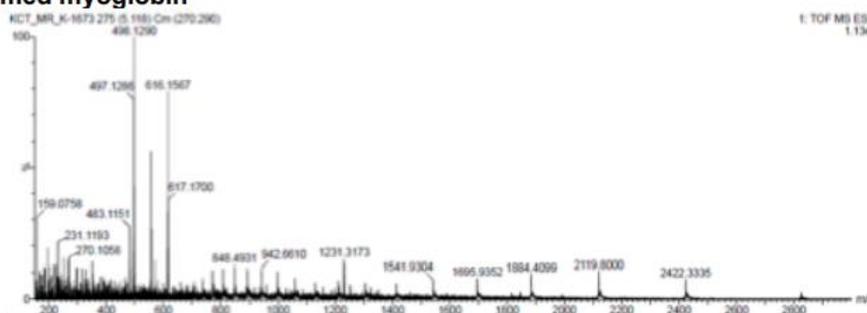
Supplementary Figure 7. Mass intensity enhancement of intact Myoglobin by oxazolation as compared to unmodified myoglobin.

Procedure for oxazolation of myoglobin. Protein Mb (0.6 μmol) in 10 mM phosphate buffer of pH 7.5 (600 μL) and 1a, Ox1 (50 equiv.) was incubated overnight. The tagged proteins were analyzed by LCMS.

Modified Mb-Ox1



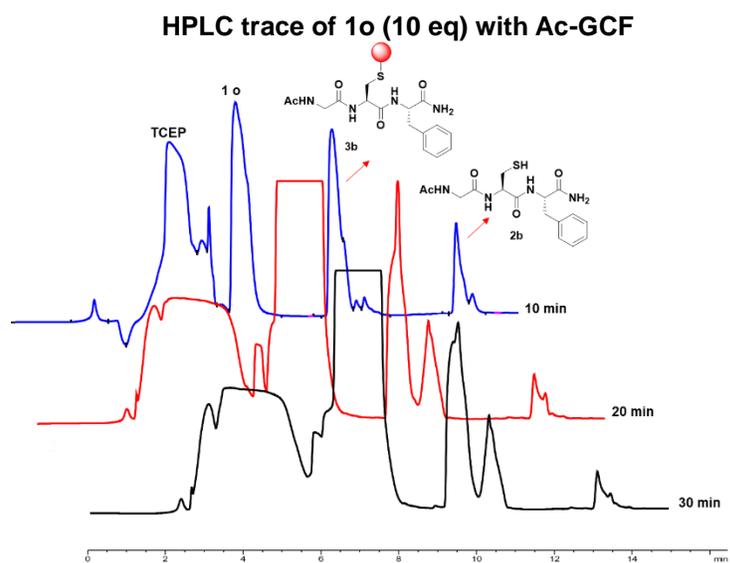
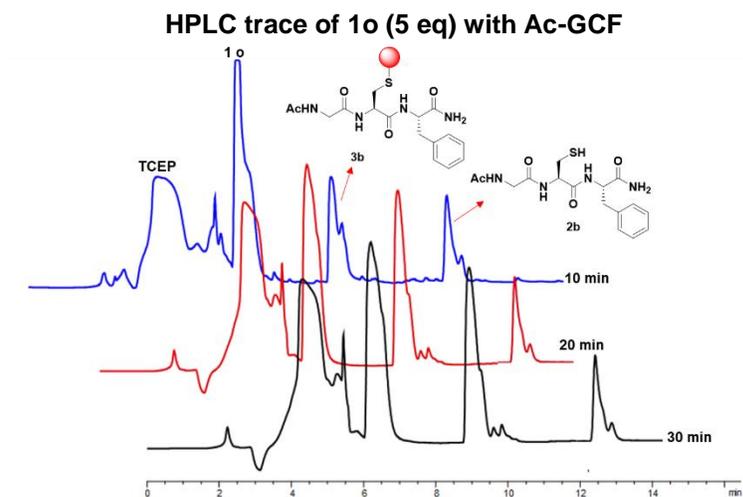
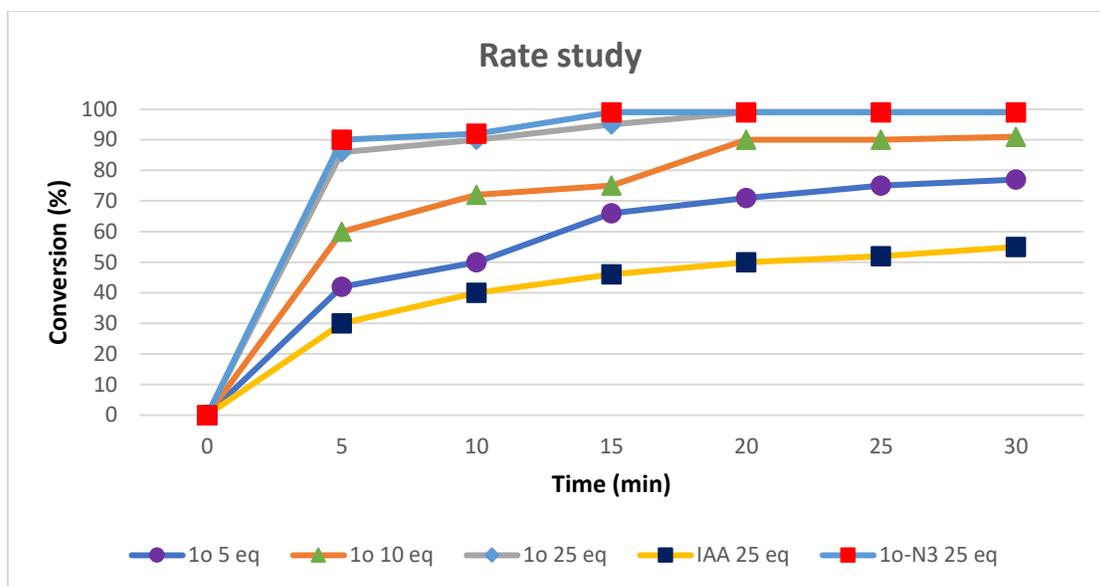
Unmodified myoglobin



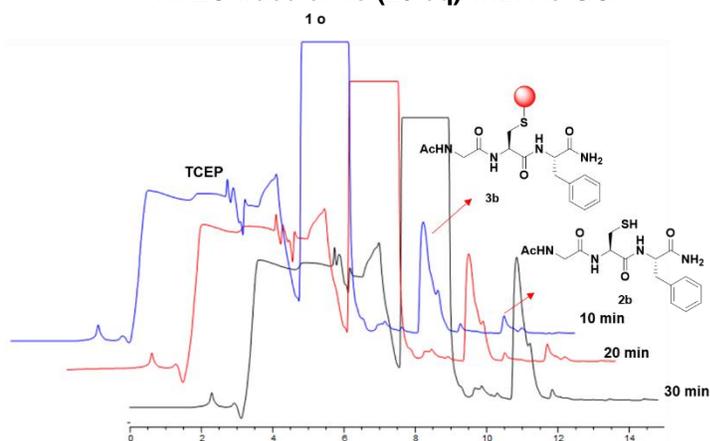
Supplementary Figure 8: Rate study of HAT bioconjugation.

Procedure for rate study of Ac-GCF with 1o, and N₃-1o. Comparison with IAA

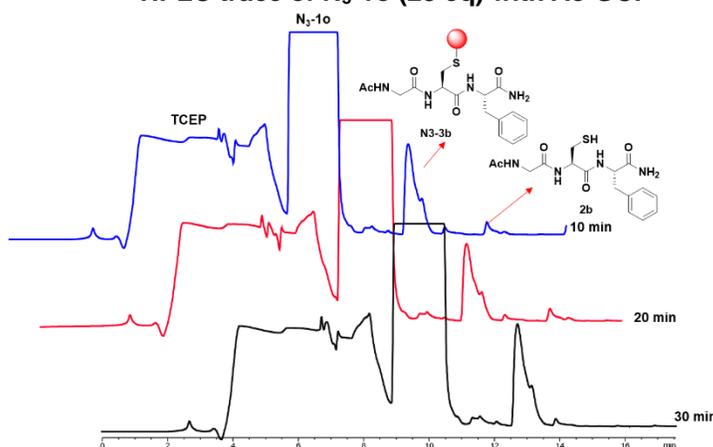
To a solution of Ac-GCF (1.14 mg, 3 mM) in 1 mL of 10 mM NaP (pH 7.5) was added probes (5-25 equiv., 15-75 mM). For time analysis, a sample (100 μ L) was taken from the mixture after regular intervals of time and quenched by freezing the sample at -80 °C. The frozen samples were then lyophilized and dissolved in 100 μ L of 1:1 H₂O/ACN and injected immediately into the HPLC for determining the % conversion to Ac-GCF-1o or Ac-GCF-N₃-1o or Ac-GCF-IAA (X Terra C18 column {5 μ m} with a gradient of 0 to 80% MeCN with 0.1% formic acid in 30 min). The rate study was done in triplet. We use average of three trials to plot the rate curve. 0 min sample is sample taken after addition of all the reagents of the bioconjugate reaction.



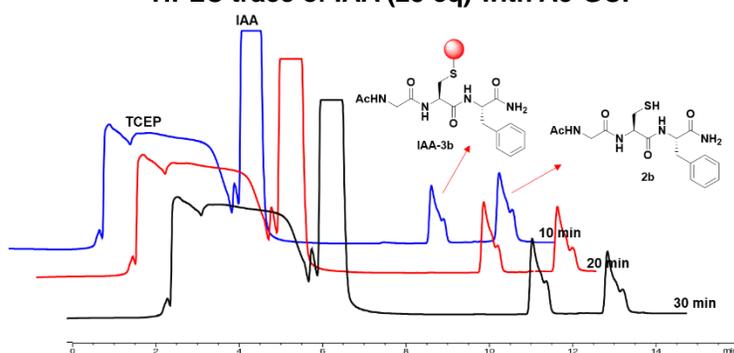
HPLC trace of 1o (25 eq) with Ac-GCF



HPLC trace of N₃-1o (25 eq) with Ac-GCF



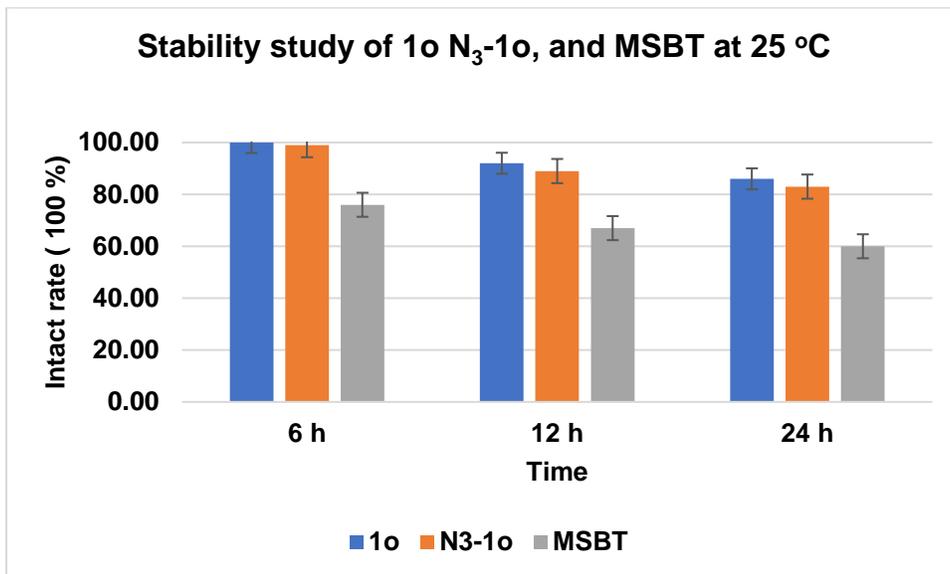
HPLC trace of IAA (25 eq) with Ac-GCF



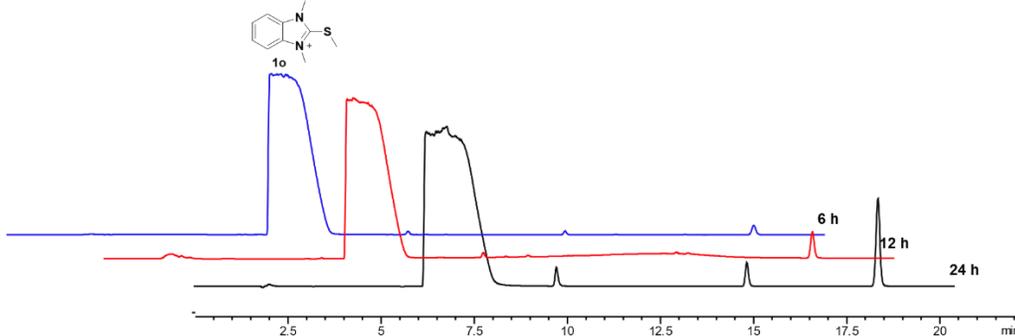
Stability study of 1o, N₃-1o and 1i

Procedure for stability study of probes. Probes **1o**, **N₃-1o** and **1i** (38.75 mM) was incubated in 400 μ L of 10 mM Nap (pH 7.5) at room temperature. A sample (50 μ L) was taken from the mixture and directly injected into HPLC. The reaction was monitored by injecting samples in

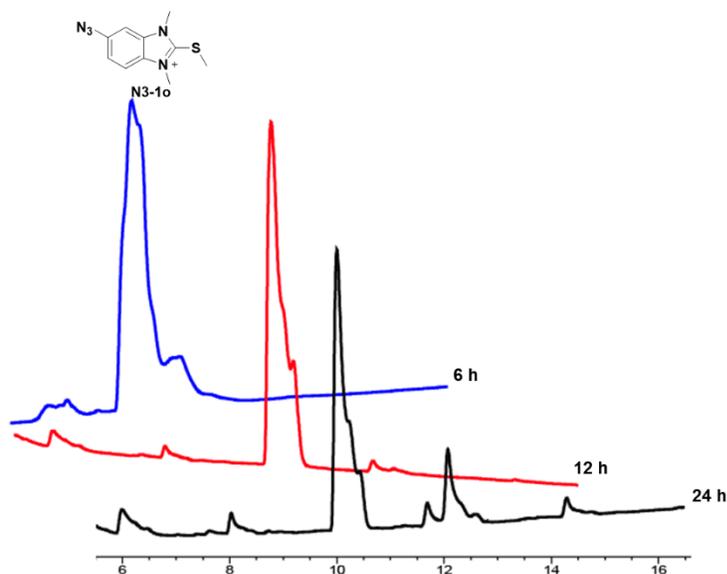
HPLC after regular intervals of time 6 h, 12 h and 24 h.



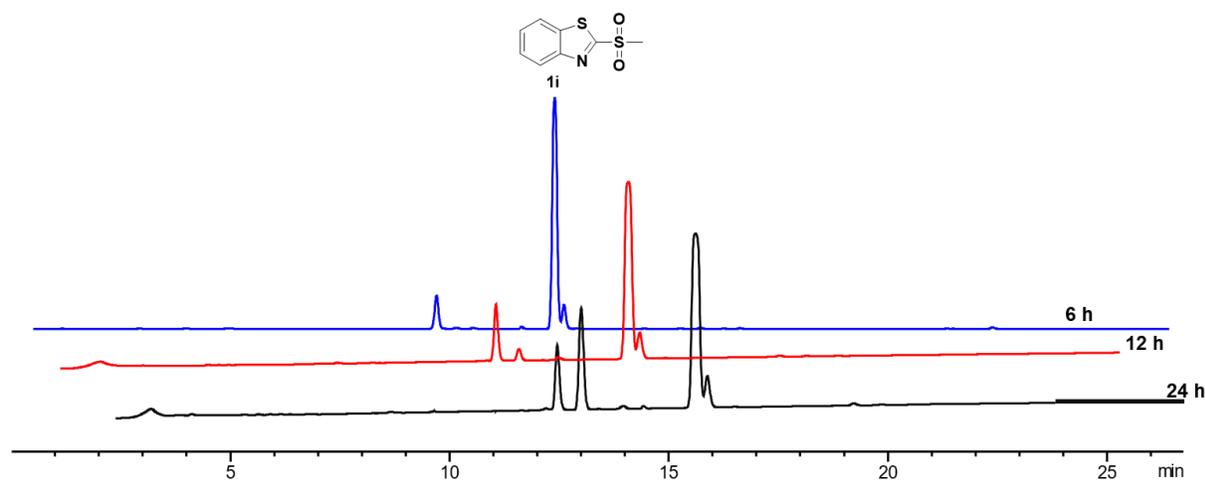
HPLC trace of 1o in Nap pH 7.5 at room temperature



HPLC trace of N3-1o in Nap pH 7.5 at room temperature

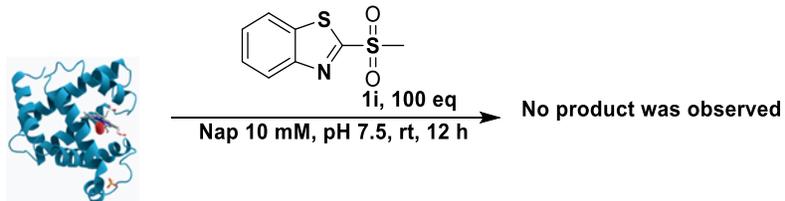
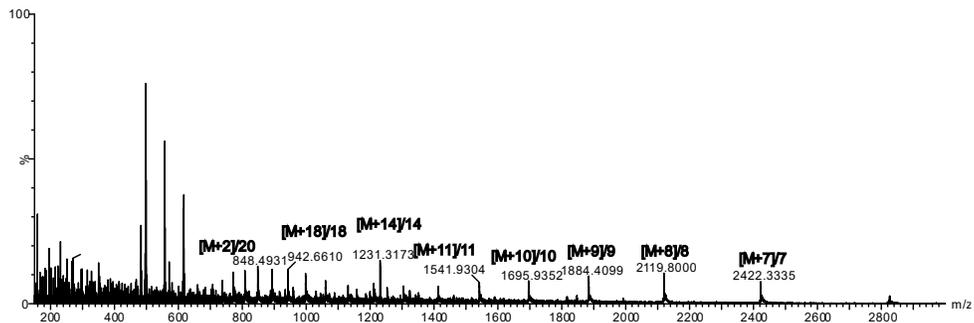
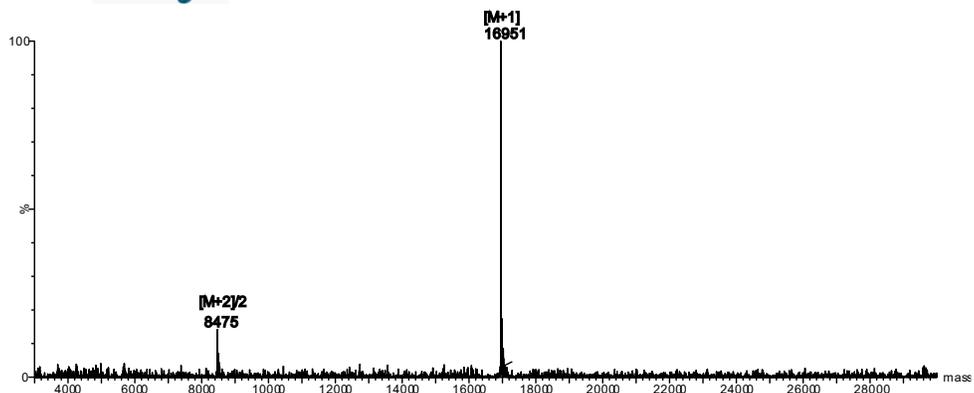
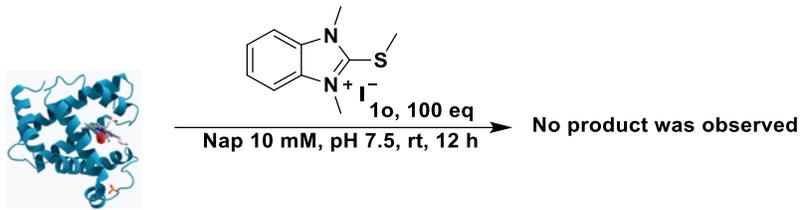


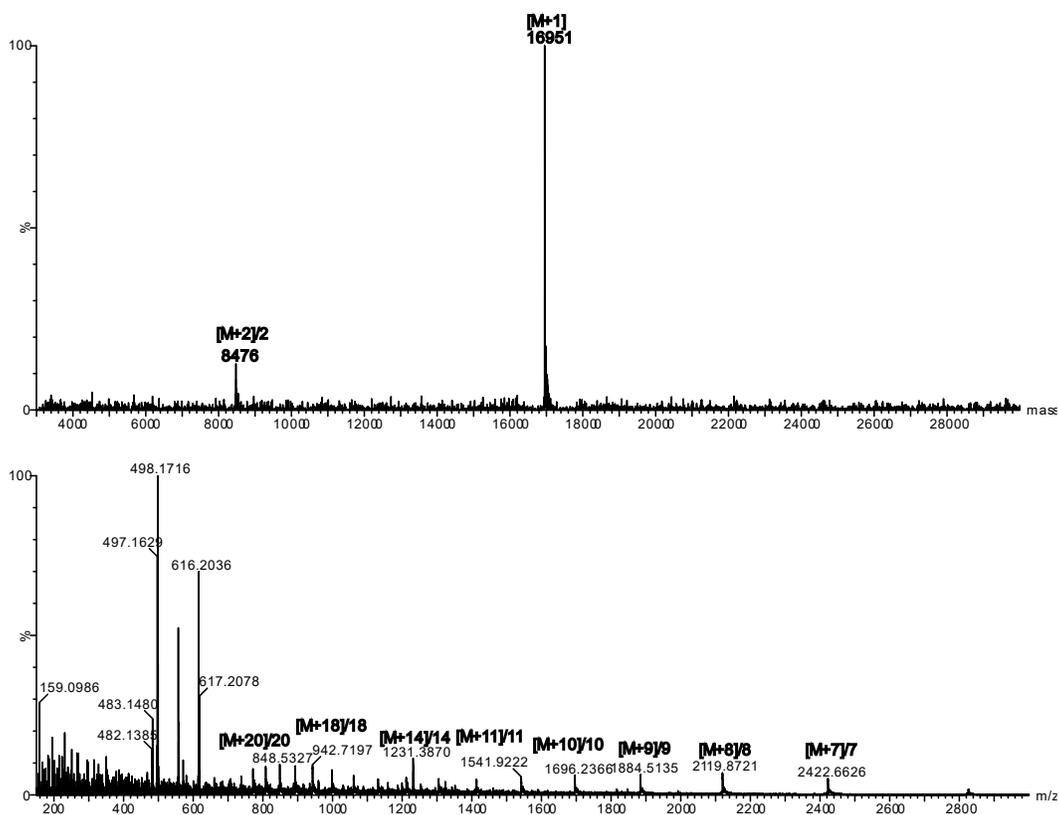
HPLC trace of **1i in Nap pH 7.5 at room temperature**



General procedure for bioconjugation of HAT probes with myoglobin

To a 1 mg of myoglobin (0.15 mM) in 400 μ L 10 mM NaP (pH 7.5), probe **1i**, or **1m** or **1o** (100 equiv. 15 mM) was added. The reaction was incubated at room temperature for 12 h. The reaction mixture was purified by molecular weight cut off and characterized by LCMS.



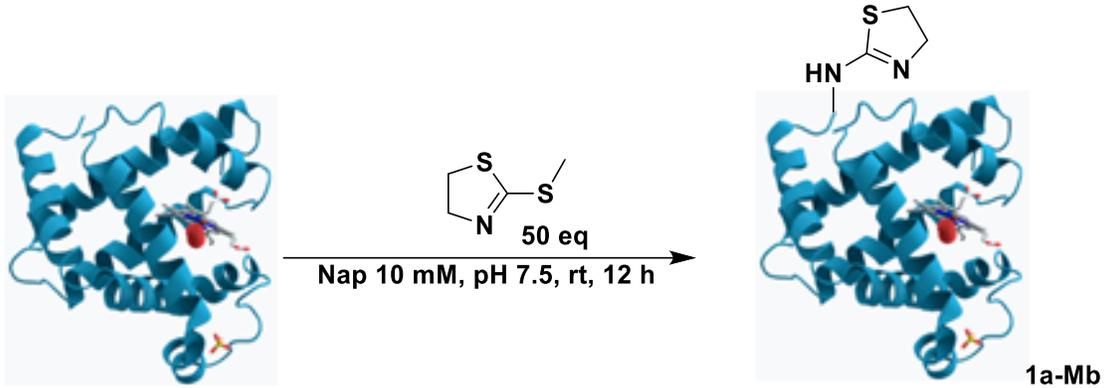


Supplementary Figure 9. Modification of myoglobin Mb with different heteroaromatic azoline compounds 1a-1c, 1j, and 1n

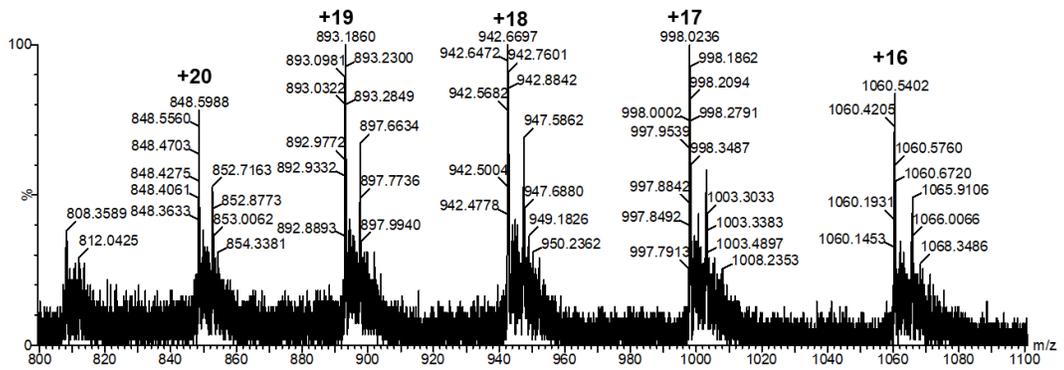
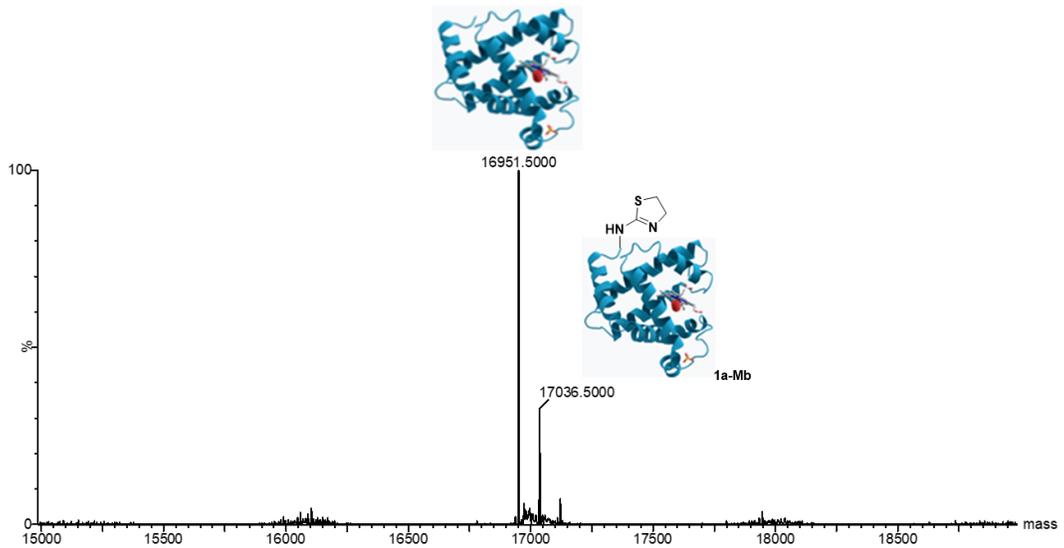
Probe ^a	Eq	Time (h)	Site of modification	Conversion (%) ^b
1a	50	12	N-terminal	33
1b	50	12	Lysine	81
1c	50	12	N-terminal	70
1j	10	8	Lysine	56
1n	10	1	Lysine	99

^aCondition: Protein Mb (60 μ mol, 0.15 mM) in 10 mM phosphate buffer of pH 7.5 (400 μ L) and probe (10-50 equiv.) was incubated for 1-12 h at 25 $^{\circ}$ C. ^bThe conversion was calculated based on the relative peak intensity of native protein and labeled protein in the deconvoluted mass spectrum.

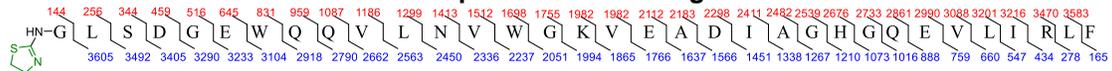
Modification of Mb with 1a

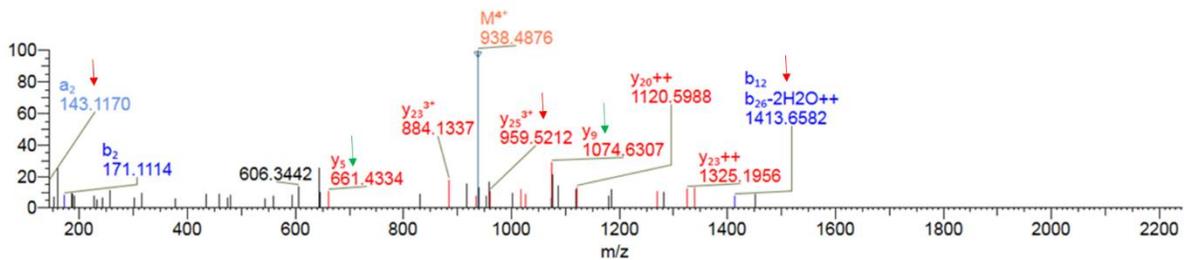


HRMS of 1a-Mb

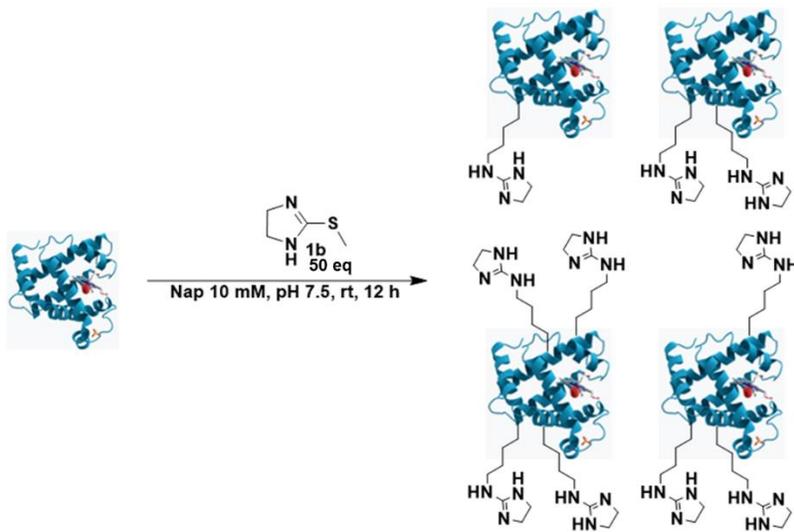


MS/MS spectrum 1a-Mb fragment

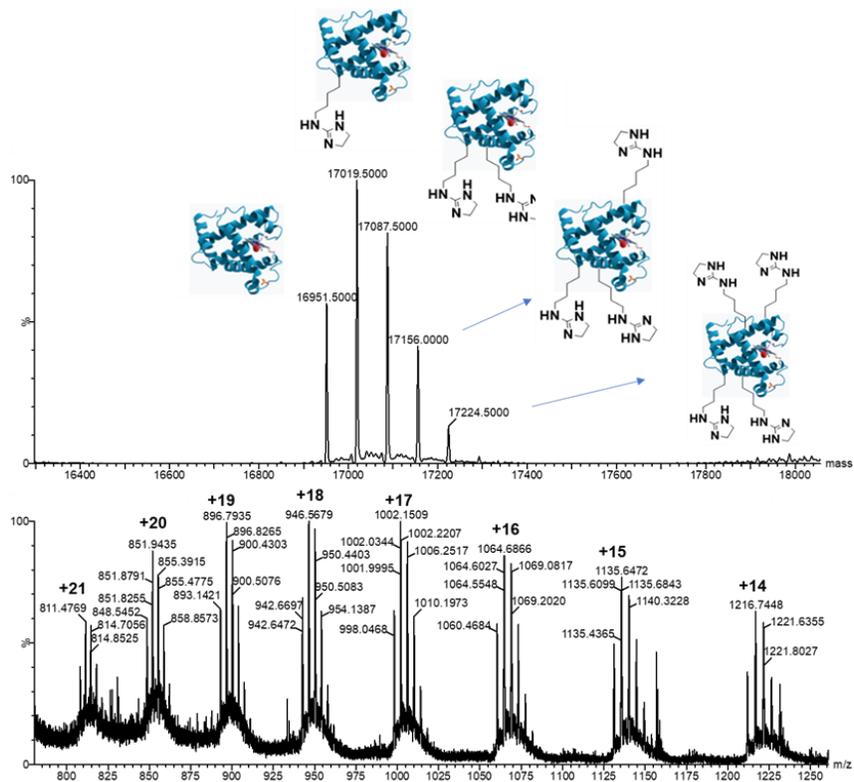




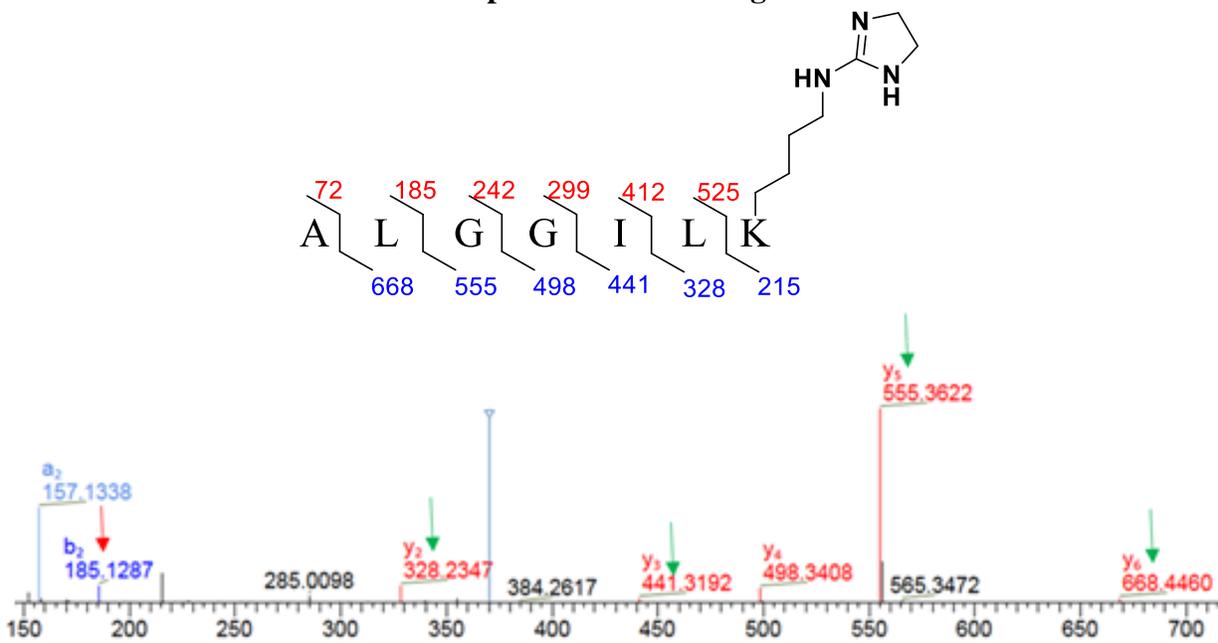
Modification of 1b-Mb



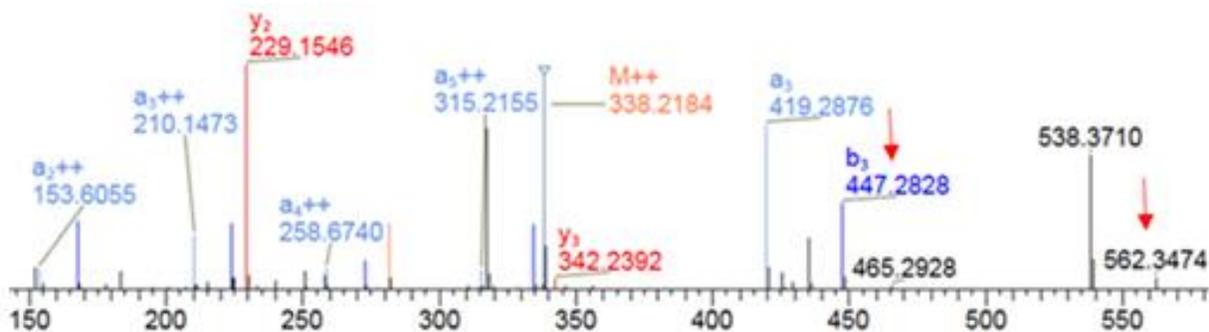
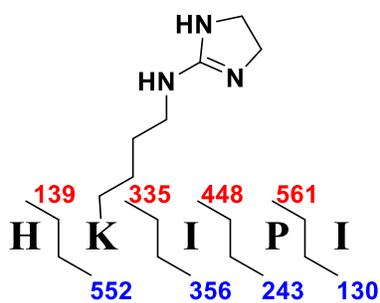
HRMS of 1b-Mb



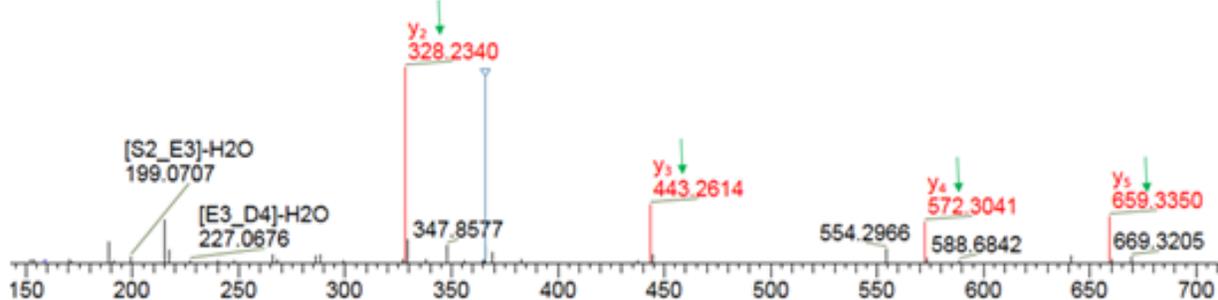
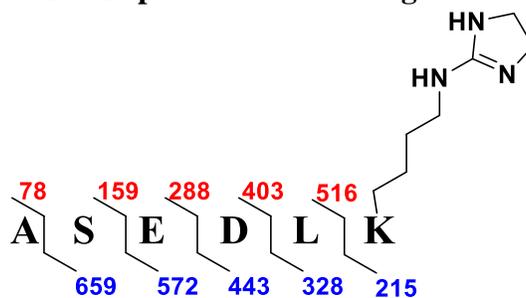
MS/MS spectrum 1b-Mb fragment I



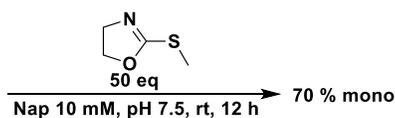
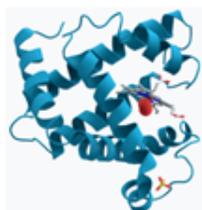
MS/MS spectrum 1b-Mb fragment II



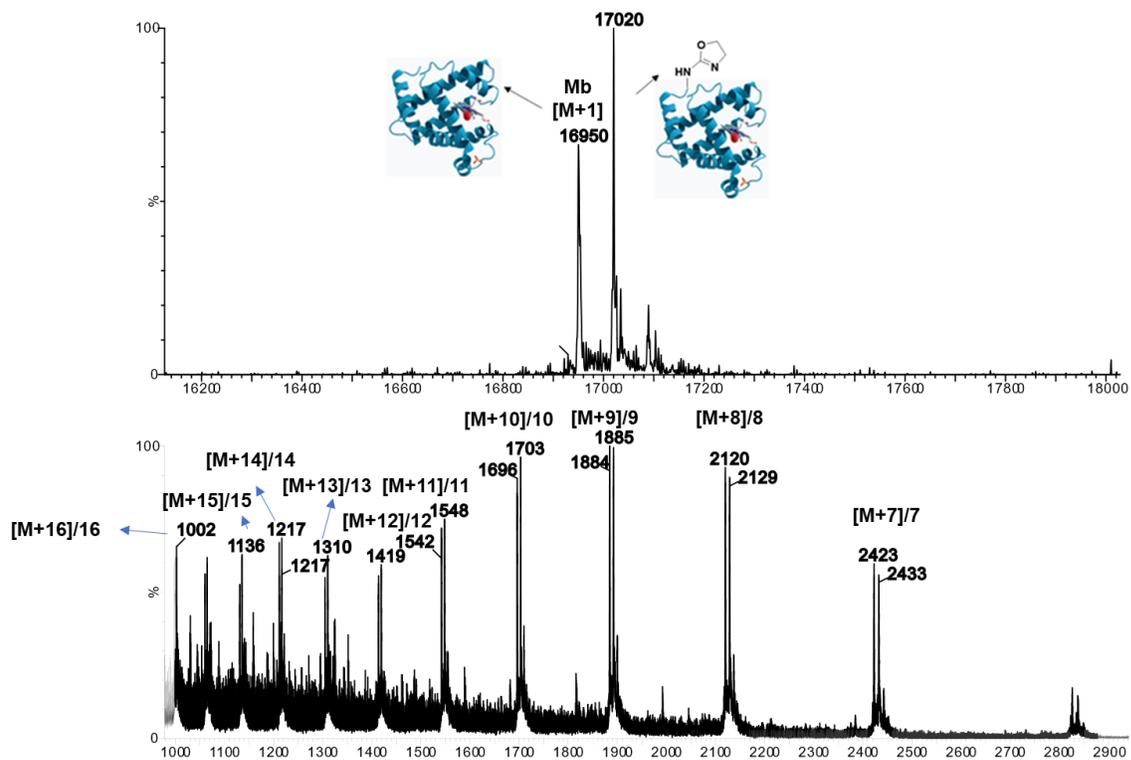
MS/MS spectrum 1b-Mb fragment III



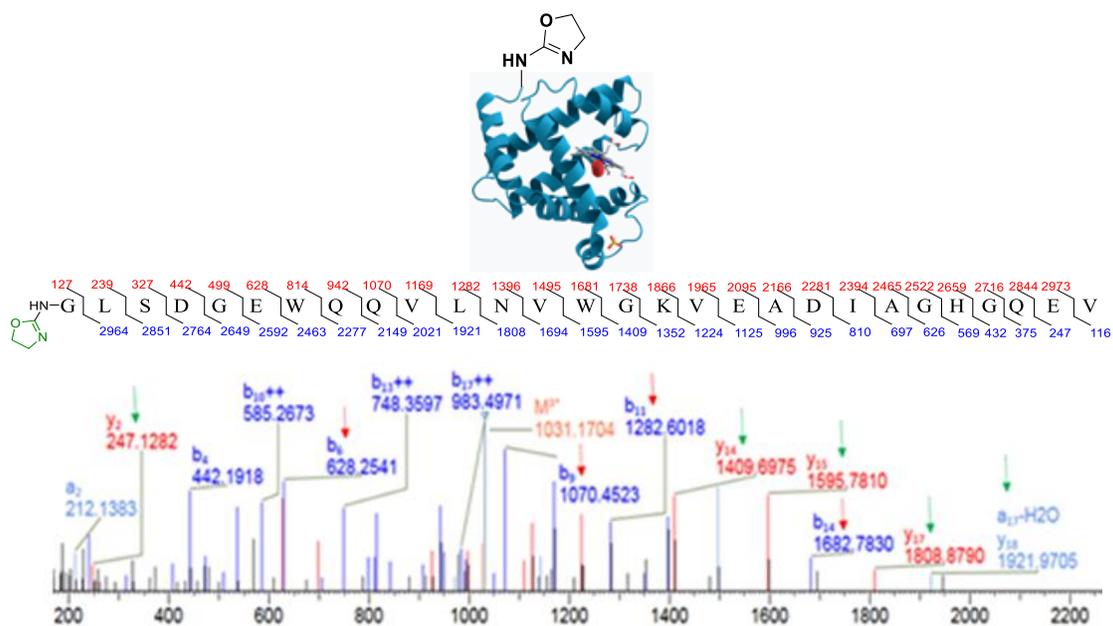
Modification of 1c with Mb



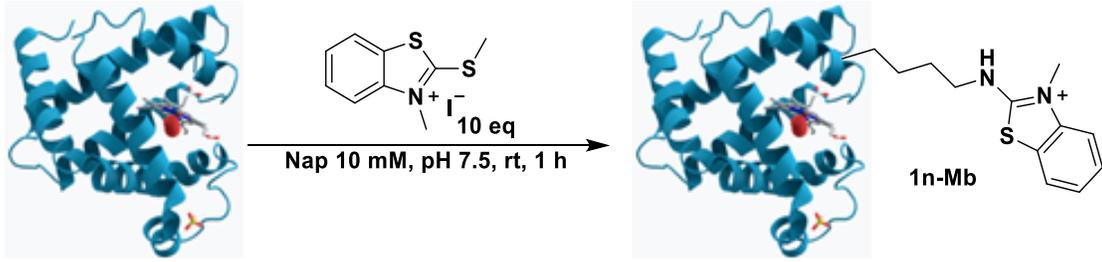
HRMS of 1c-Mb



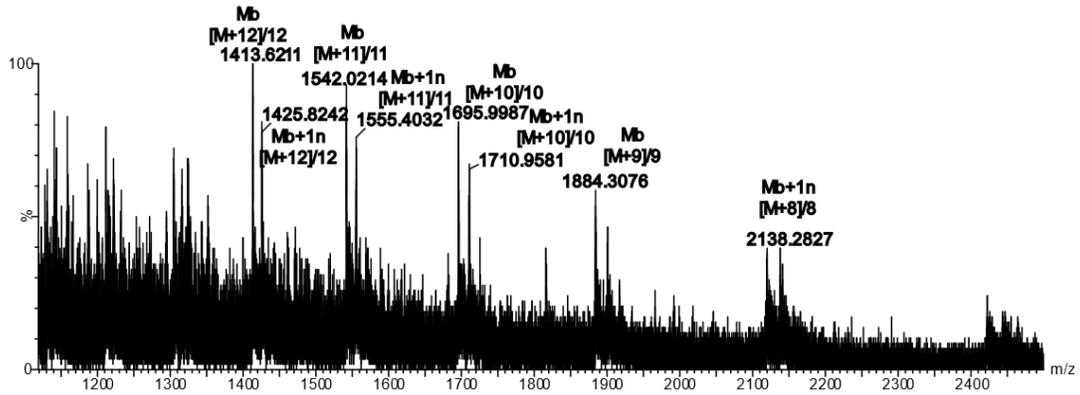
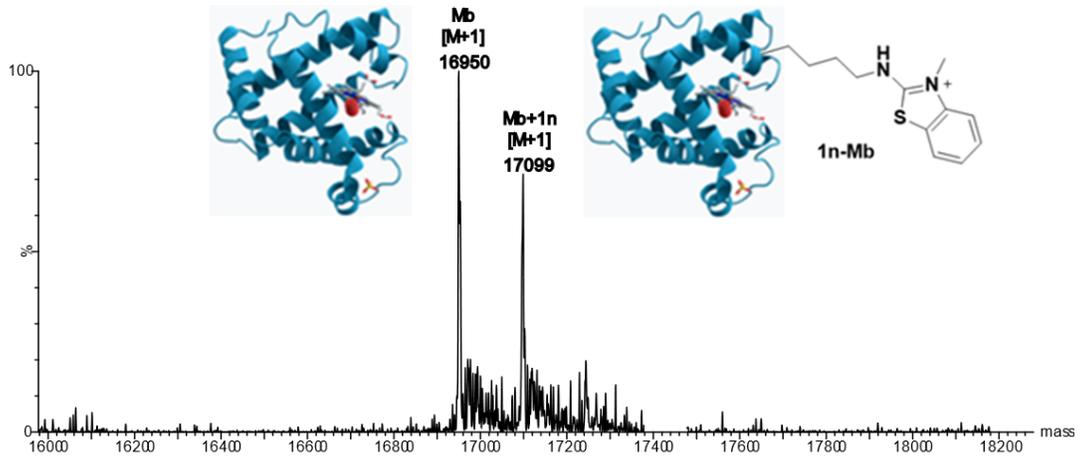
ESI MS-MS spectrum of 1c-Mb



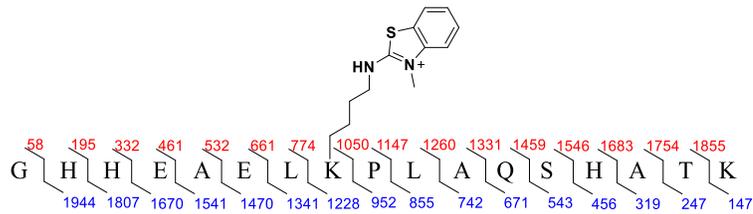
Modification of Mb with 1n

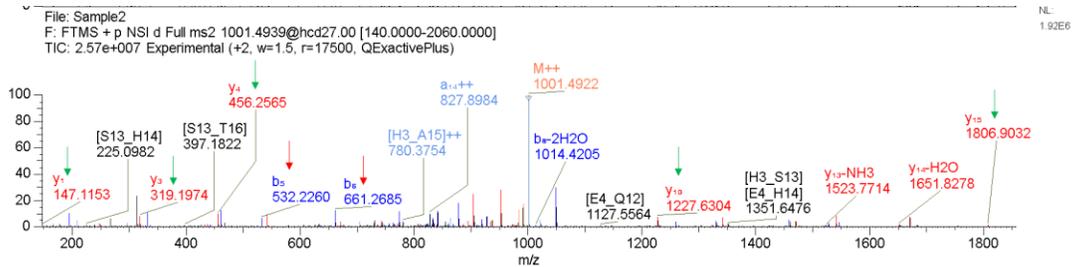


HRMS of 1n-Mb

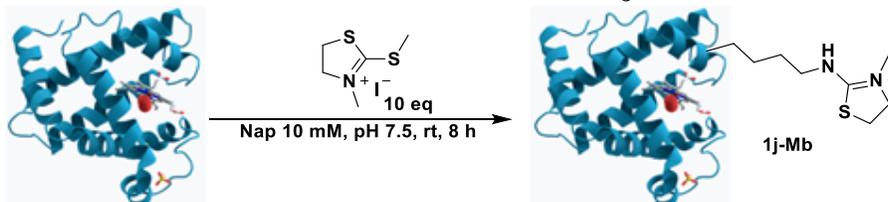


MS/MS spectrum 1n-Mb fragment

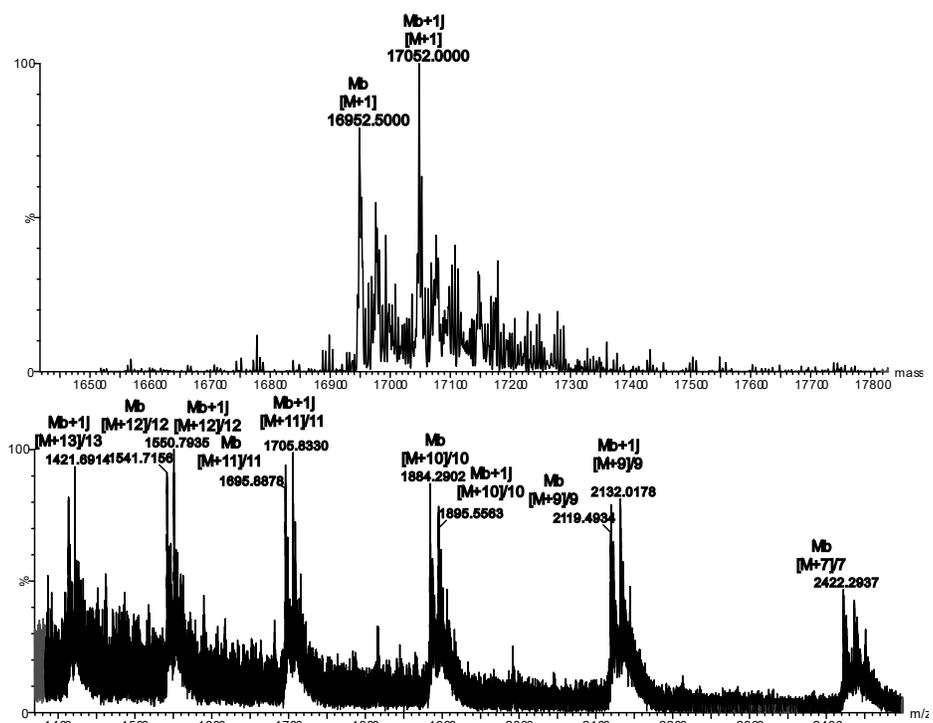




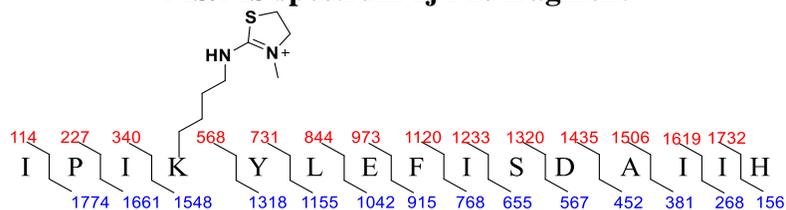
Modification of Mb with 1j

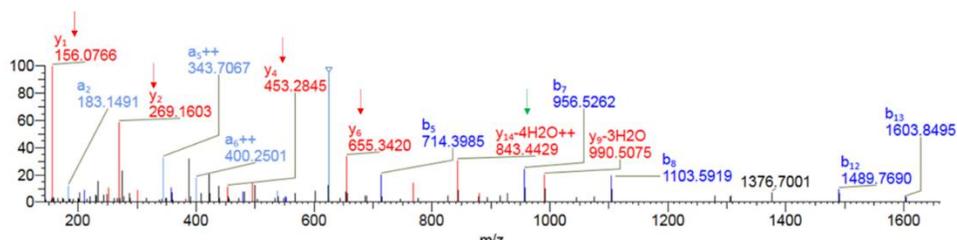


HRMS of 1j-Mb



MS/MS spectrum 1j-Mb fragment





Supplementary Figure 10. Selective cysteine bioconjugation of Insulin a chain and b chain with compound **1o**

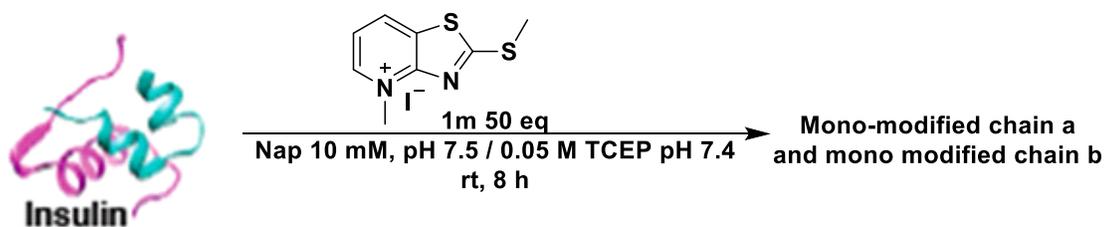
Procedure for labeling of the reduced insulin.

To 0.35 mg of Insulin (60 μ mol, 0.15 mM) in 400 μ L 10 mM Nap (pH 7.5), 20 μ L of TCEP solution (pH 7.5, 50 mM) was added to reduce the insulin. The mixture was incubated at room temperature for 20 min. Probe **1o** (50 equiv.) was added into this mixture and the reaction was allowed to react at room temperature for 8 h. The **1o-chain a**, and **1o-chain b** bioconjugates were characterized by LCMS.

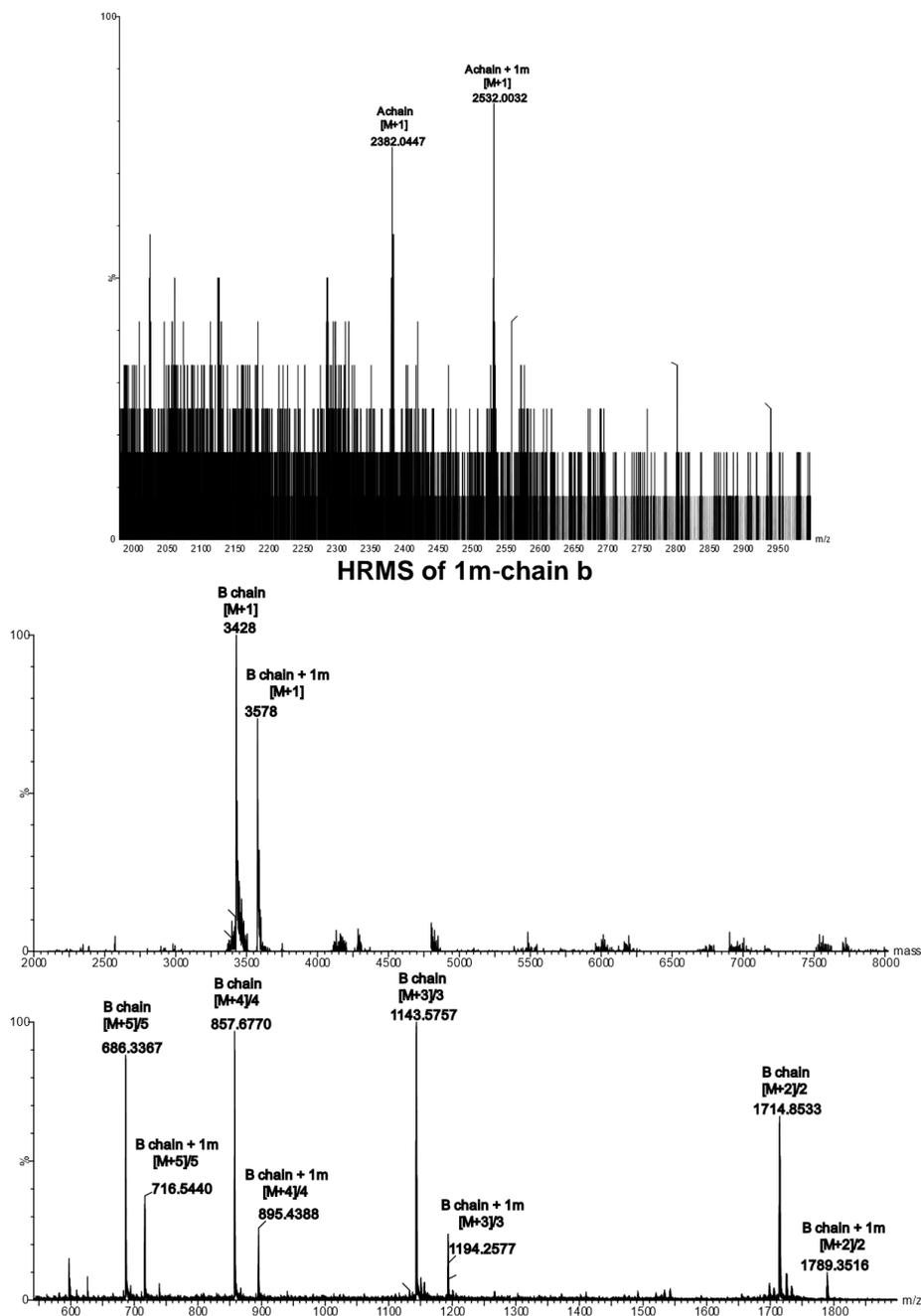
Supplementary Figure 11. Modification of reduced insulin with **1m**

Procedure for labeling of the reduced insulin with **1m**

To 0.35 mg of Insulin (60 μ mol, 0.15 mM) in 400 μ L 10 mM Nap (pH 7.5), 20 μ L of TCEP solution (pH 7.5, 50 mM) was added to reduce the insulin. The mixture was incubated at room temperature for 20 min. Probe **1m** (50 equiv.) was added into this mixture and the reaction was allowed to react at room temperature for 8 h. The **1m-chain a**, and **1m-chain b** bioconjugates were characterized by LCMS.



HRMS of **1m-chain a**

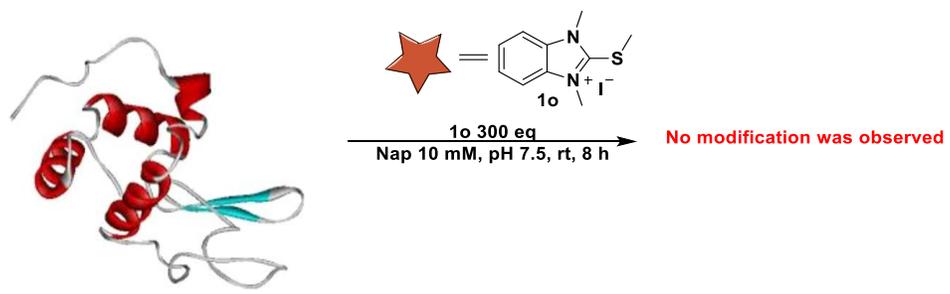
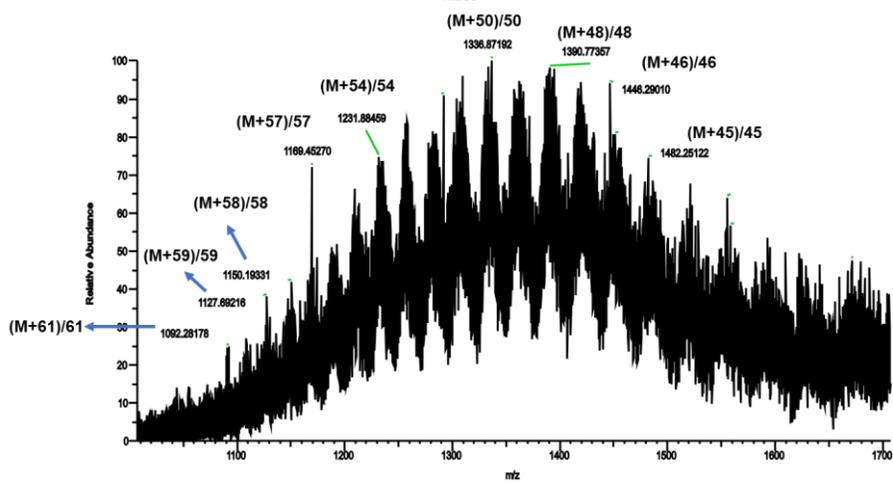
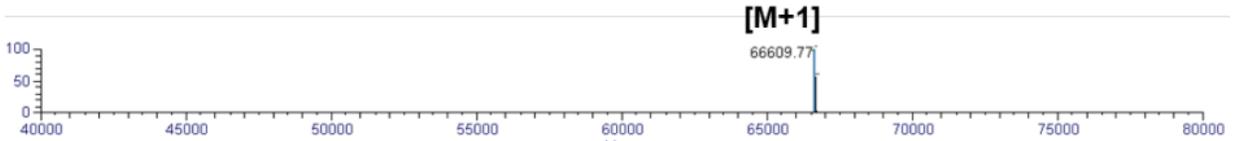
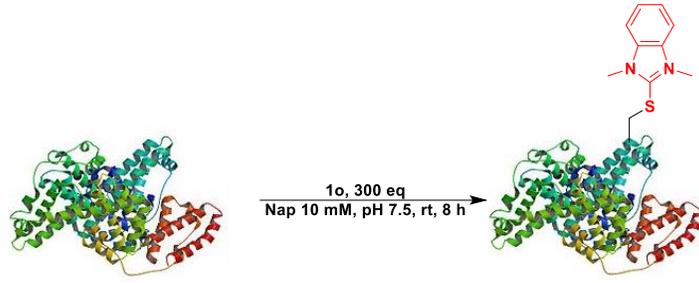


Supplementary Figure 12. Modification of bovine serum albumin (BSA) and lysozyme with HAT probe 1o

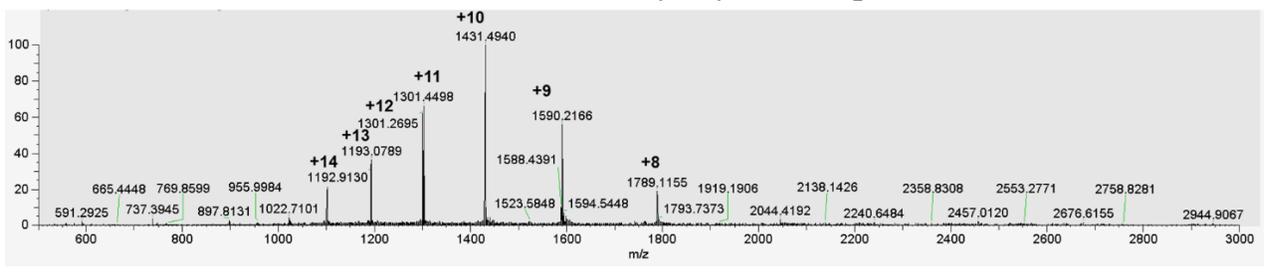
Procedure for the labeling of commercial proteins with compound 1o

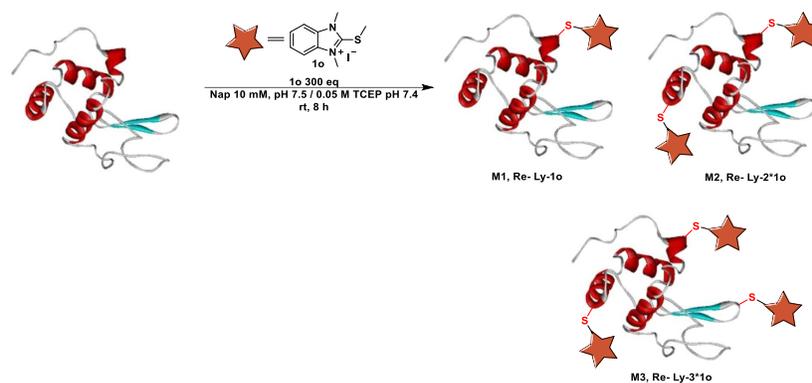
To a mixture of protein (60 μ mol, 0.15 mM) in 400 μ L 10 mM Nap (pH 7.5), probe **1o** (5.76 mg, 18 mmol, 300 equiv.) was added. The reaction mixture was incubated at room temperature for 8

h. The crude was analyzed by LCMS.

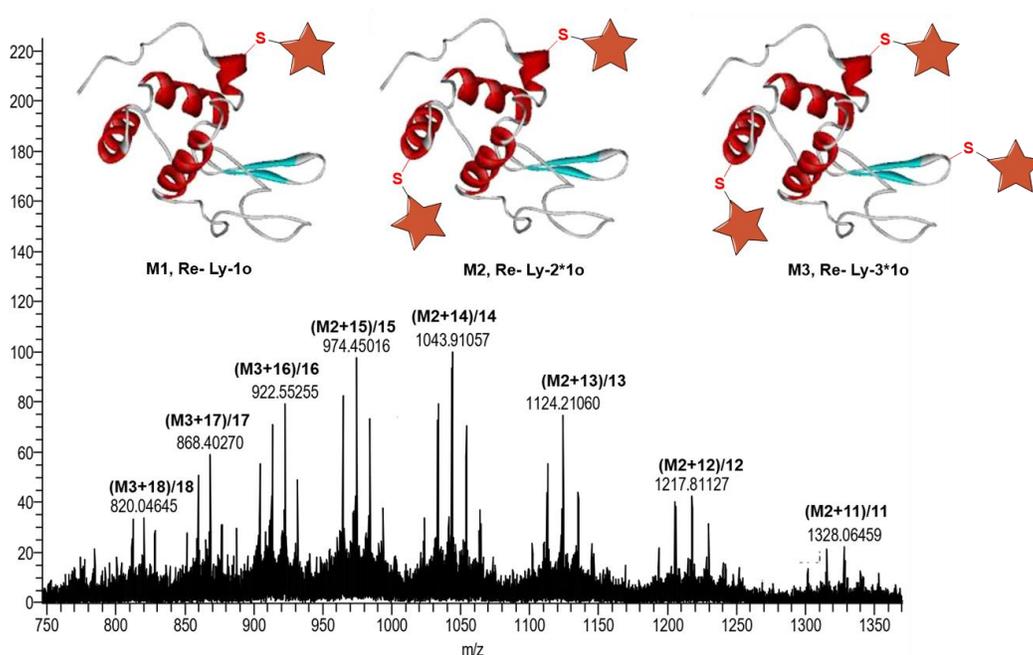


HRMS of unreduced lysozyme example





HRMS of modified lysozyme



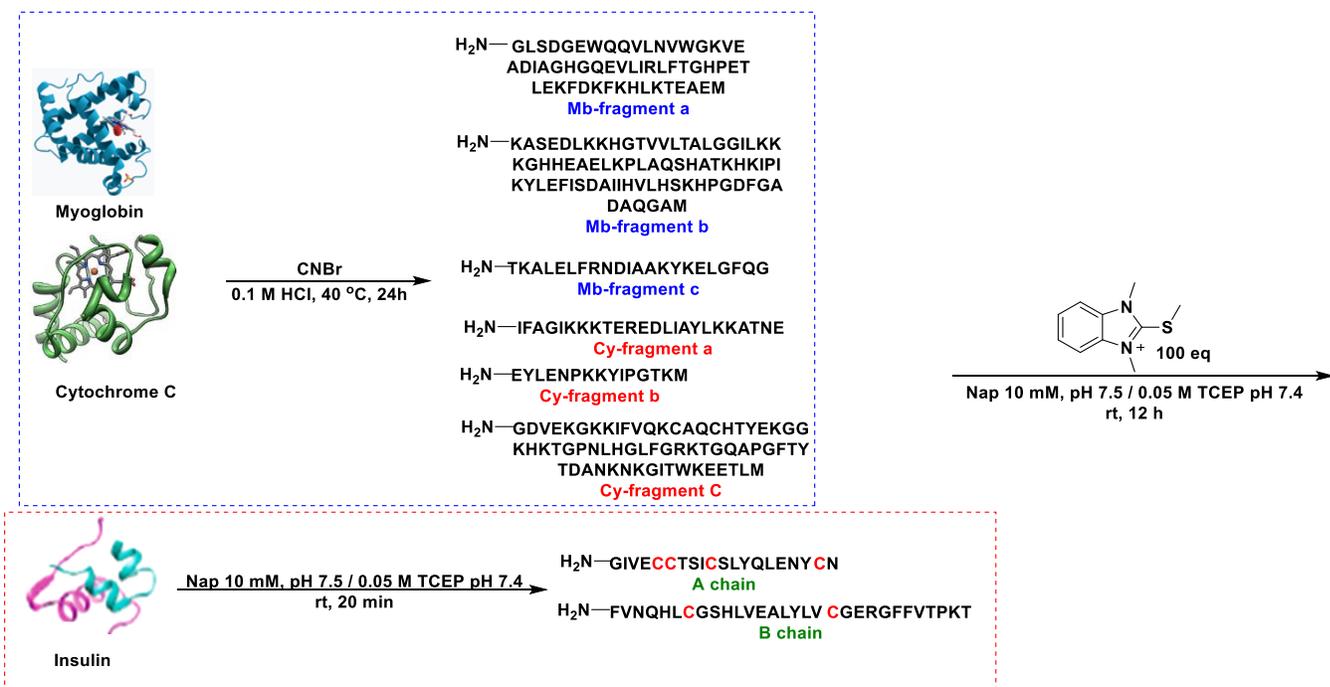
Supplementary Figure 13. Enrichment of cysteine containing peptides with 1o in mixture of proteolytic fragments.

Procedure for the digestion of cytochrome C and Myoglobin by CNBr⁴

Proteins (81 μ mol, 0.25 mM) and 0.3 mg of CNBr (0.1875 mM) were mixed in 324 μ L 0.1 M HCl, the reaction mixture was incubated at 40 $^{\circ}$ C for 24 h. The reaction mixture was quenched by freezing the sample at -80 $^{\circ}$ C. The frozen samples were then lyophilized to afford the dry peptide fragments.

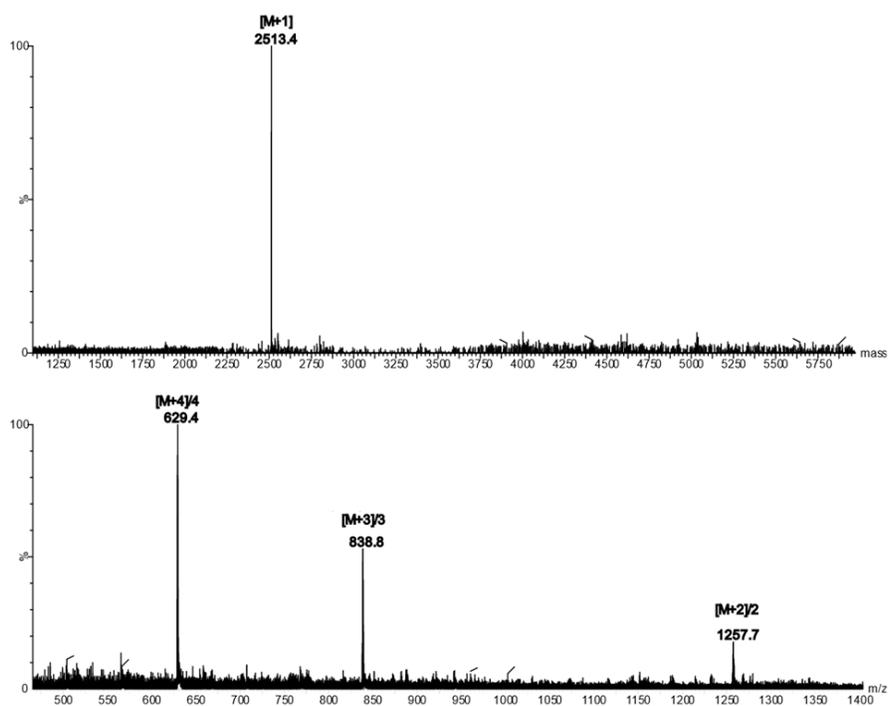
Procedure of enrichment of cysteine fragments by compound 1o:

To the mixture of proteolytic fragments (0.81 μmol of cytochrome C and myoglobin) in 0.4 mL of 10 mM sodium phosphate buffer (Nap, pH 7.5), insulin (81 μmol , 0.25 mM) in 10 mM Nap (400 μl , pH 7.5) and 20 μL of TCEP solution (pH 7.4, 50 mM) was added to generate reduced insulin, the mixture was incubated at room temperature for 20 min. Probe **1o** (100 eq) was added in the mixture of proteolytic fragments at room temperature and reaction was left for 12 h. The tagged proteolytic fragments were analyzed by LCMS without purification.

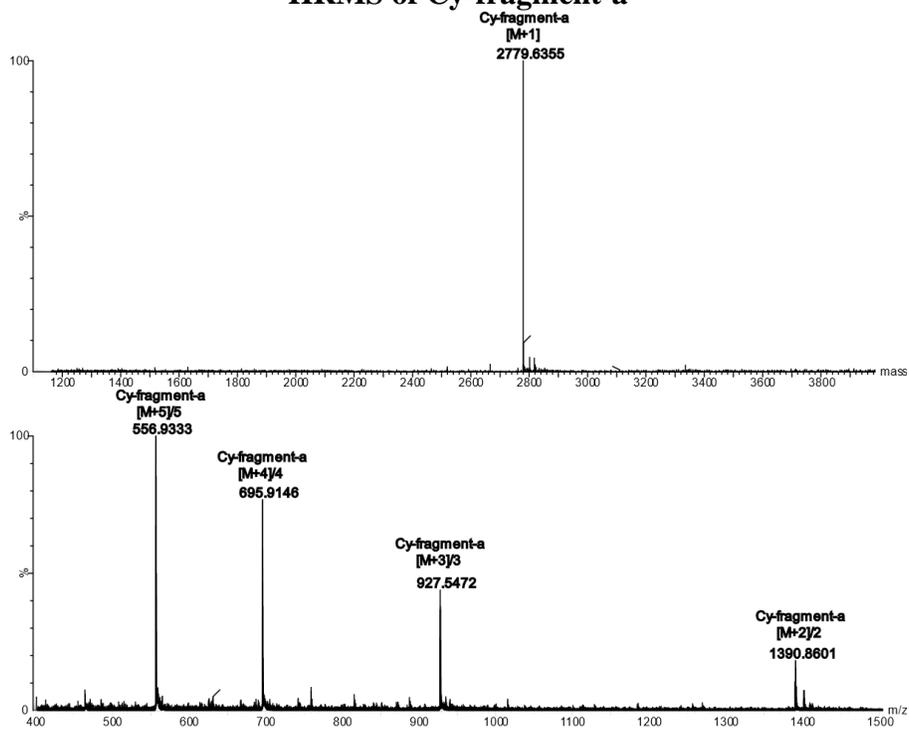


HRMS of Mb-fragment-c

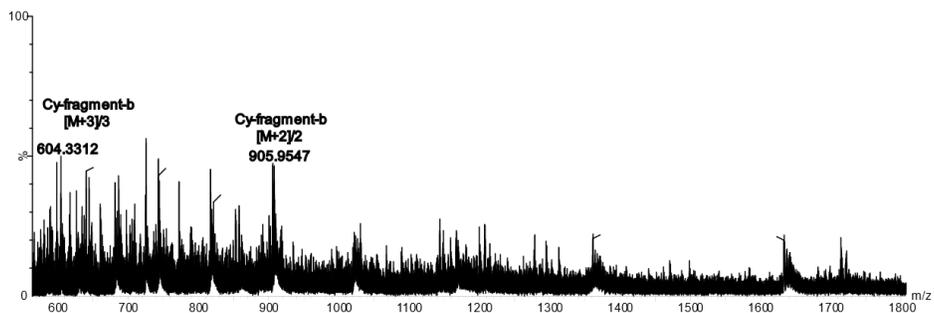
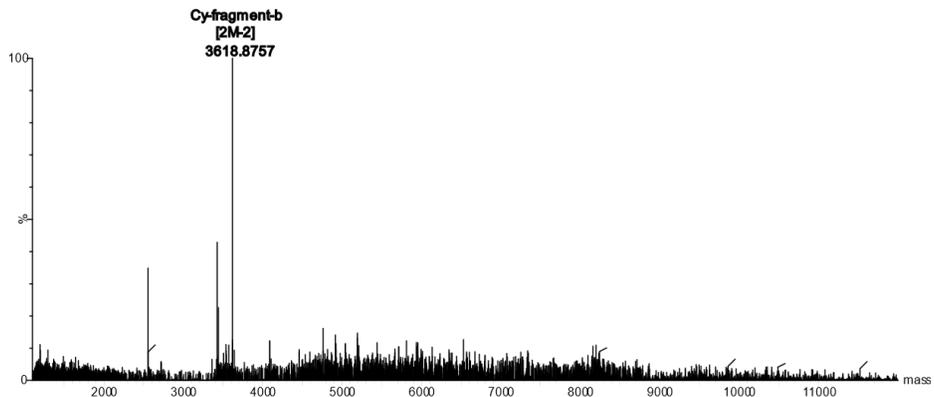
Mb-fragment c



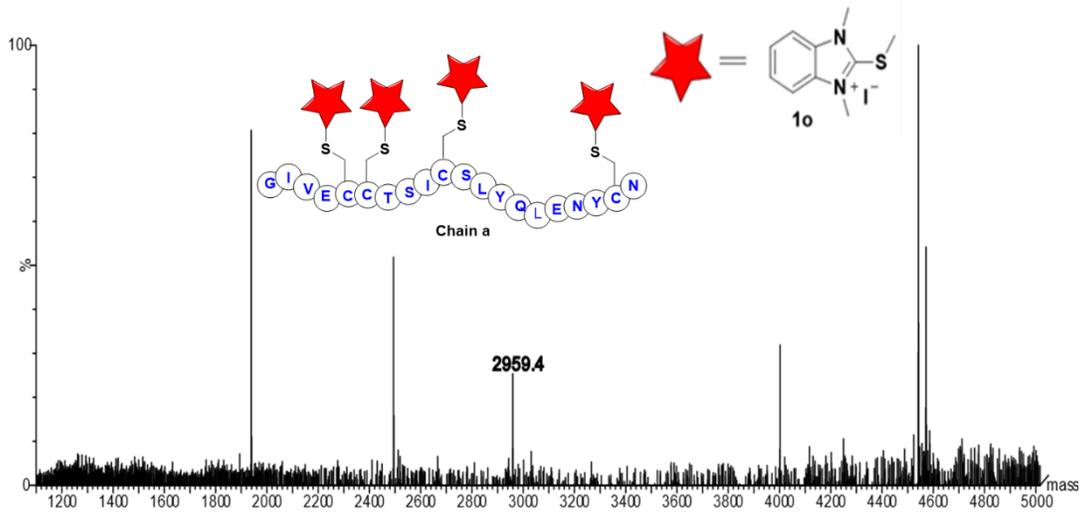
HRMS of Cy-fragment-a



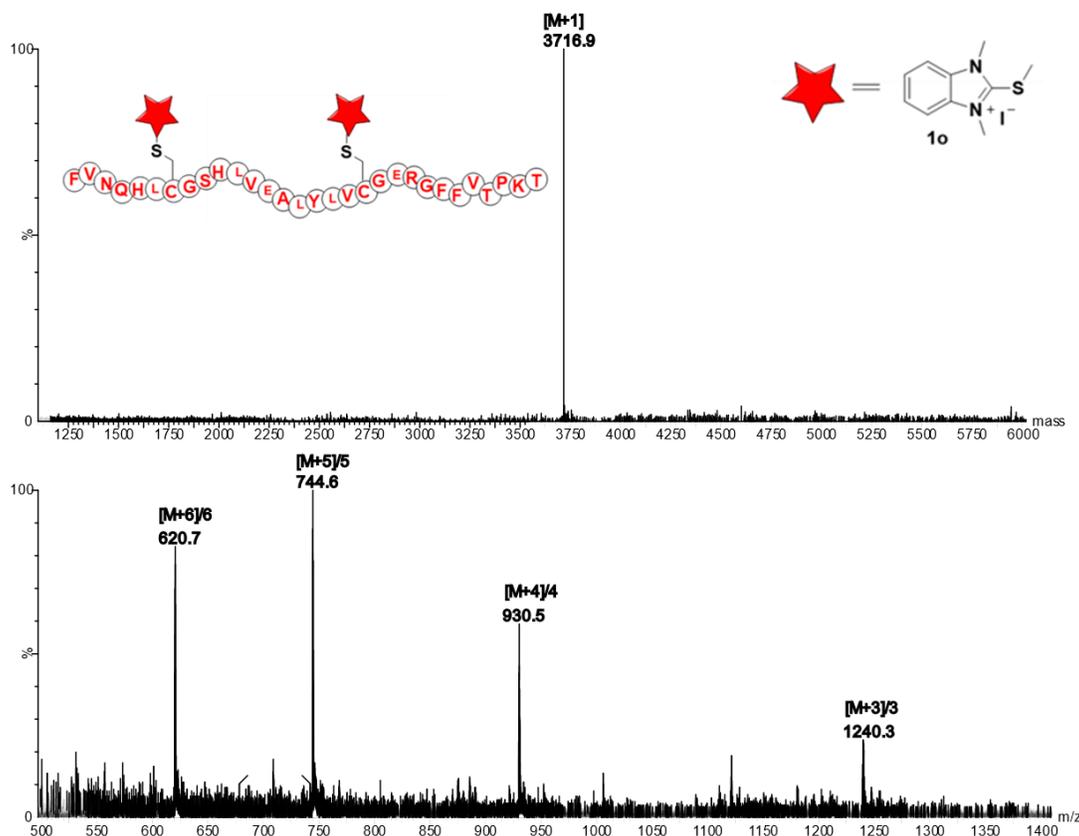
HRMS of Cy-fragment-b



HRMS of 1o-insulin chain a



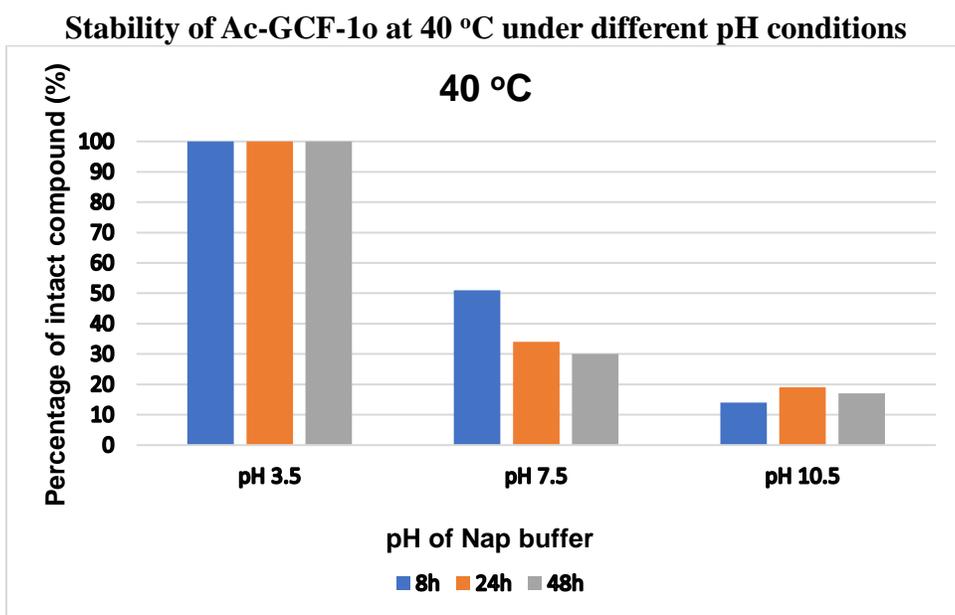
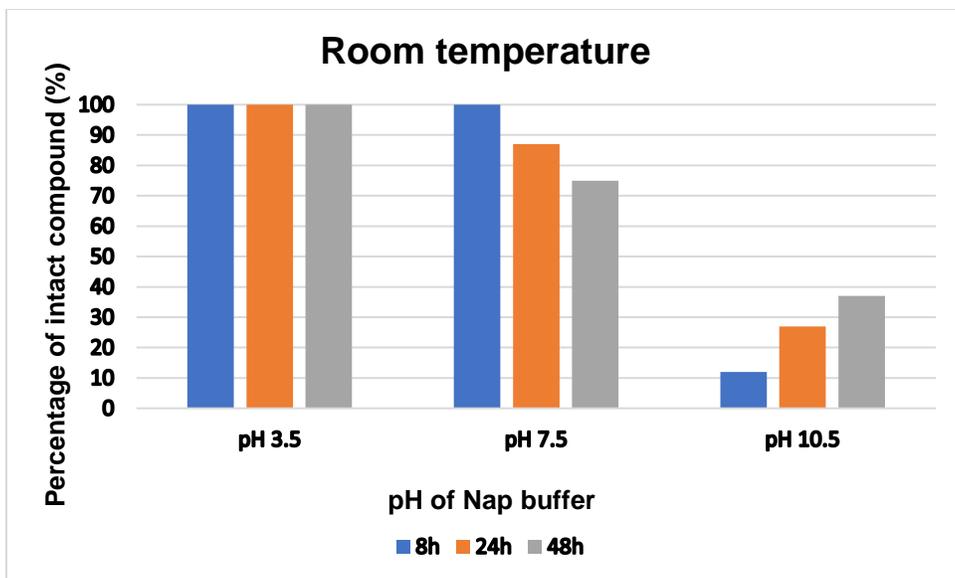
HRMS of 1o-insulin chain b



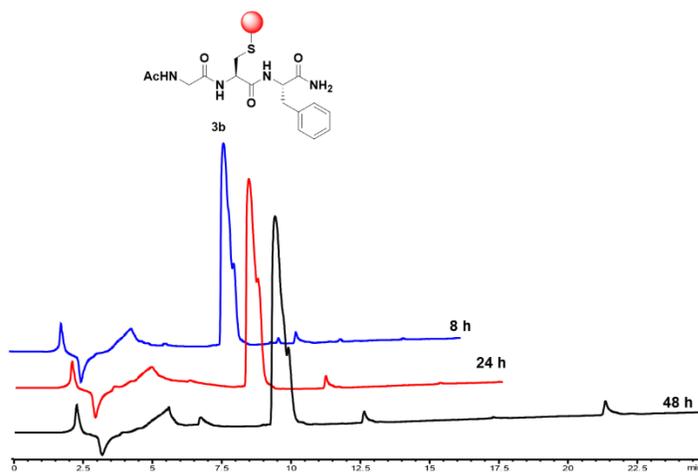
Supplementary Figure 14. Stability study of Ac-GCF-1o under different pH conditions.

Ac-GCF-1o conjugate **3b** (3.75 mM) was incubated in 10 mM Nap at different pH ranging from 3.5 to 10.5 at room temperature and at 40 °C. The samples (50 μ L) were taken from the reaction mixtures and directly injected into the HPLC. The reactions were monitored by injecting samples in HPLC after regular intervals of time 8h, 24h and 48 h. The bioconjugate product showed high stability under pH 3.5 at both room temperature and 40 °C. We observed high stability of the conjugate for 24h under physiological conditions (pH 7.5) at room temperature.

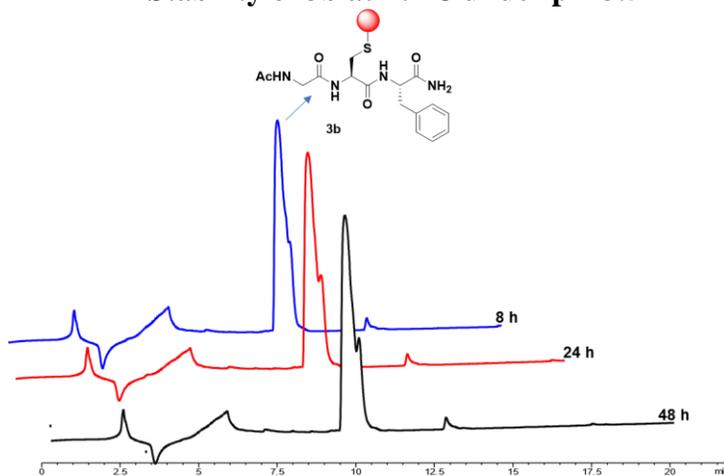
Stability of Ac-GCF-1o at 25 °C under different pH conditions



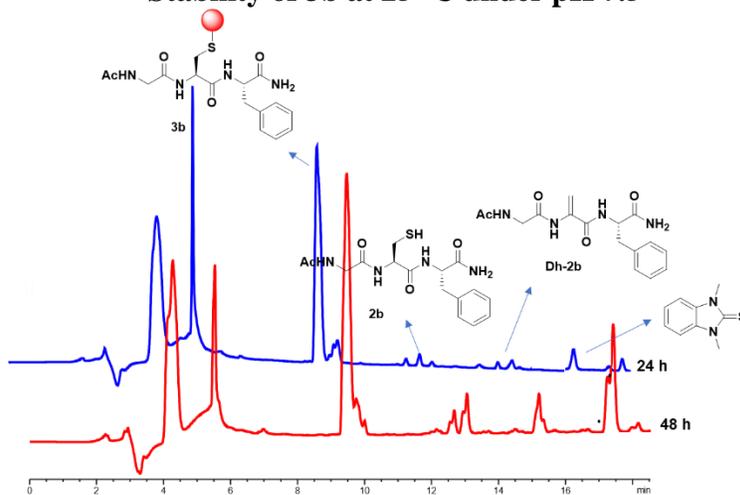
Stability of 3b at 25 °C under pH 3.5



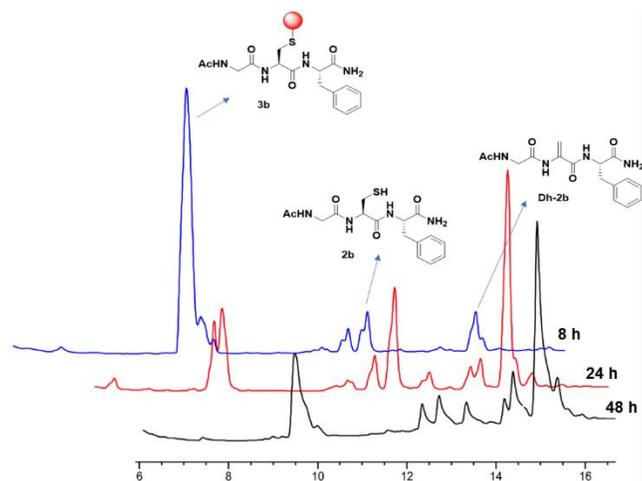
Stability of 3b at 40 °C under pH 3.5



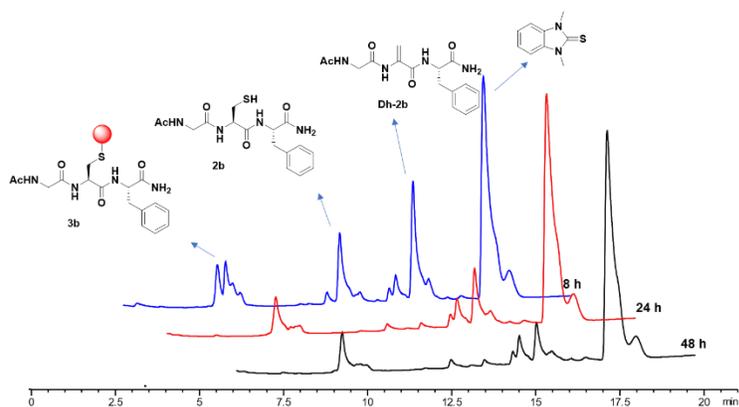
Stability of 3b at 25 °C under pH 7.5



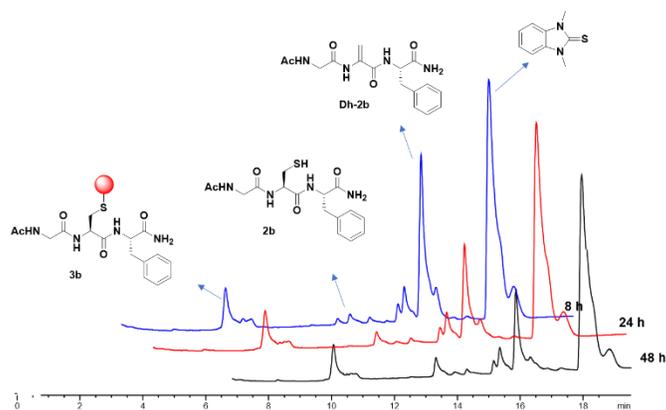
Stability of 3b at 40 °C under pH 7.5



Stability of 3b at 25 °C under pH 10.5



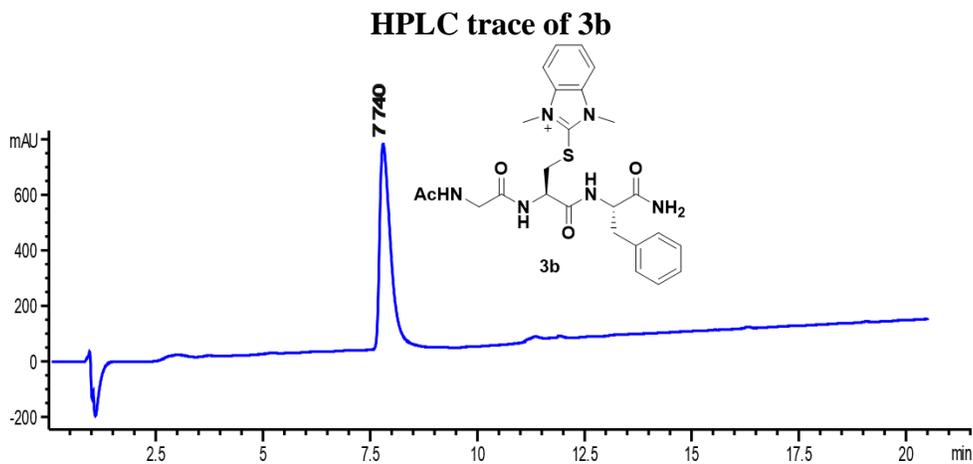
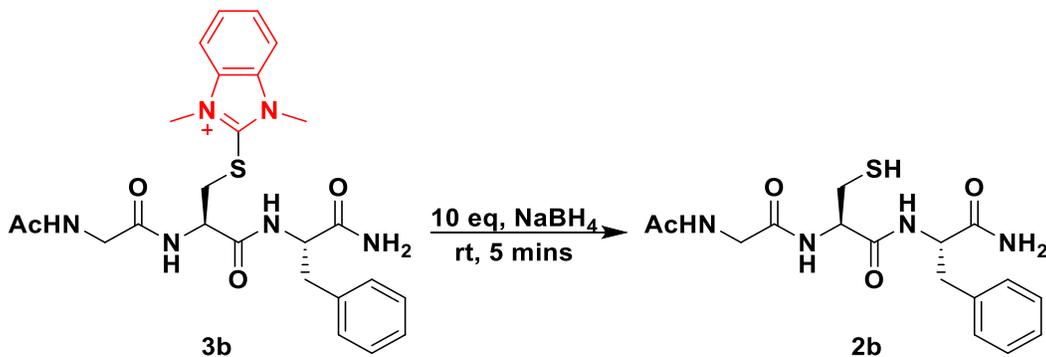
Stability of 3b at 40 °C under pH 10.5



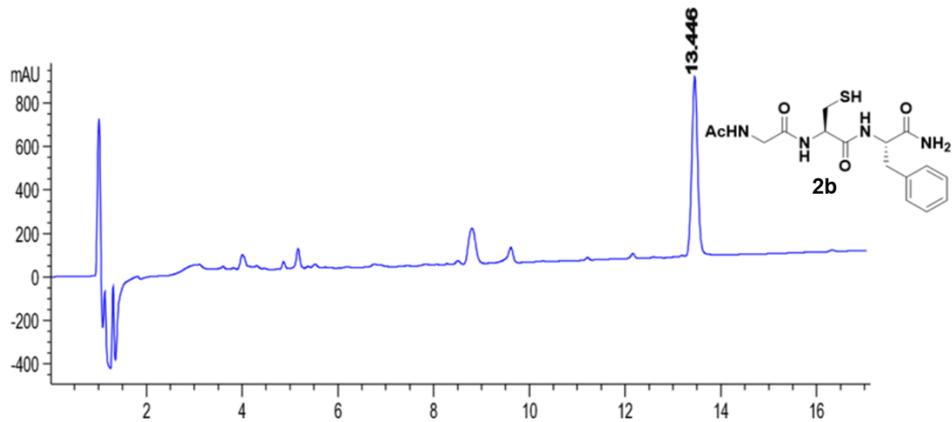
Supplementary Figure 15. Reversible study of Cys-HAT biconjugate with NaBH₄

Procedure for reversible study of 3b by NaBH₄

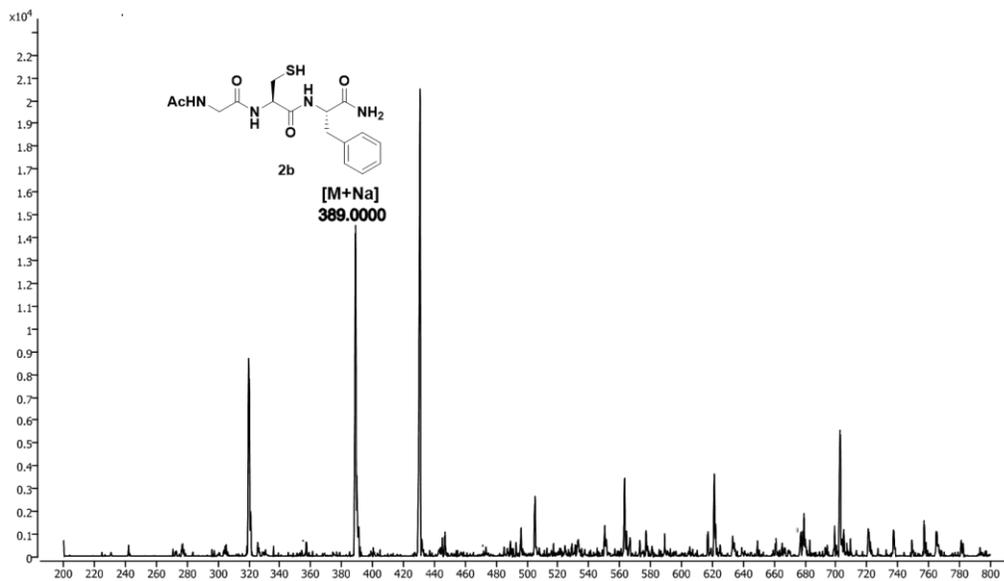
In a solution of peptide conjugate **3b** (6.3 mM) in 0.4 mL of 10 mM Nap (pH 7.5) was added 10 equiv. of NaBH₄ (63 mM). The mixture was stirred at room temperature for 5 min. The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



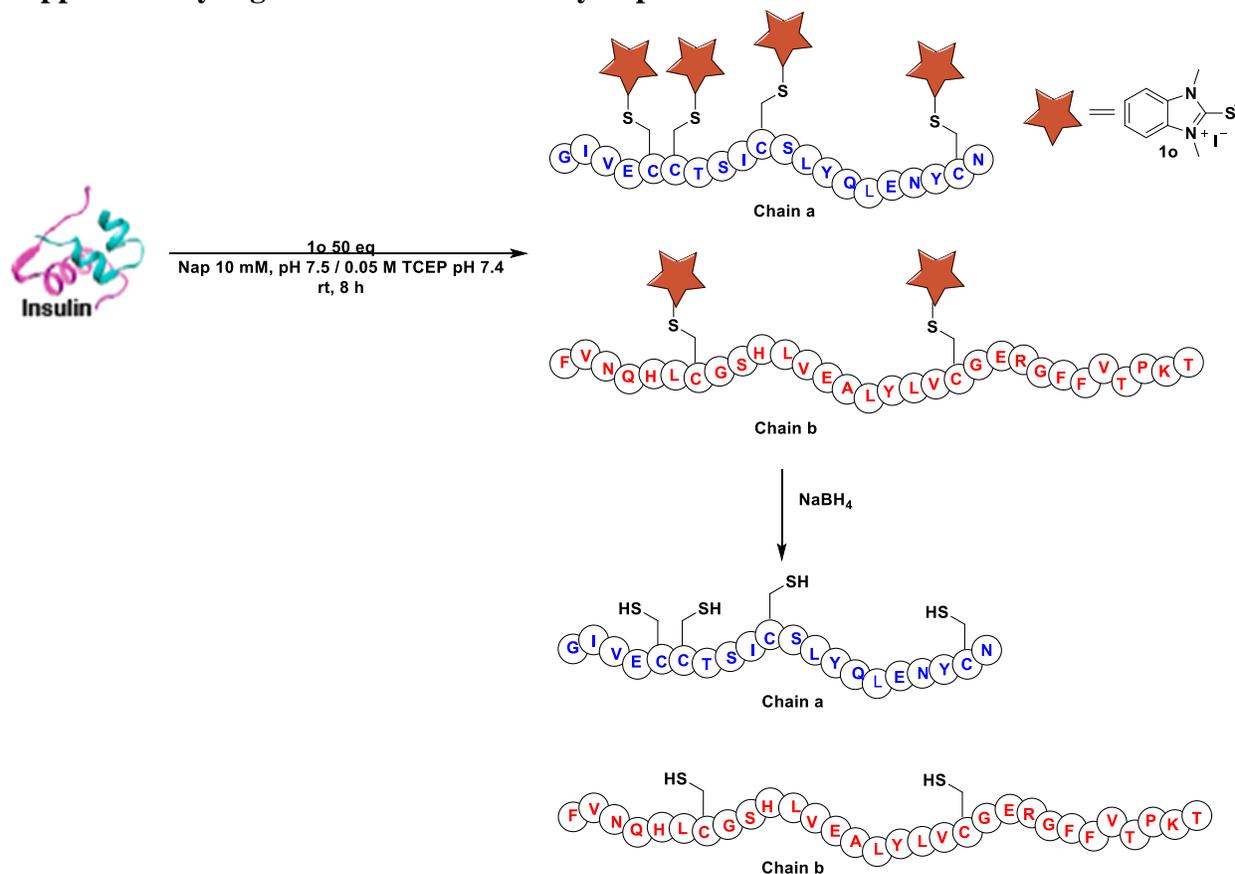
HPLC trace after addition of NaBH₄ for 5 min



ESI-MS after addition of NaBH₄



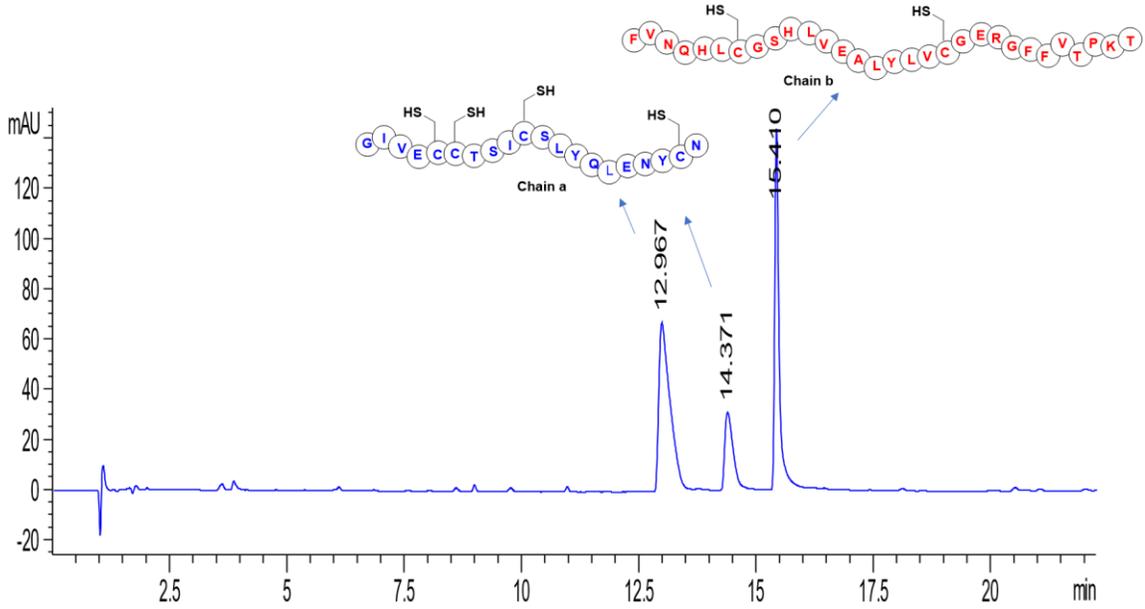
Supplementary Figure 15. Reversible study of probe 1o



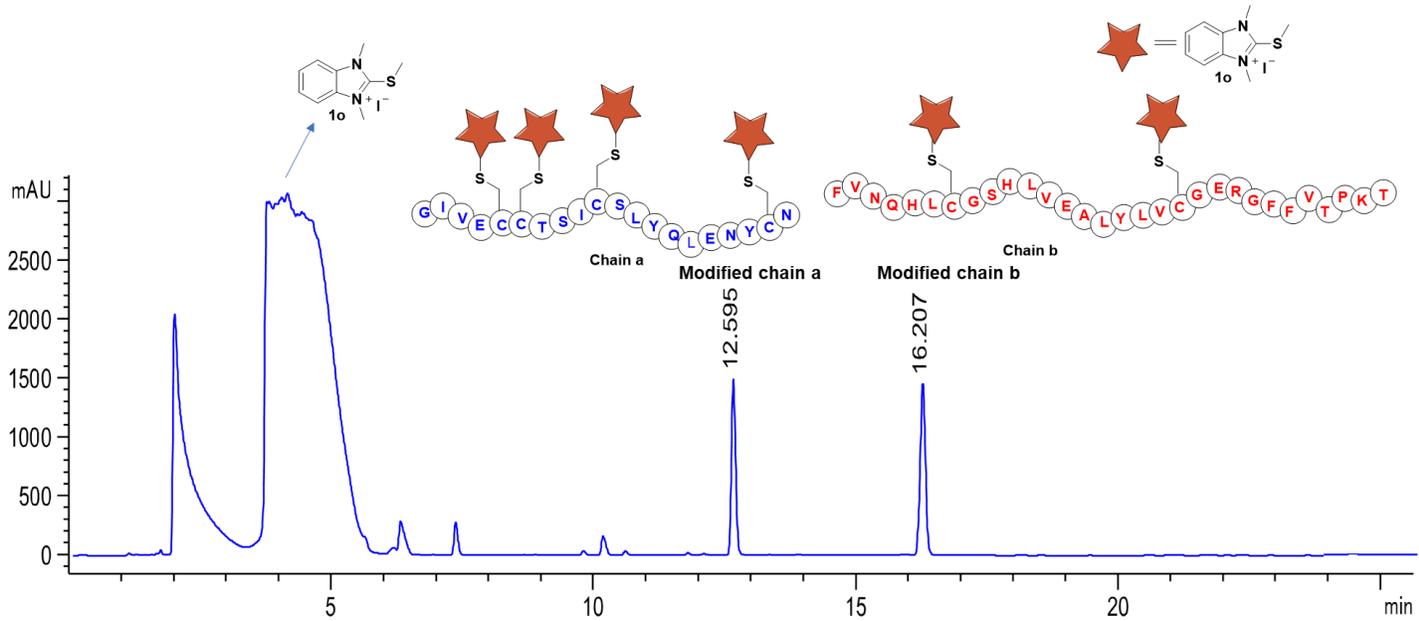
Procedure for reversible study of modified reduced insulin by NaBH₄

In a solution of modified reduced insulin (0.15 mM) in 400 μ L of 10 mM Nap (pH 7.5) was added 10 equiv. of NaBH₄ (1.5 mM). The mixture was stirred at room temperature for 5 min. The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

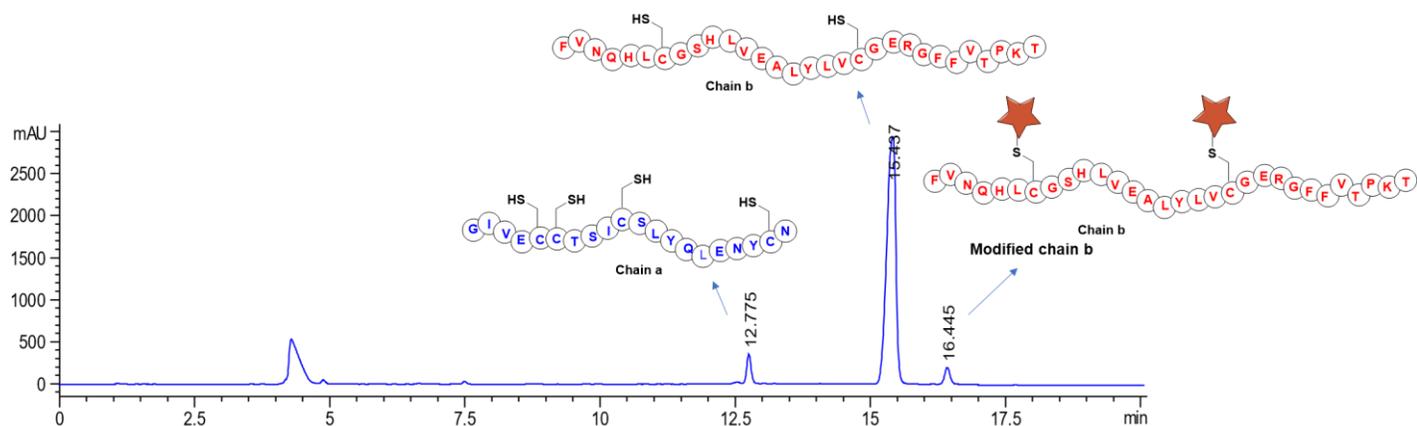
HPLC trace of reduced insulin



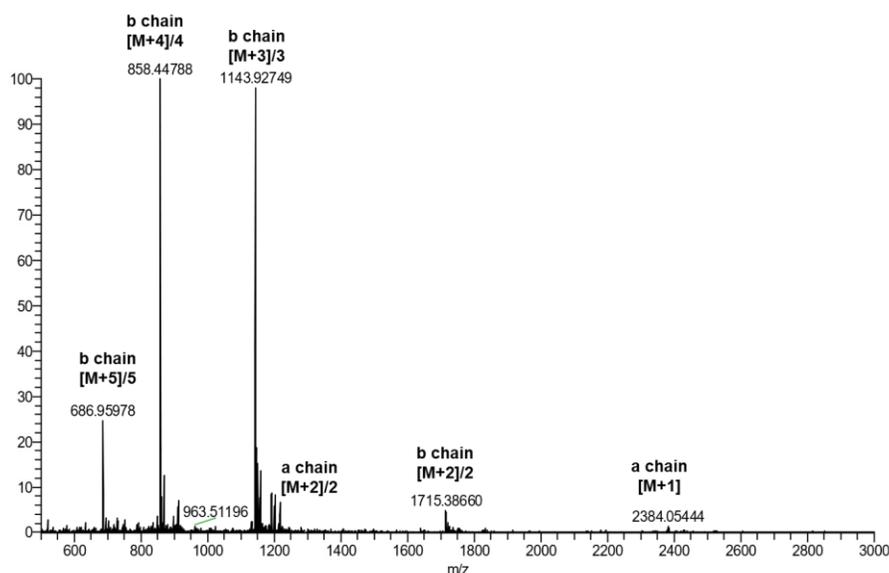
HPLC trace of modified chain a and chain b



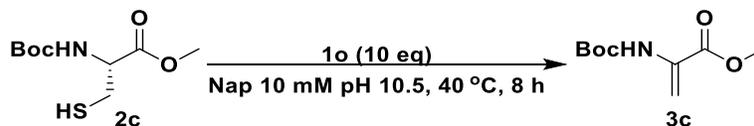
HPLC trace after addition of NaBH₄



HRMS after addition of NaBH₄

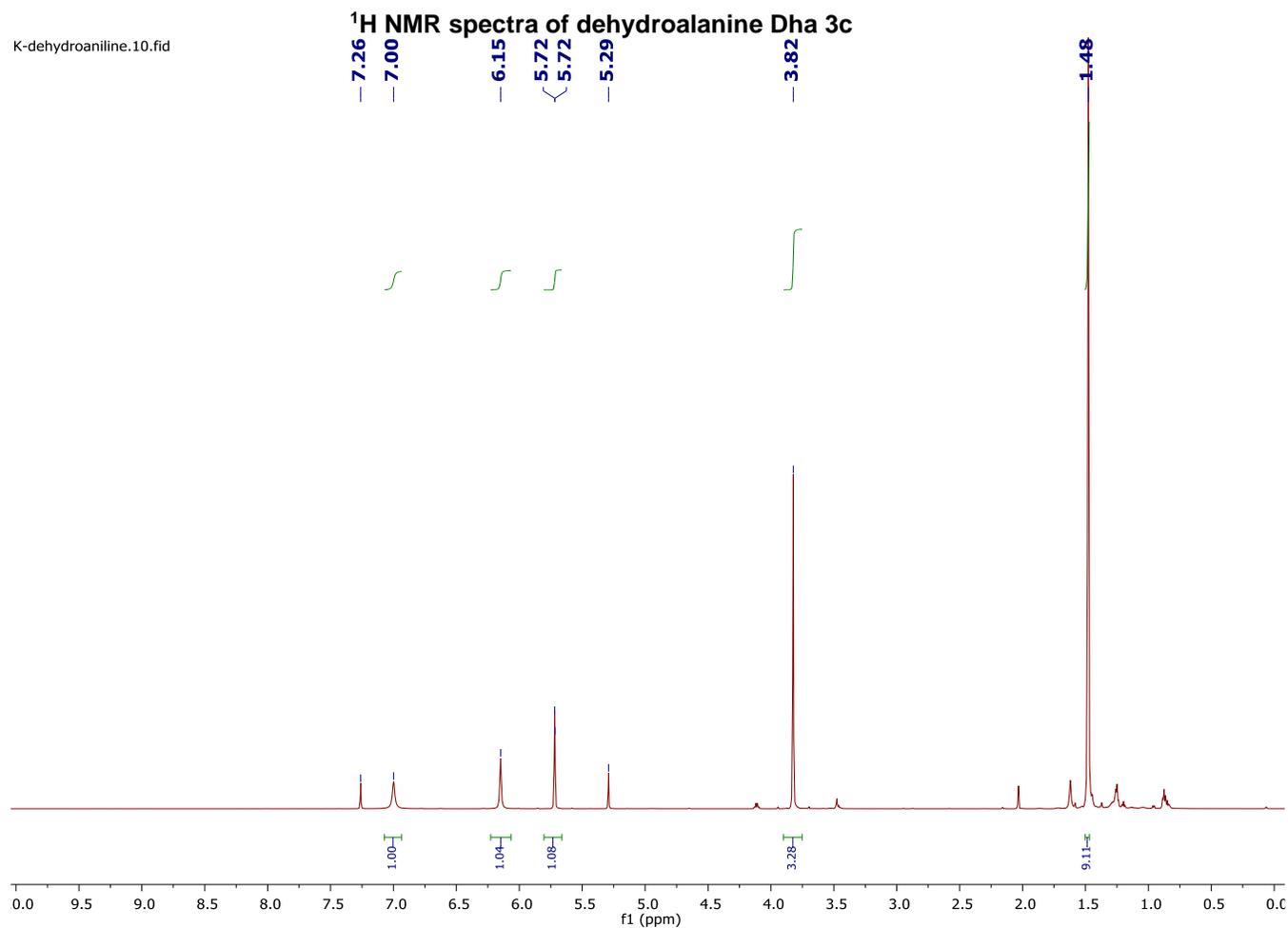


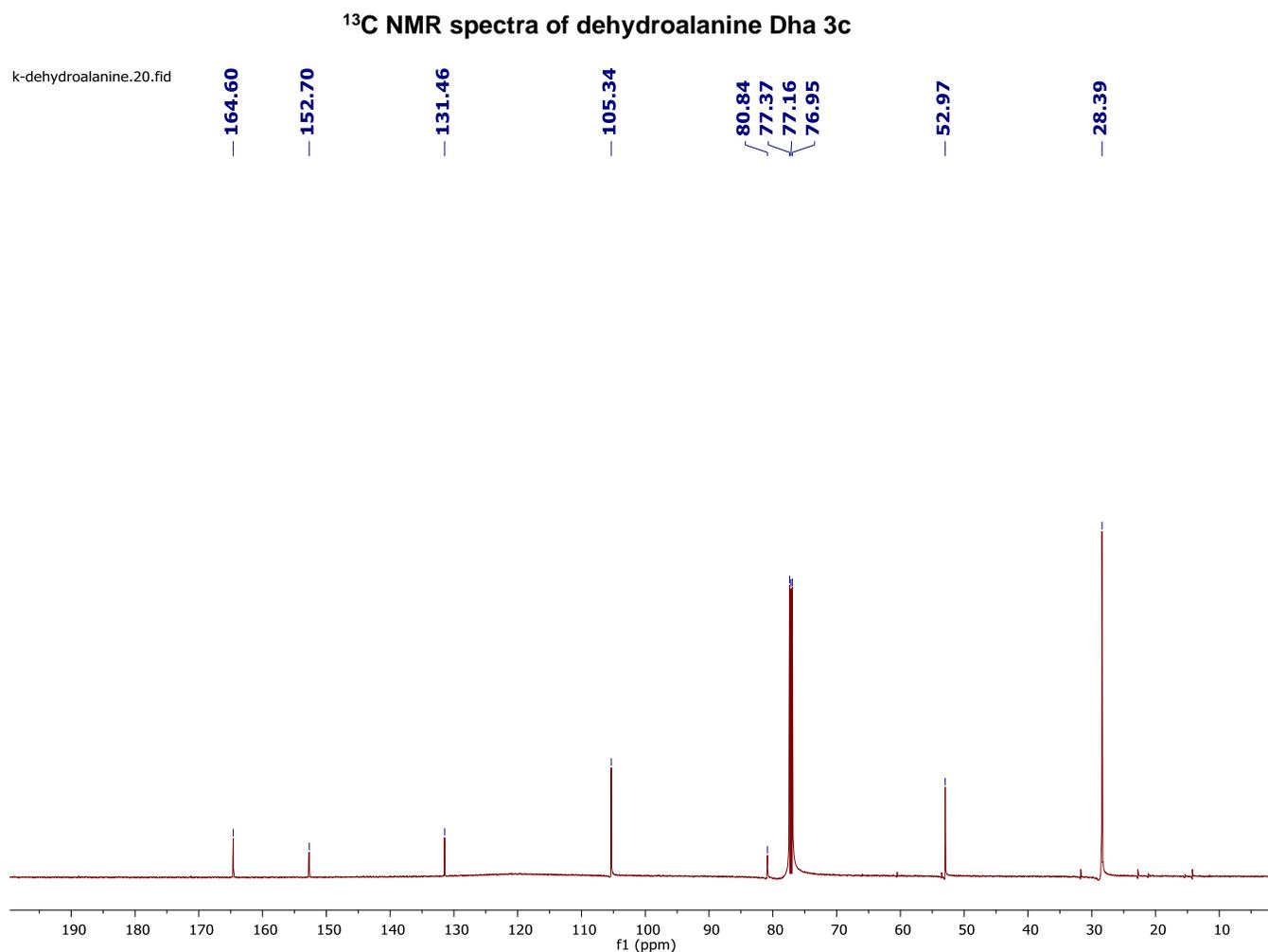
Supplementary Figure 16. Dehydroalanine synthesis from cysteine **2c using **1o** at high pH**
Procedure for synthesis of dehydroalanine from cysteine **2c**



To 500 mg (2.12 mmol) of Boc-Cys-OMe **2c** in 1 mL of 10 mM Nap (pH 10.5), probe **1o** (50 equiv.) was added and the reaction was allowed to react at 40 °C for 8 h. The reaction solution was washed with ethyl acetate and brine for 3 times. The residue was concentrated under reduced pressure and purified by the column chromatography (hexane: ethyl acetate 3:1) to obtain

dehydroalanine **3c** as colorless oil (312.2 mg, 73 %). NMR. ^1H NMR (600 MHz, Chloroform-*d*) δ 7.00 (br, 1H), 6.15 (s, 1H), 5.72 (d, 1H, $J = 1.5$ Hz), 3.82 (s, 3H), 1.49 (s, 9H). ^{13}C NMR (151 MHz, Chloroform-*d*) δ 164.6, 152.7, 131.7, 105.3, 80.8, 53.0, 28.4.





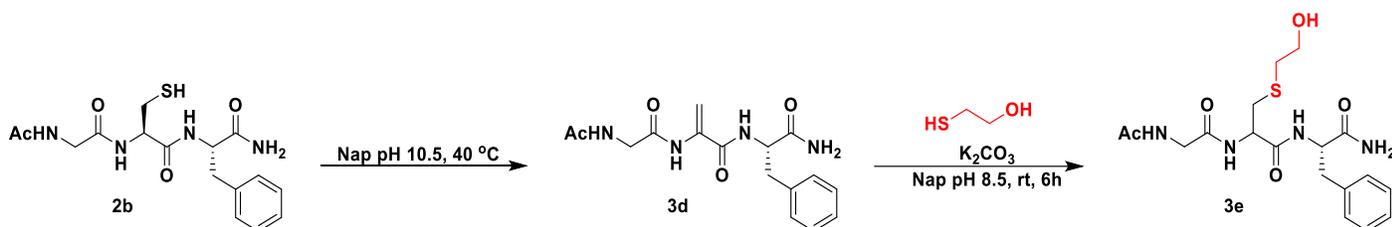
Supplementary Figure 17. Aza-Michael addition and thiol-ene reaction of dehydroalanine.

General procedure for synthesis of Dha peptide 3d

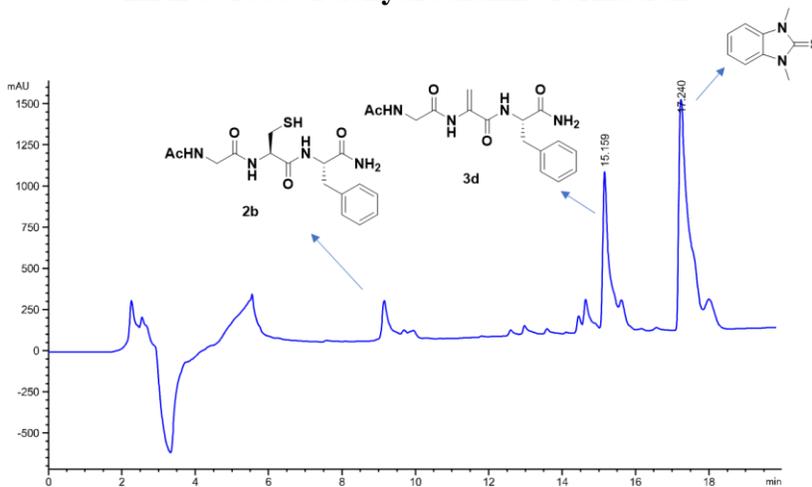
To a solution of Ac-GCF **2b** (0.57 mg, 0.0015 mmol) in 400 μ L of 10 mM Nap (pH 10.5) was added **1o** (25 equiv., 0.0375 mmol). The reaction mixture was incubated at 40 $^{\circ}$ C for 12 h. The crude compound was purified by HPLC and lyophilized to afford dehydroalanine (88 % conversion). The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

Procedure for Dha peptide thiol-ene reaction

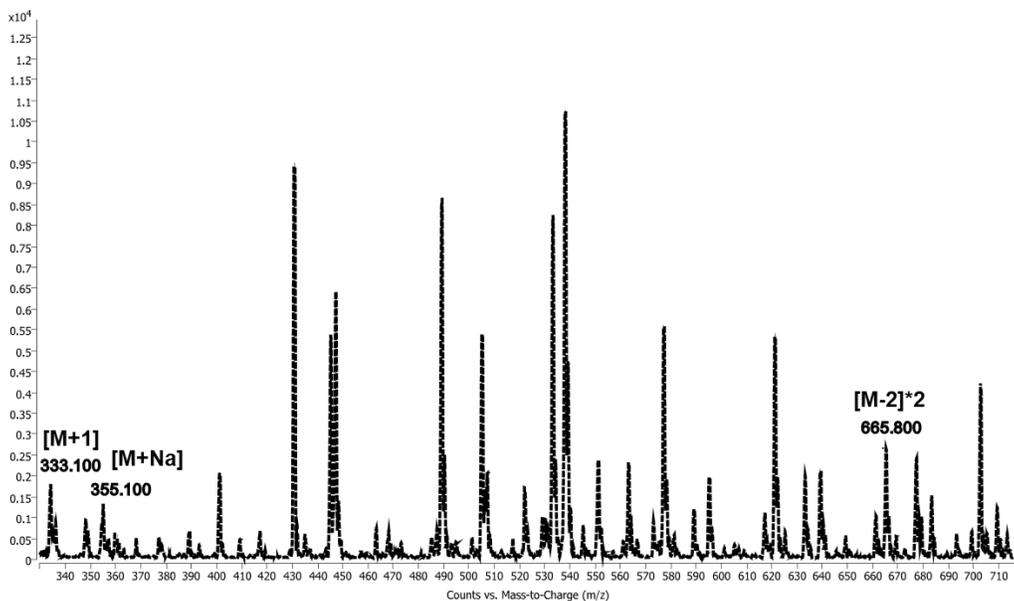
To a solution of Dha peptide **3d** (0.0015 mmol) in 400 μL of 10 mM Nap (pH 8.5) was added K_2CO_3 (4 equiv., 0.006 mmol) and 2-mercaptoethanol (10 eq, 0.015 mmol). The reaction mixture was incubated at rt for 6 h to generate **3e** (>99 % conversion). The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



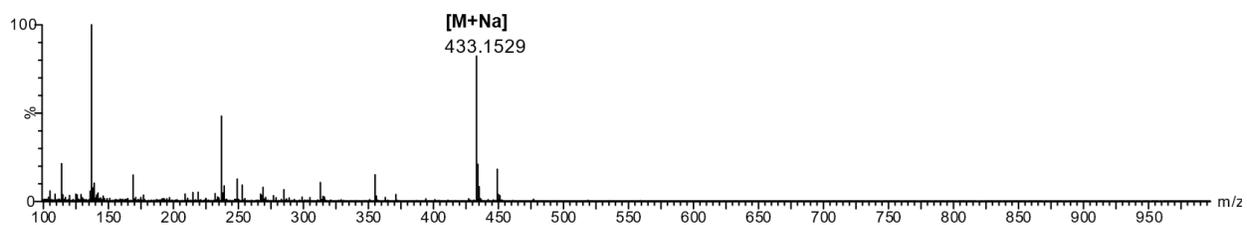
HPLC trace of dehydroalanine formation



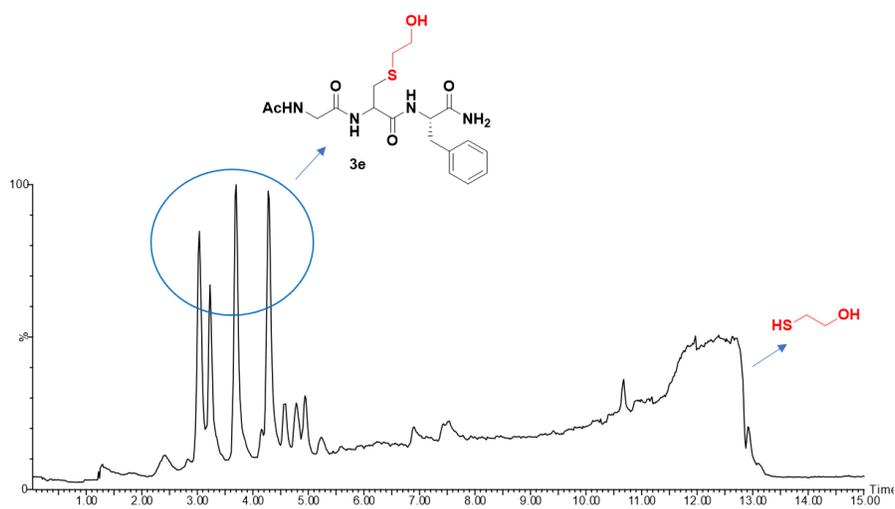
HRMS of 3d



HRMS of 3e



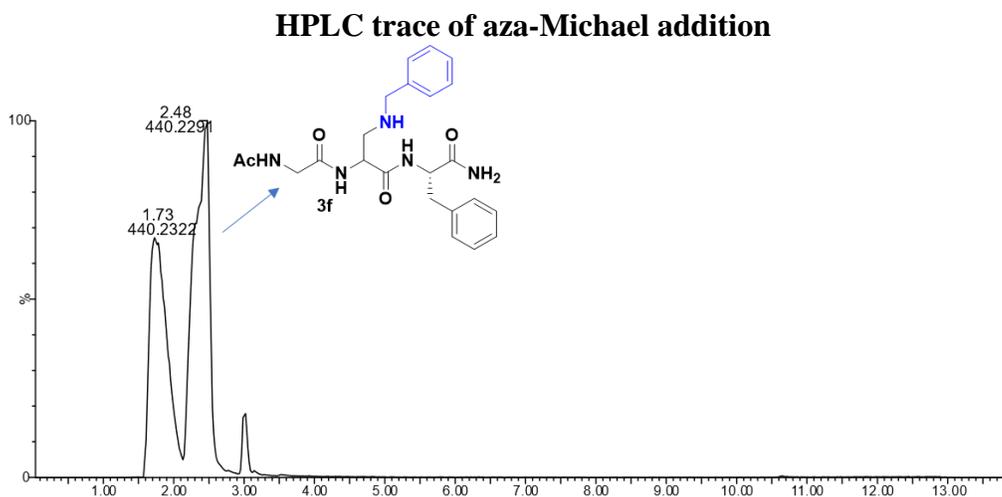
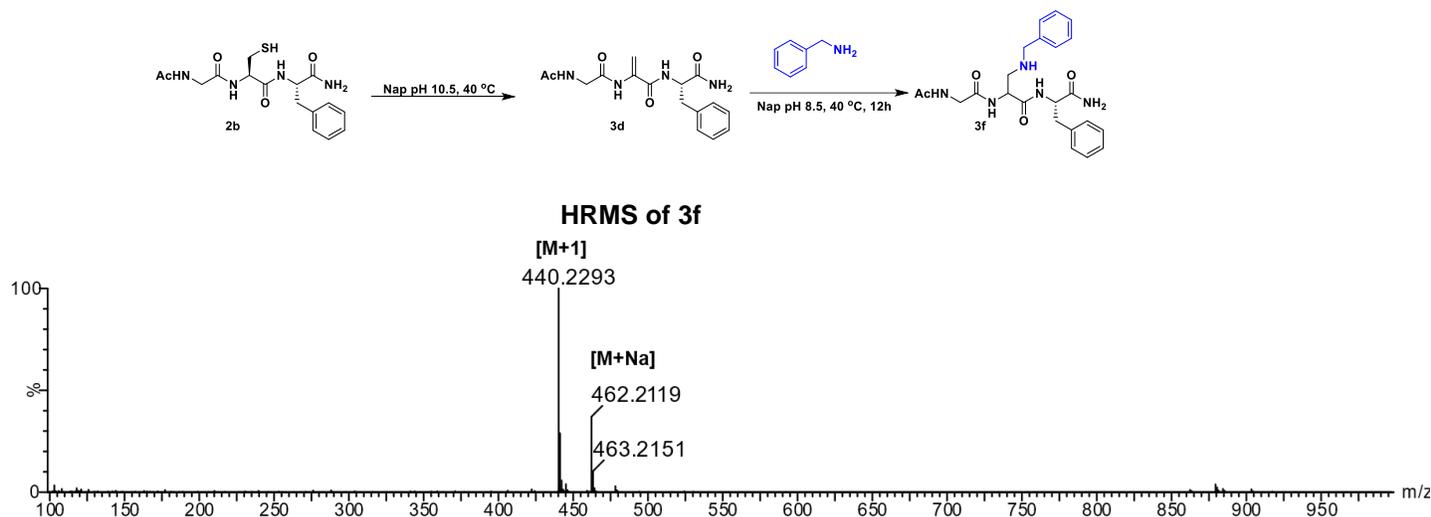
HPLC trace of thio-ene reaction



Procedure for dehydroalanine aza-Michael addition

Dha peptide **3d** (0.0015 mmol) was dissolved in 0.4 mL of 10 mM Nap (pH 8.5), 4 equiv. of

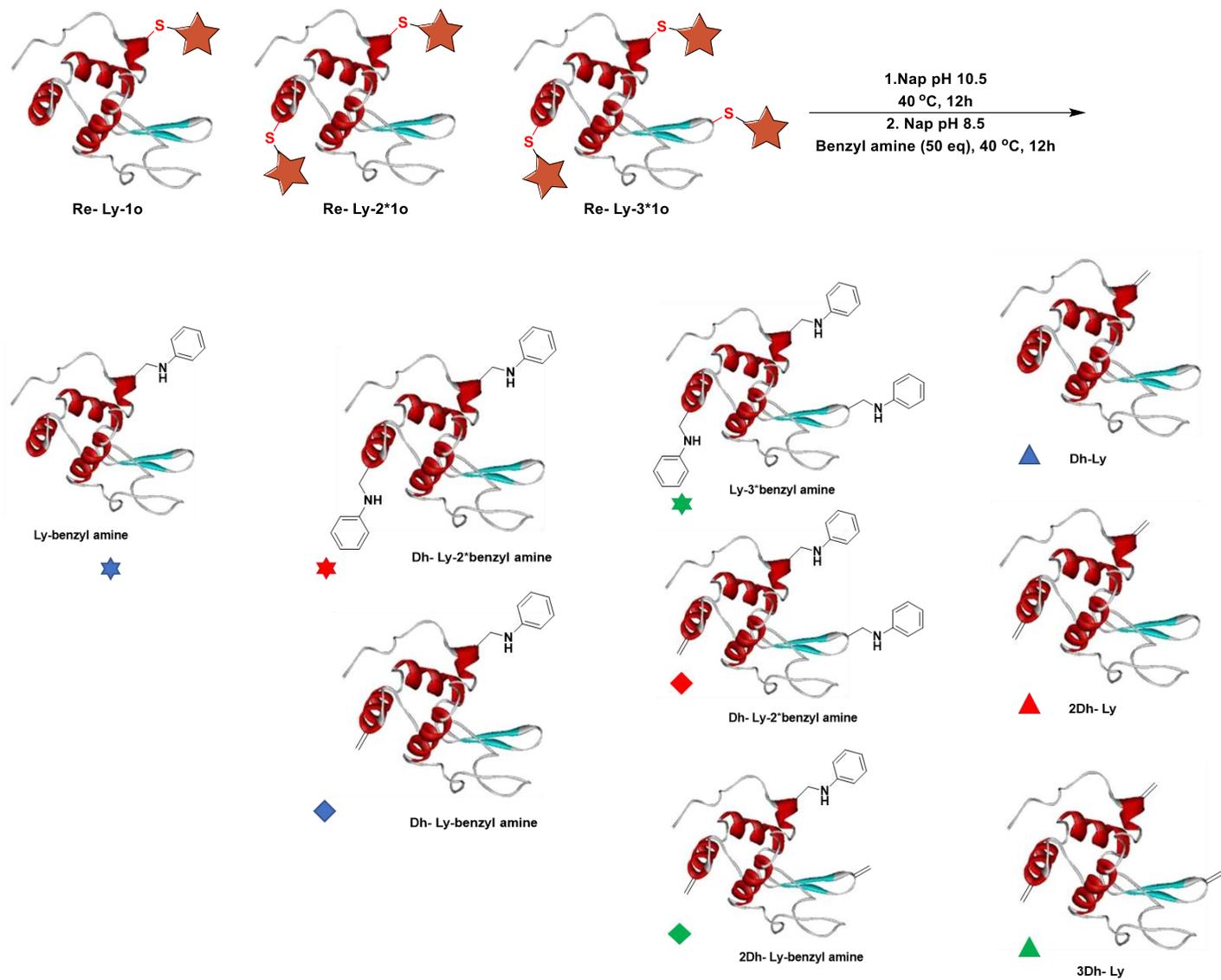
benzylamine was added and the reaction mixture was stirred at 40 °C for 12 h to generate 3f (> 99% conversion). The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



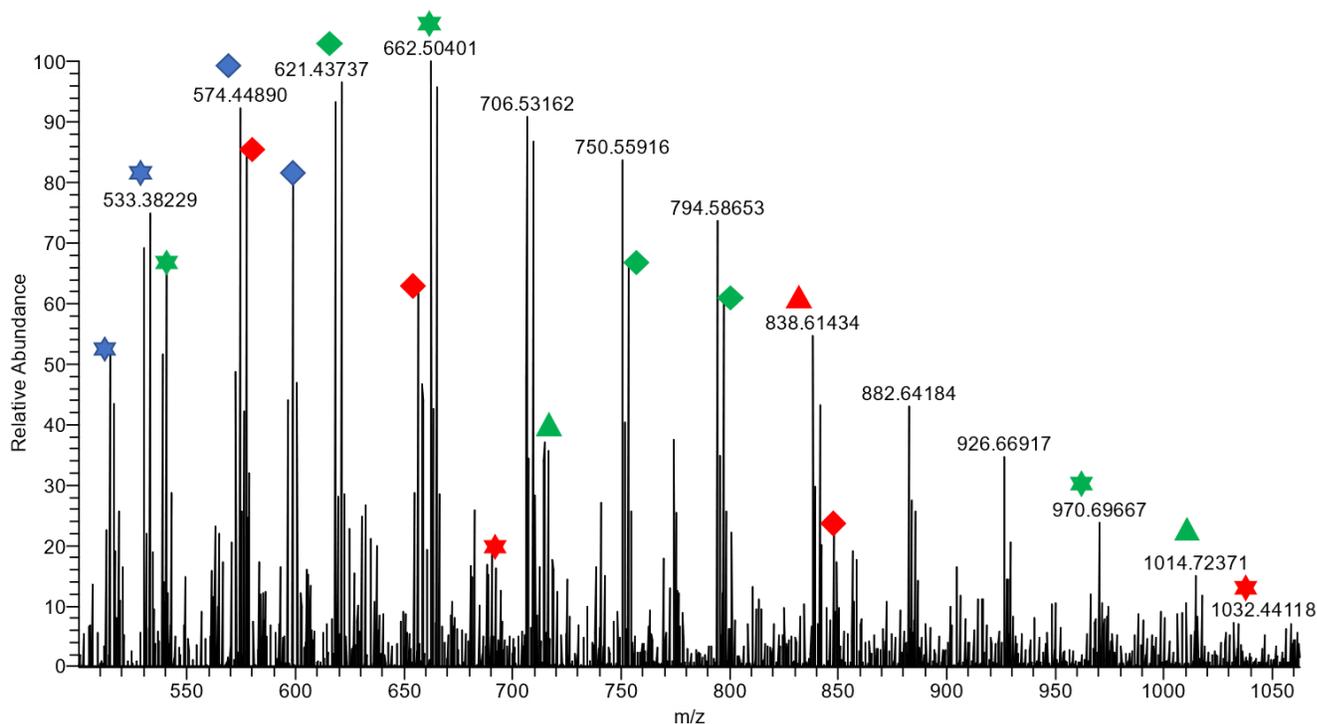
Supplementary Figure 18. Aza-Michael addition of dehydroalanine lysozyme

To a mixture of lysozyme (60 μmol, 0.15 mM) in 400 μL 10 mM Nap (pH 7.5), probe **1o** (5.76 mg, 18 μmol, 300 equiv.) was added. The reaction mixture was incubated at room temperature for 8 h and lyophilized to obtain modified lysozyme. To a solution of modified lysozyme (0.15 mM) in 400 μL of 10 mM Nap (pH 10.5) was incubated at 40 °C for 12 h. The crude compound was

purified by molecular weight-cutoff and lyophilized to afford dh-lysozyme. In a solution of dh-lysozyme (0.15 mM) in 400 μ L of 10 mM Nap (pH 8.5) was incubated with benzyl amine (7.5 mM) at 40 $^{\circ}$ C for 12 h. The reaction was analyzed by ESI-MS.



HRMS of dehydroalanine lysozyme Aza-Michael addition



Supplementary Figure 19. Mass sensitivity booster capability of 1o-modified peptides.

Procedure for checking intensity ratios of IAA-Ac-GCF, 1i-Ac-GCF and 1o-AcGCF

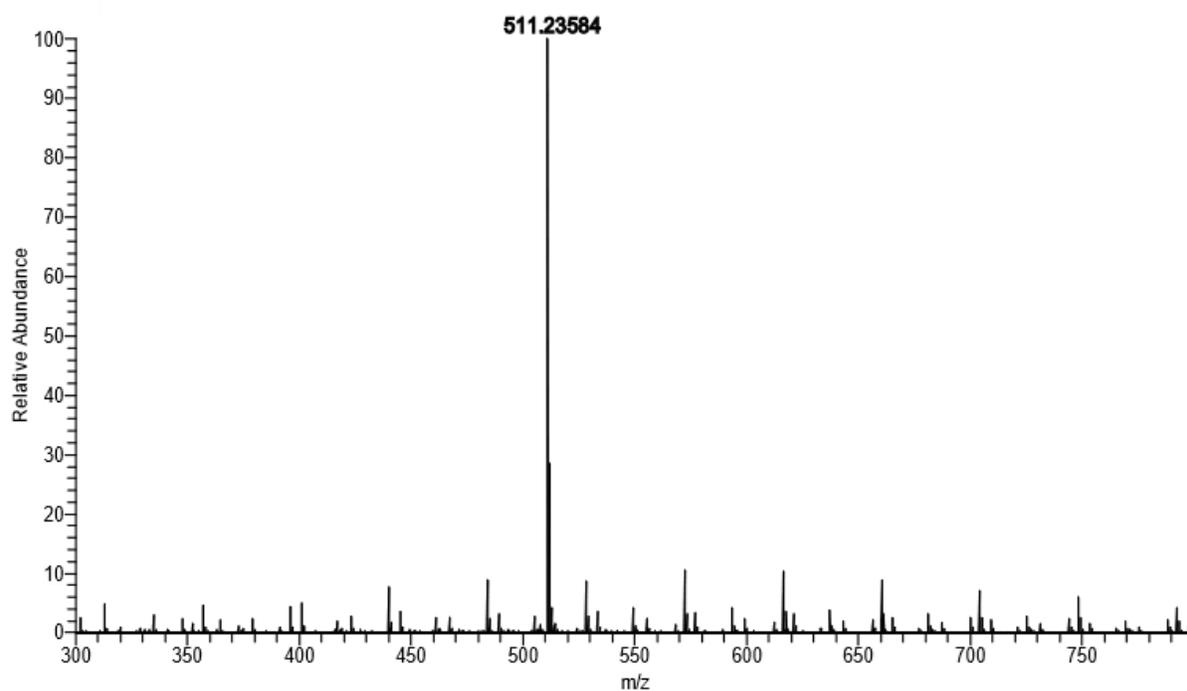
The peptides 5 μmol (IAA-Ac-GCF, 1i-Ac-GCF and Ac-GCF) were taken in an eppendorf tube containing acetonitrile (100 μL). The 1o-AcGCF 5 μmol were taken in another eppendorf tube containing acetonitrile (100 μL). Equal volumes (50 μL) of each solution were taken from the stock solution in another eppendorf tube. The mixture was vortexed, and 50 μL was transferred to the HPLC vial for ESI-MS. Subsequently, the intensity ratios were analyzed by MS.

Supplementary Figure 20. Mass sensitivity of 1o-Ac-GCF (low concentration)

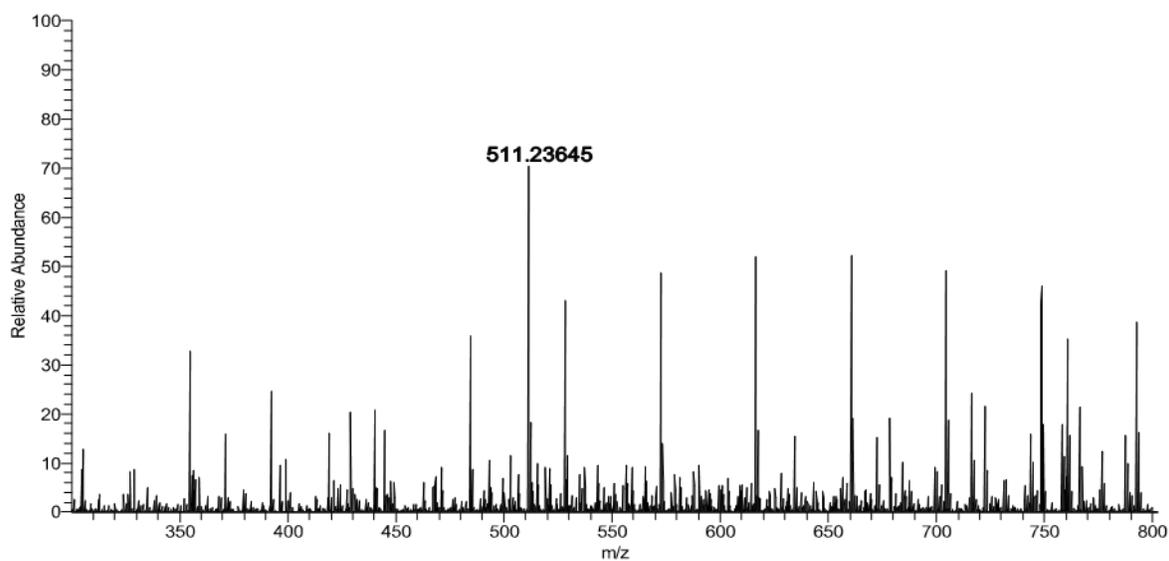
General procedure for checking mass intensity of 3b at low concentration

The modified peptide 3b (5 μmol) in 100 μL Nap pH 7.5 was diluted to 5 nM, 0.5 nM with DI water. The mixture was vortexed, and 50 μL was transferred to the HPLC vial for ESI-MS. Subsequently, the mass intensity of each concentration were analyzed by MS.

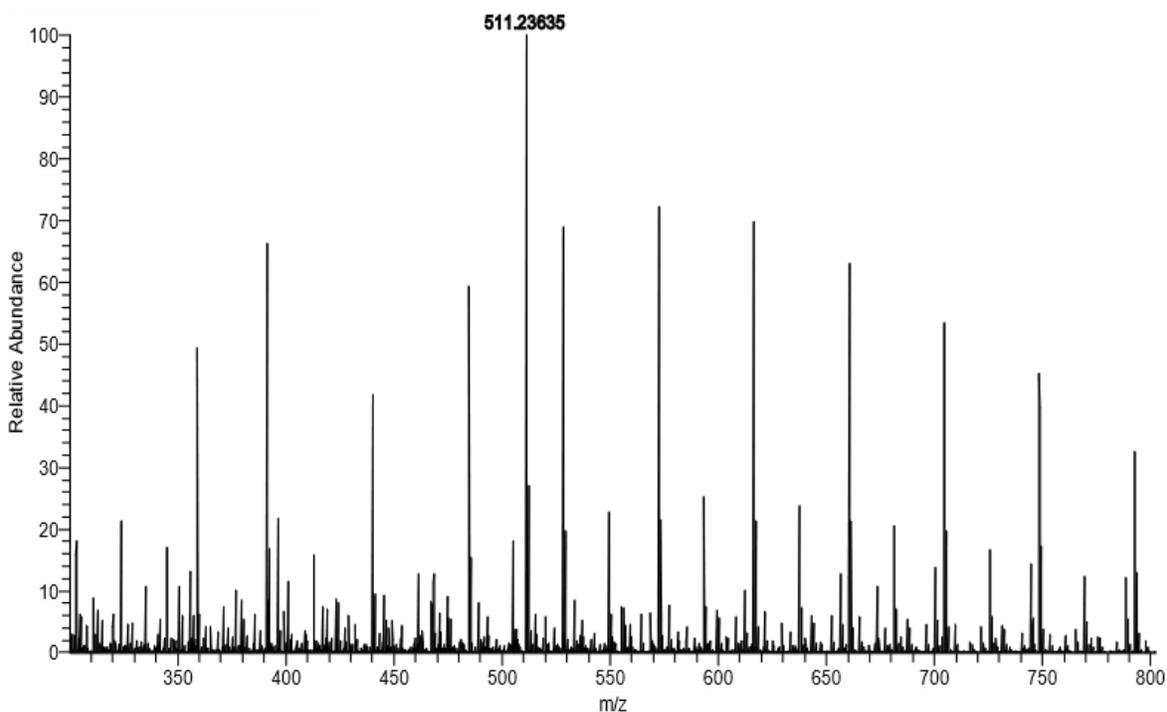
Mass spectra of 3b at 0.05 μM



Mass spectra of 3b at 5 nM



Mass spectra of 3b at 0.5 nM



Supplementary Figure 21. Mass intensity enhancement of N₃-1_o-reduced insulin bioconjugate products.

Procedure for checking intensity ratios of N₃-1_o-reduced insulin and IAA-reduced insulin.

To 0.35 mg of insulin (60 μ mol, 0.15 mM) in 400 μ L 10 mM Nap (pH 7.5), 20 μ L of TCEP solution (pH 7.4, 50 mM) was added to reduce the insulin. The mixture was incubated at room temperature for 20 min. Probe N₃-1_o or IAA (50 equiv.) was added into this mixture and the reaction was allowed to react at room temperature for 8 h. Equal volume (50 μ L) of each solution was taken from the reaction mixture then transferred into the HPLC vial for ESI-MS. Subsequently, the intensity ratios were analyzed by MS.

Supplementary Figure 22. HAT probes for Gel-based ABPP

Cell culture and preparation of cell lysates. Cell culture reagents including Dulbecco's phosphate-buffered saline (DPBS), Dulbecco's modified Eagle's medium (DMEM)/high glucose

media, trypsin-EDTA and penicillin/streptomycin (Pen/Strep) were purchased from Fisher Scientific. Fetal Bovine Serum (FBS) were purchased from Avantor Seradigm (lot # 214B17).

HEK293T (ATCC: CRL-3216) cells were cultured in DMEM supplemented with 10% FBS and 1% antibiotics (Penn/Strep, 100 U/mL). Media was filtered (0.22 μ m) prior to use. Cells were maintained in a humidified incubator at 37 °C with 5% CO₂. Cell lines were validated prior to use and tested regularly for myoplasma.

HEK293T cells were harvested once cells were grown to 90 – 95% confluence by centrifugation (4500g, 5 min, 4 °C), washed twice with cold DPBS, resuspended in 300 μ L DPBS, sonicated, and clarified by centrifuging (21,000 g, 10 min, 4 °C). The lysates were then transferred to an eppendorf tube. Protein concentrations were determined using a Bio-Rad DC protein assay kit using reagents from Bio-Rad Life Science (Hercules, CA) and the lysate diluted to the working concentrations indicated below.

Gel-based ABPP with N3-1o. HEK293T proteome (50 μ L of 1.5 mg/mL, prepared as described above) was labeled with various concentration of **N3-1o** (stock solutions in DMSO, final concentration as indicated), IAA (1 μ L of 5 mM stock solution in DMSO, final concentrations = 1-200 μ M) or DMSO for vehicle control for 1h at ambient temperature followed by adding 1 μ M IA-Rh. Samples were allowed to react for another hour at ambient temperature at which point the reactions were quenched with 4 \times Laemmli buffer (20 μ L). Samples were then denatured (5 min, 95 °C) and then resolved by SDS-PAGE. SDS-PAGE gels were imaged on the Bio-Rad ChemiDoc™ Imager using rhodamine channel.

Gel-based ABPP with N3-1o. HEK293T proteome (50 μ L of 1.5 mg/mL) was labeled with different amount of **N3-1o** (stock solutions in DMSO, final concentration as indicated), STP-alkyne (1 mM) or DMSO for vehicle control for 1h at ambient temperature followed by adding 1

μ M NHS-Rh. Samples were allowed to react for another hour at ambient temperature at which point the reactions were quenched with 4 \times Laemmli buffer (20 μ L). Samples were then denatured (5 min, 95 $^{\circ}$ C) and then resolved by SDS-PAGE. SDS-PAGE gels were imaged on the Bio-Rad ChemiDocTM Imager using rhodamine channel.

Supplementary Figure 23. Modification of myoglobin Mb by 1d and 1e

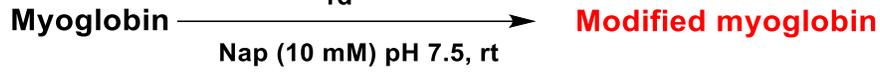
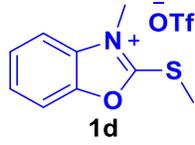
To a 1 mg of myoglobin (3 mM) in 0.2 mL of 10 mM phosphate buffer (Nap, pH 7.5), compound 1d (1 or 10 equiv.) was added. The reaction was stirred at room temperature for 1 h. The Mb-1d protein bioconjugate was purified by molecular weight cut off and characterized by LC-MS/MS.

Procedure for in-solution digestion and analysis of modified myoglobin

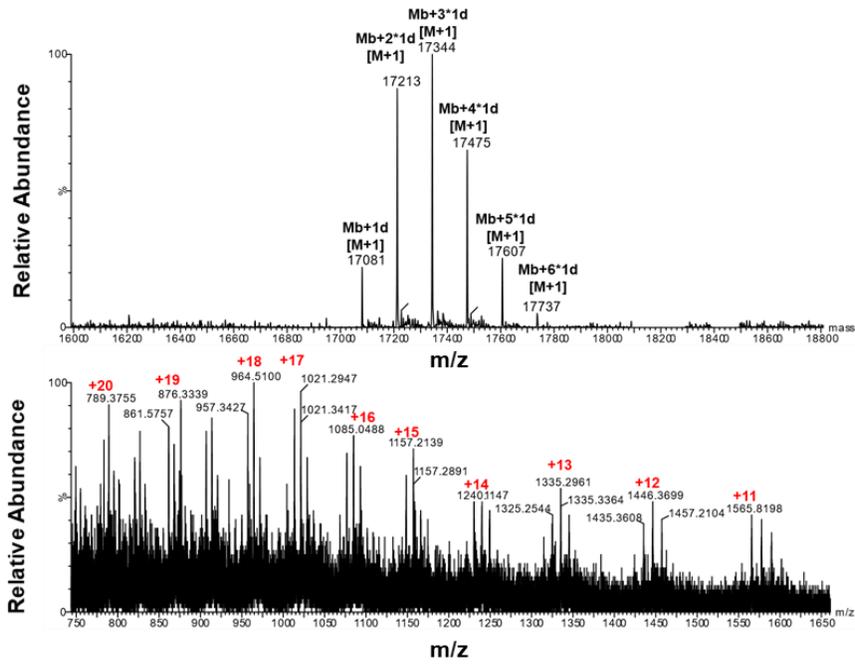
Protein (1 mg) in 10 mM tris (100 μ L, pH 7.5) with urea (6 M) was incubated for 30 min at 37 $^{\circ}$ C. To reduce the urea concentration to 0.6 M, the sample was diluted with grade I water. To this solution, 100 μ L of enzyme (α chymotrypsin/trypsin) solution (0.1 mg, enzyme/protein (1:20); enzyme in 1 mM HCl was dissolved in 0.1 M tris and 0.01 M CaCl₂) was added and the mixture was incubated at 37 $^{\circ}$ C for 18 h. The pH of digested mixture was adjusted to < 6 (confirmed by pH paper) with trifluoroacetic acid (0.5%). Subsequently, the sample was used for peptide mapping by MS and sequencing by MS/MS. The protein fragments were analyzed by Thermo Ultimate 3000 nanoLC/Orbitrap Q-Exactive Plus MS with positive mode. The identification of protein sequences was achieved by Thermo BioPharma Finder.

Myoglobin from equine heart amino acid sequence:

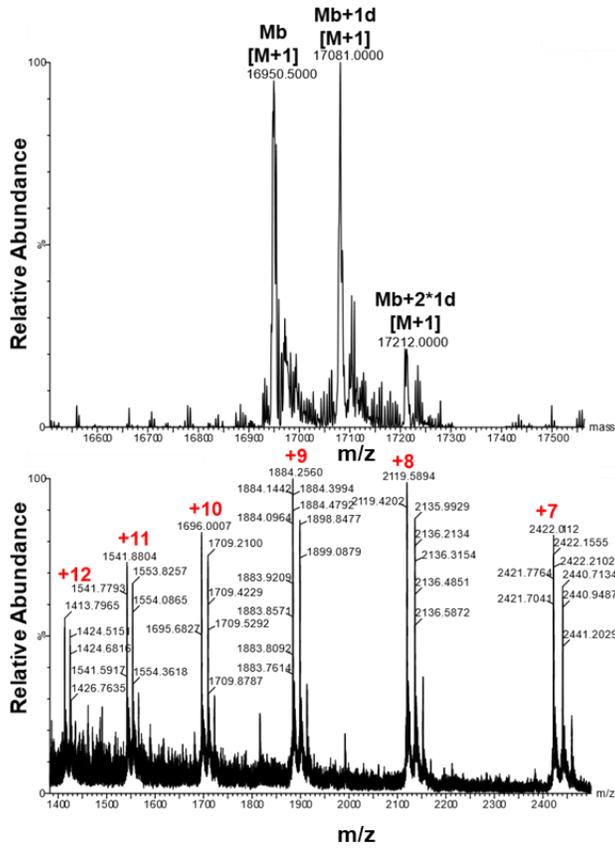
GLSDGEWQQVLNVWGKVEADIAGHGQEVLRIRLFTGHPETLEKFDKFKHLKTEAEMKAS
EDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISDAIIHVLHSK
HPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG (no cysteine or disulfide)



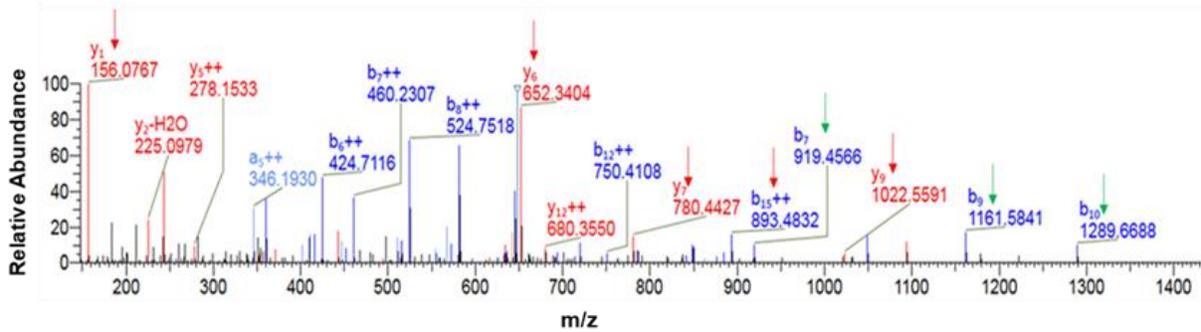
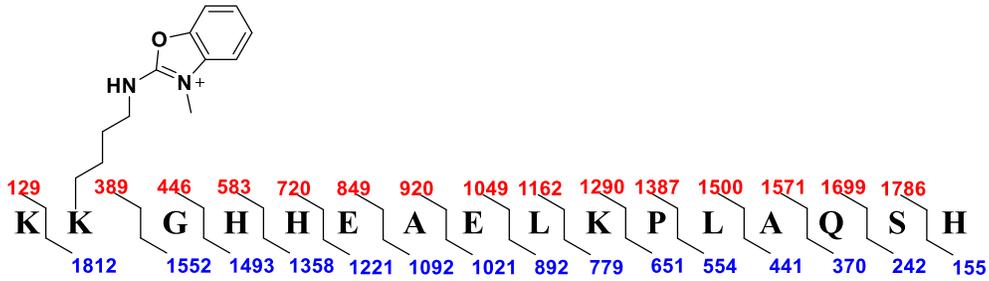
HRMS of 1d-Mb (10 eq, 1h)



HRMS of 1d-Mb (1 eq, 1h)



MS/MS spectrum Mb-1d fragment (1eq, 1h)

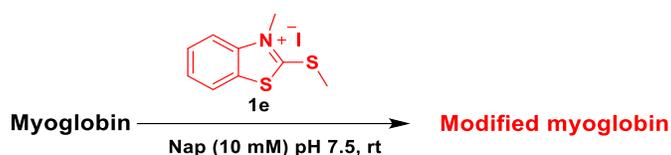


Protein	Peptide	Modification	Site	Confidence	ID	M/Z	Charge	Mono Mass	Avg
---------	---------	--------------	------	------------	----	-----	--------	-----------	-----

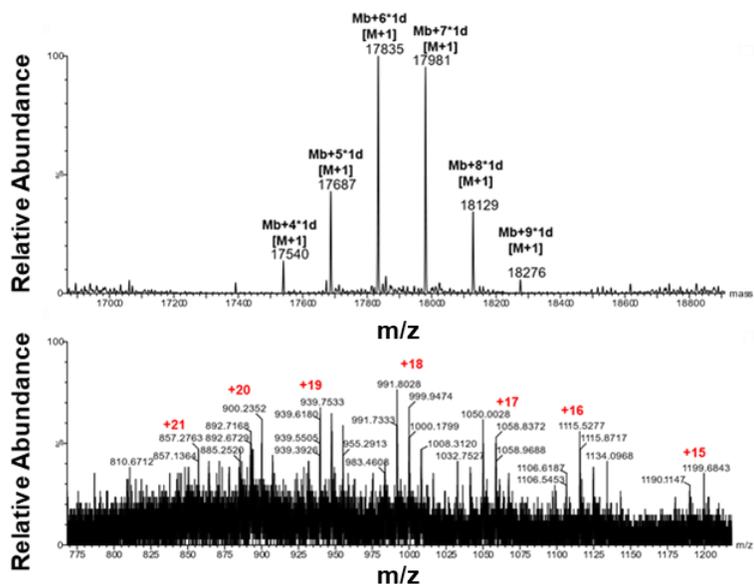
	Sequence			Score	Type		State	Exp.	Mass Exp.
Modified Myoglobin	KKGHHEALK PLAQSH	131.0337	~K79	100	MS2	486.259	4	1940.004	1941.15

General procedure of tagging multiple lysines with compound 1e

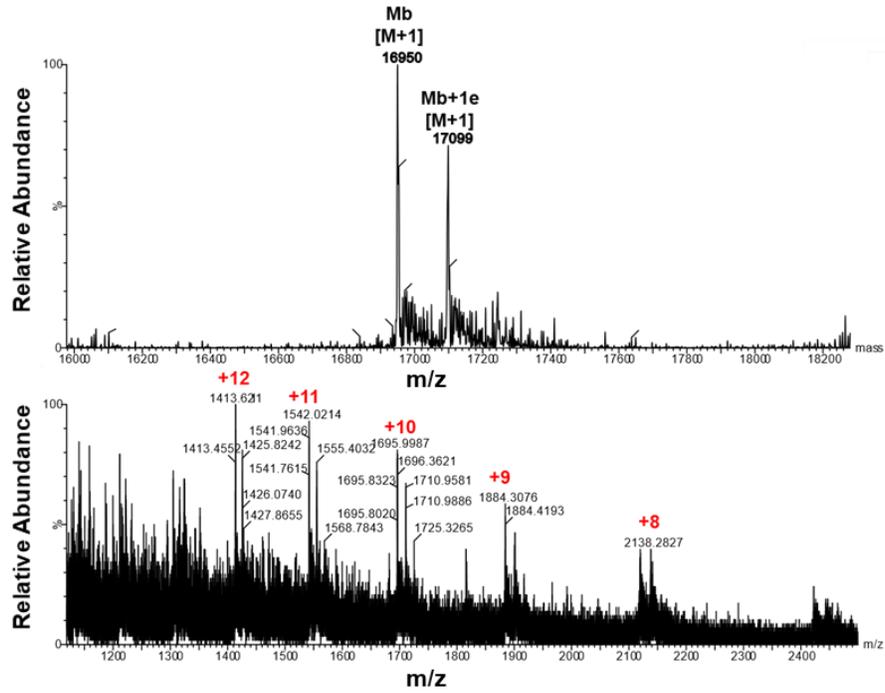
To a 1 mg of myoglobin (3 mM) in 0.2 mL of 10 mM phosphate buffer (Nap, pH 7.5), compound 1e (10 or 100 equiv.) was added. The reaction was stirred at room temperature for 1 or 12 h. The Mb-1e protein bioconjugate was purified by molecular weight cut off and characterized by LC-MS/MS.



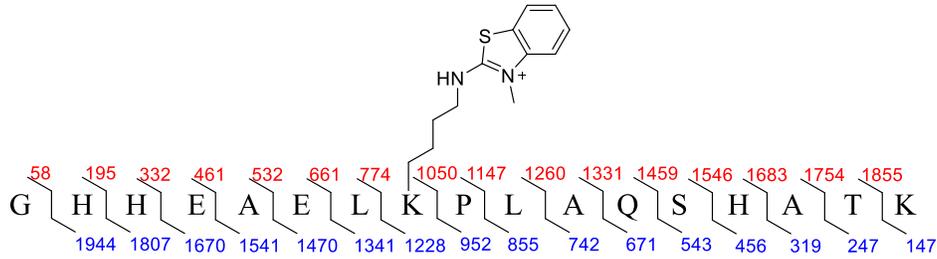
HRMS of 1e-Mb (100 eq, 12h)



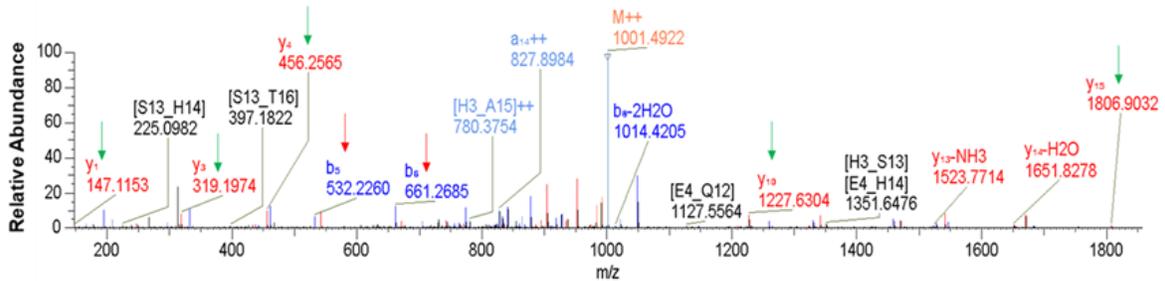
HRMS of 1e-Mb (10 eq, 1h)



MS/MS spectrum Mb-1e fragment (10 eq, 1h)



Protein	Peptide Sequence	Modification	Site	Confidence Score	ID Type	m/z	Charge State	Mono Mass Exp.	Avg Mass Exp.
Modified Myoglobin	GHHEAELKPLAQSHTK	147.0146	~K87	100	MS2	1000.992	2	1999.9698	2001.03



Lysine-tagging of α -lactalbumin (Lb) and cytochrome C (CyC) with probe 1d.

To protein (Lb or CyC) (3 mM) in 0.2 mL 10 mM phosphate buffer (pH 7.5), 1d (1 equiv.) was added. The reaction was stirred at room temperature for 1 h. The Lb-1d and CyC-1d protein bioconjugates were purified by molecular weight cut off and characterized by LC-MS.

Cytochrome C from equine heart amino acid sequence:

GDVEKGKKIFVQKCAQCHTVEKGGKHKHTGPNLHGLFGRKTGQAPGFTYTDANKNK

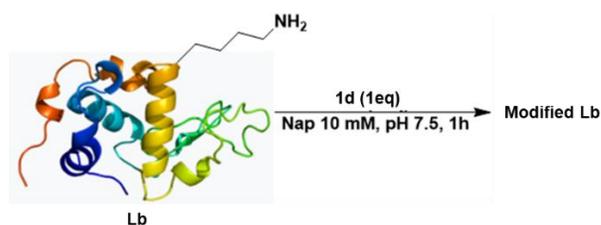
GITWKEETLMEYLENPKKYIPGTKMIFAGIKKKTEREDLIAYLKKATNE (1-disulfide bond)

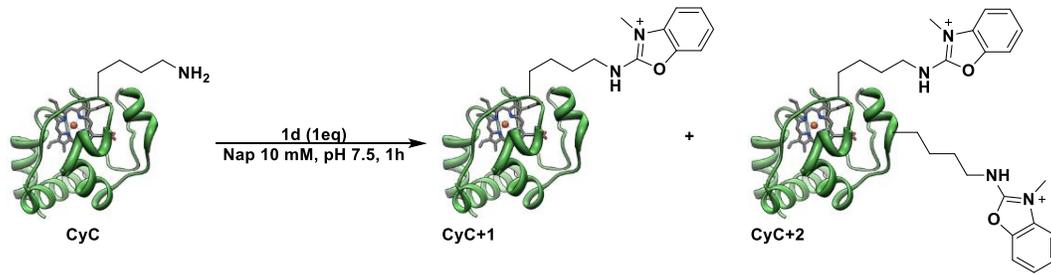
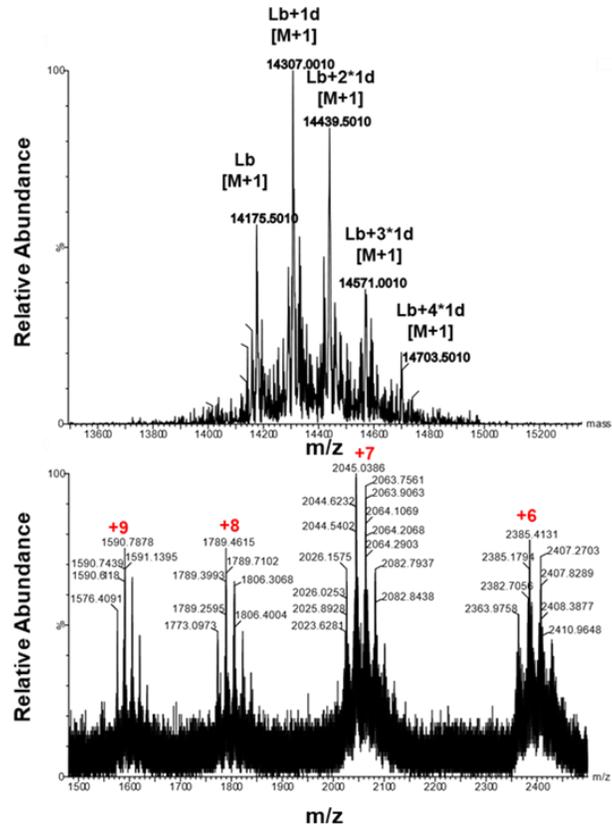
α -Lactalbumin from bovine milk amino acid sequence:

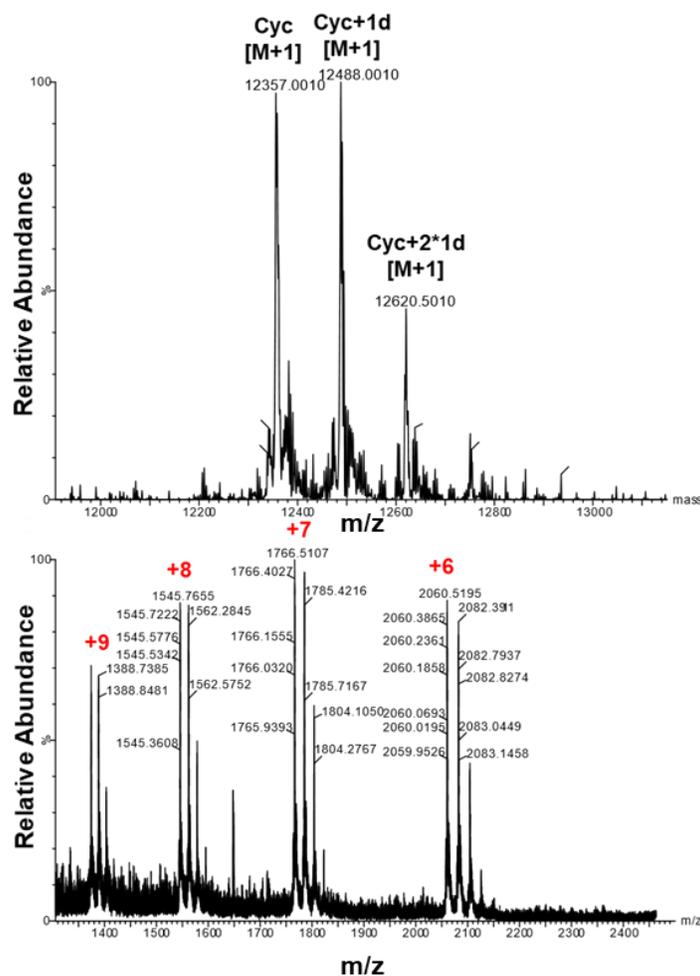
EQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQAIVQNNDSTEYGLFQINNKI

WCKNDQDPHSSNICNISCDKFLNNDLTNNIMCVKKILDKVGINYWLAHKALCSEKLDQ

WLCEKL (8-disulfide bonds)



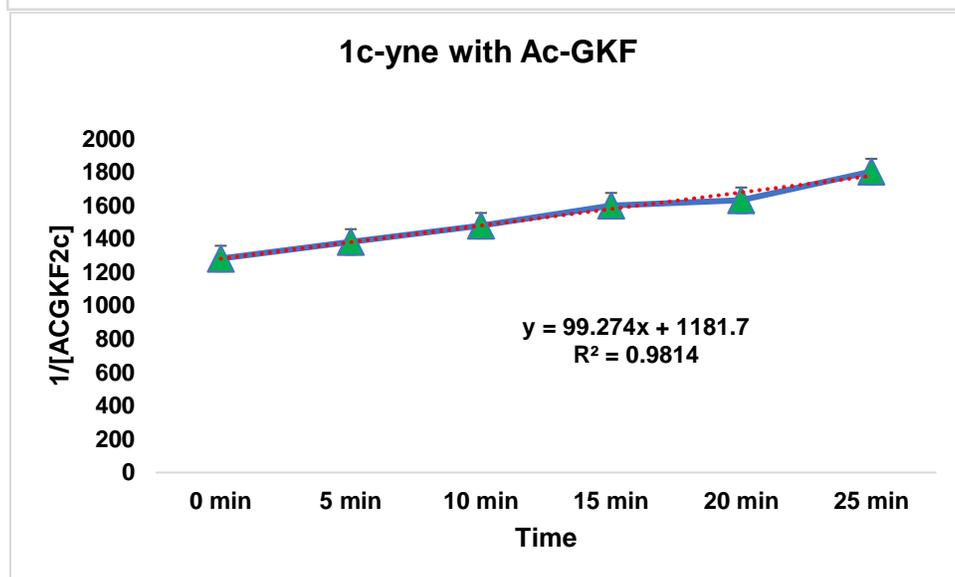
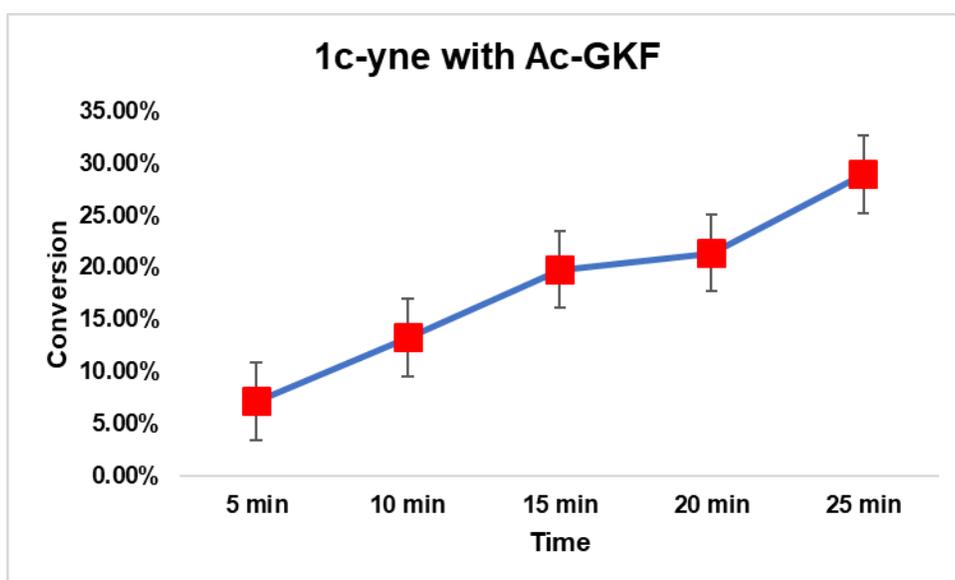
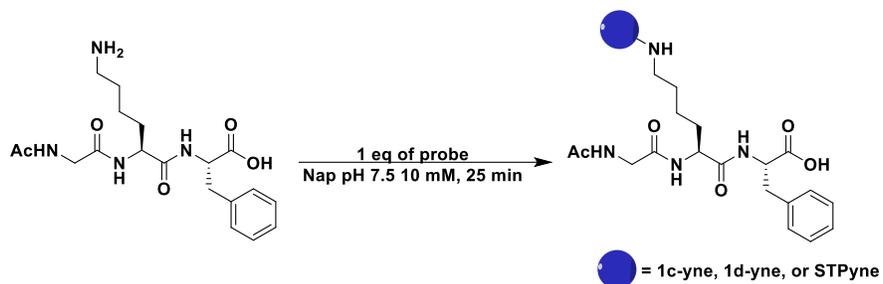


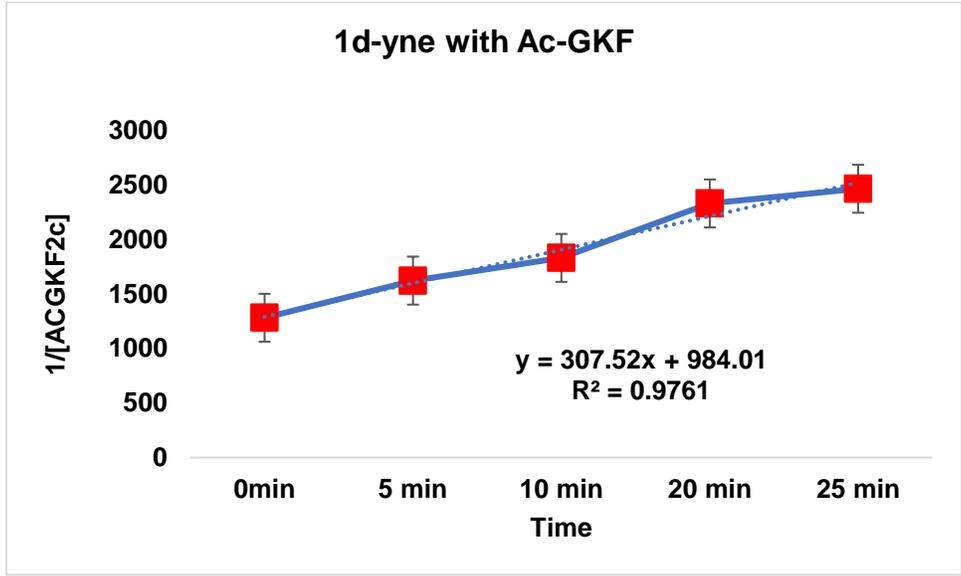
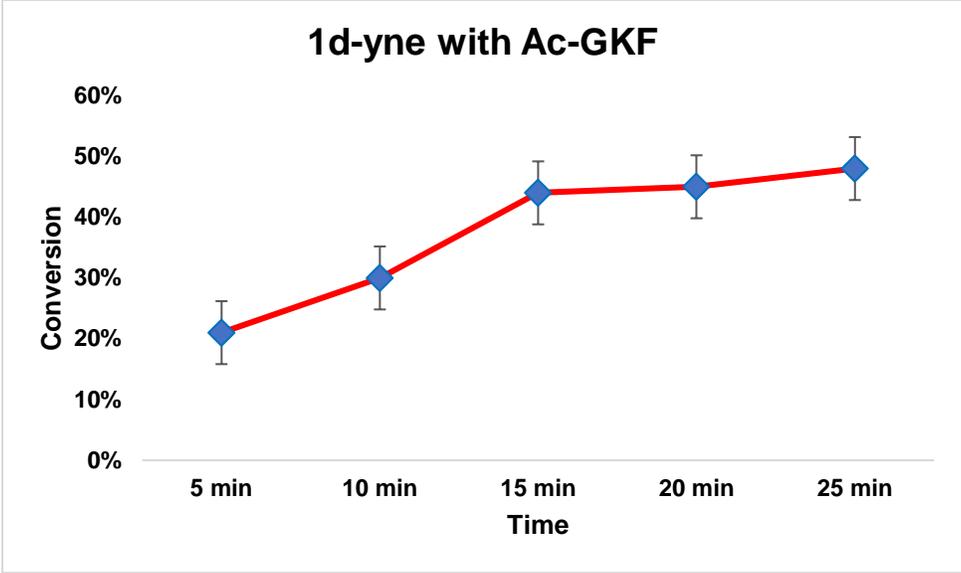


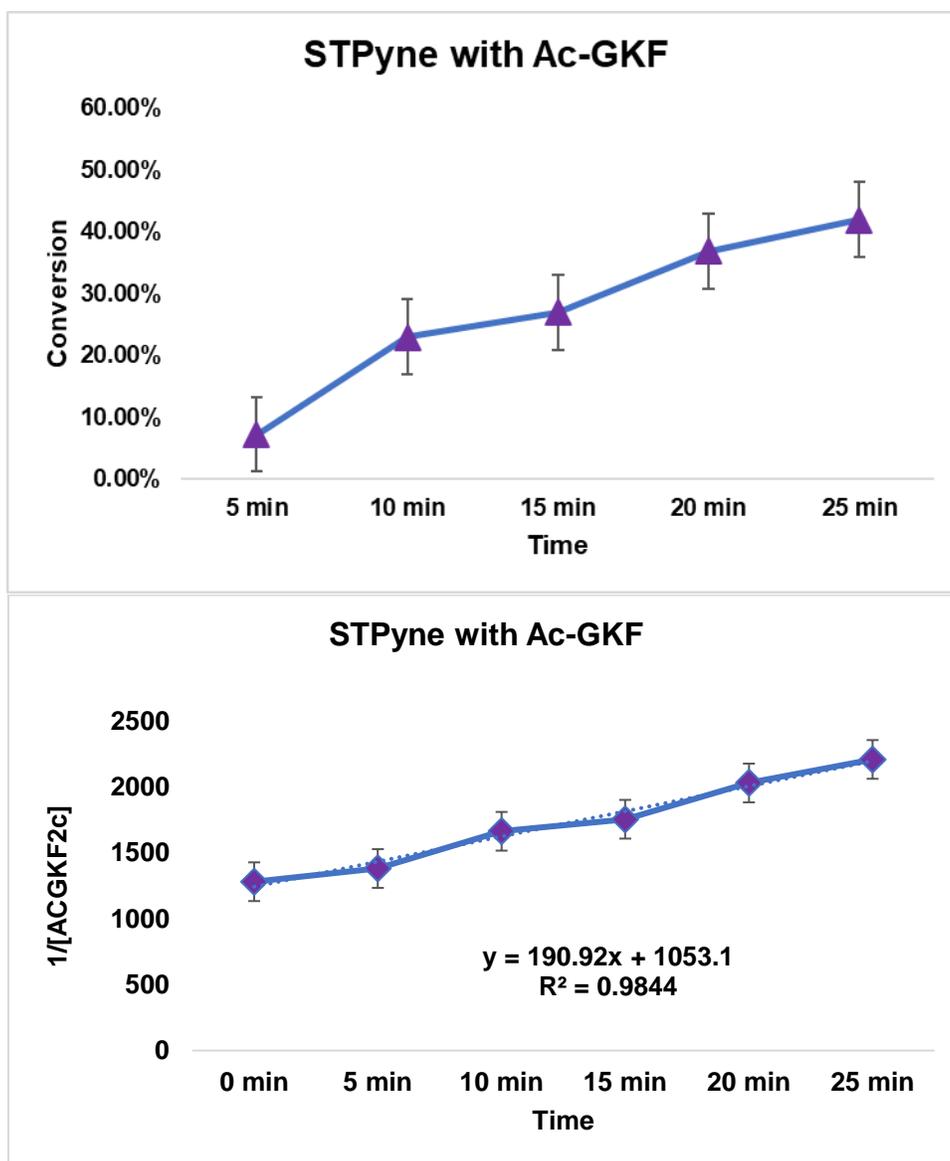
Supplementary Figure 24. Rate study of 1c-yne and 1d-yne with peptide Ac-GKF 2c.

To a solution of Ac-GKF **2c** (0.56 mg, 0.0014 mmol) in 2.0 mL of 10 mM phosphate buffer (Nap, pH 7.5) was added **1c-yne** or **1d-yne** (1 equiv., 0.0014mmol). For analysis, a sample (200 μ L) was taken from the mixture after regular intervals of time and quenched by freezing the sample at -80 $^{\circ}$ C. The frozen samples were then lyophilized and dissolved in 100 μ L of 1:1 H₂O/ACN and injected immediately into the HPLC for determining the % conversion to the modified product (X Terra C18 column {5 μ m} with a gradient of 0 to 80 % MeCN with 0.1 % formic acid in 30 min). The rate study was done in triplicate. We use the average of three trials to plot the rate curve. 0 min sample was taken immediately after addition of all the reagents of the

bioconjugation reaction. The result showed that 1c-yne, 1d-yne, STPyne bioconjugate reactions are second order reaction with $k = 99.27 \text{ M}^{-1}\text{S}^{-1}$, $307.52 \text{ M}^{-1}\text{S}^{-1}$, and $190.92 \text{ M}^{-1}\text{S}^{-1}$ respectively.





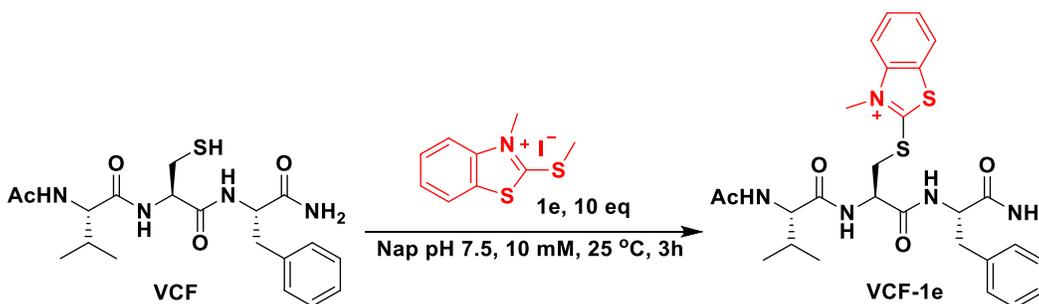


Supplementary Figure 25. Stability study of 1c-yne, 1d-yne, NHS ester and STPyne

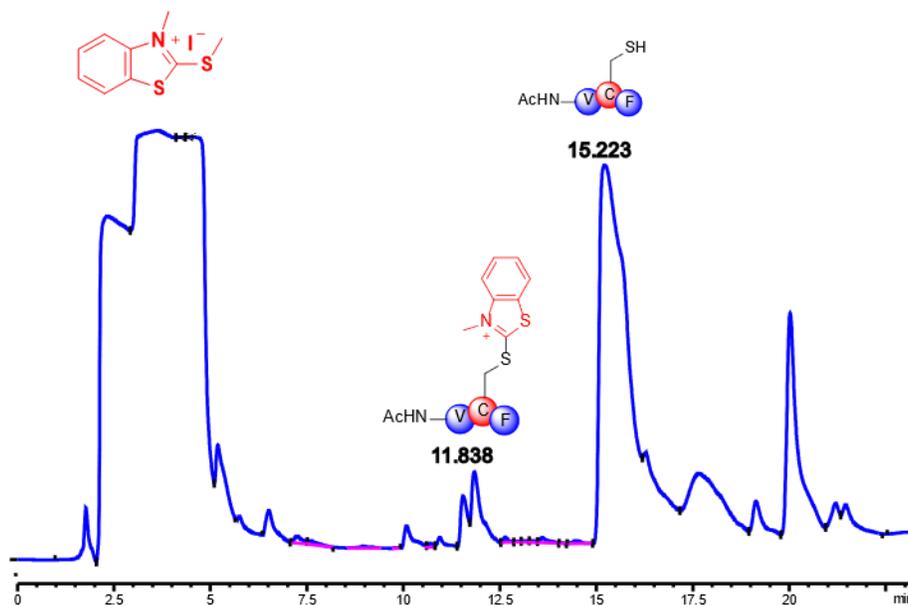
1c-yne, 1d-yne, NHS ester, and STPyne (0.035 mmol) were incubated in 400 μ L of 10 mM Nap (pH 7.5) at room temperature. A sample (50 μ L) was taken from the mixture and directly injected into HPLC. The reaction was monitored by injecting samples in HPLC after regular intervals of time 2 h, 4h, 6 h, 8h, 10h, and 12 h.

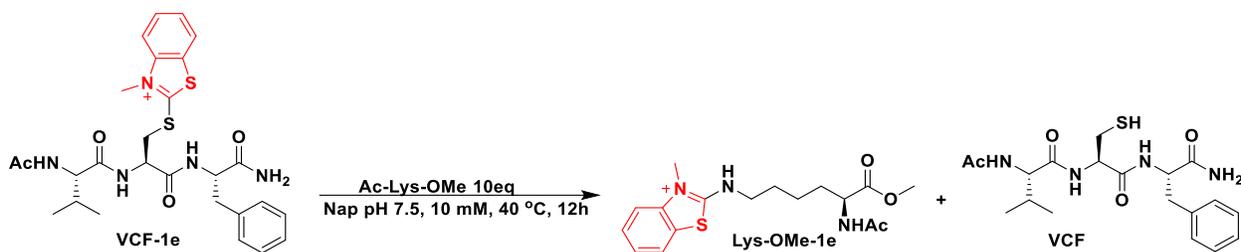
Supplementary Figure 26. Reversibility study of 1e-Cys-conjugate (VCF-1e) and 1d-thio-conjugate with lysine methylester.

Ac-VCF-NH₂ (6.3 mM) was dissolved in 0.4 mL of 10 mM phosphate buffer (Nap, pH 7.5), and then 10 equiv. of 1e was added. The mixture was stirred at room temperature for 3 hours. The reaction mixture was purified by HPLC then dried by lyophilized to afford pure VCF-1e. To a solution of pure VCF-1e (6.3 mM) in Nap buffer pH 7.5, 400 μ L, 10 eq of lysine methyl ester was added. The reaction was incubated at 40 °C for 12 h. The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

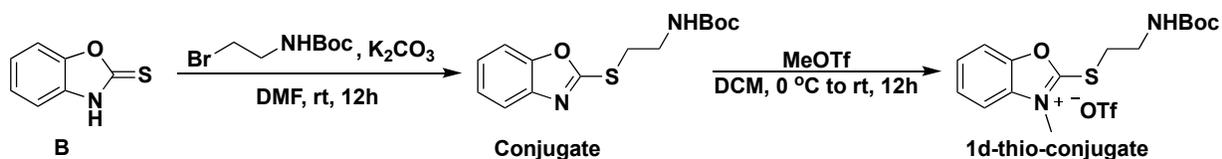
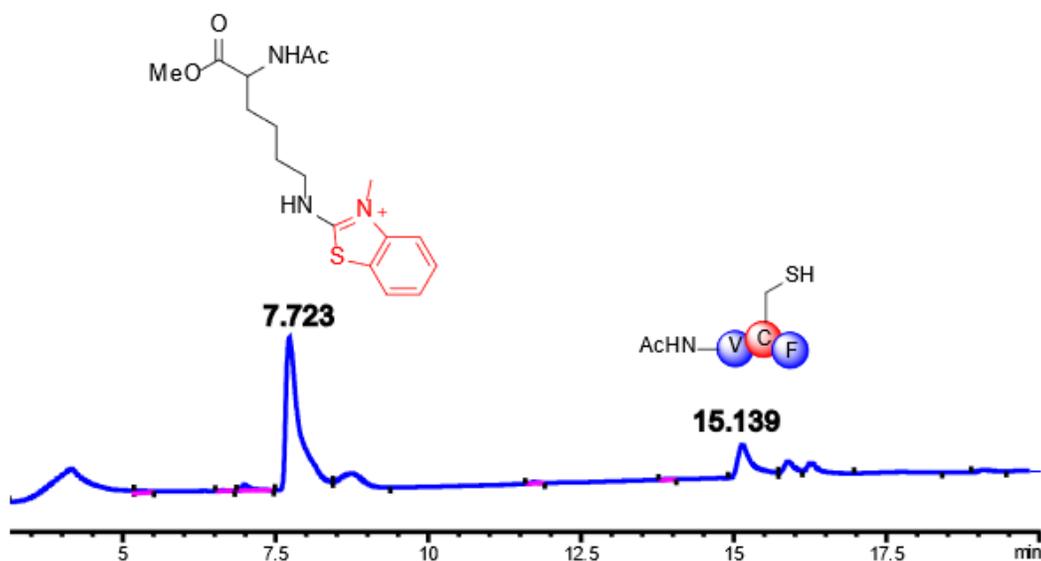


HPLC trace of modified VCF by probe 1e





HPLC trace after incubating with Ac-Lys-OMe for 12 h at 40 °C

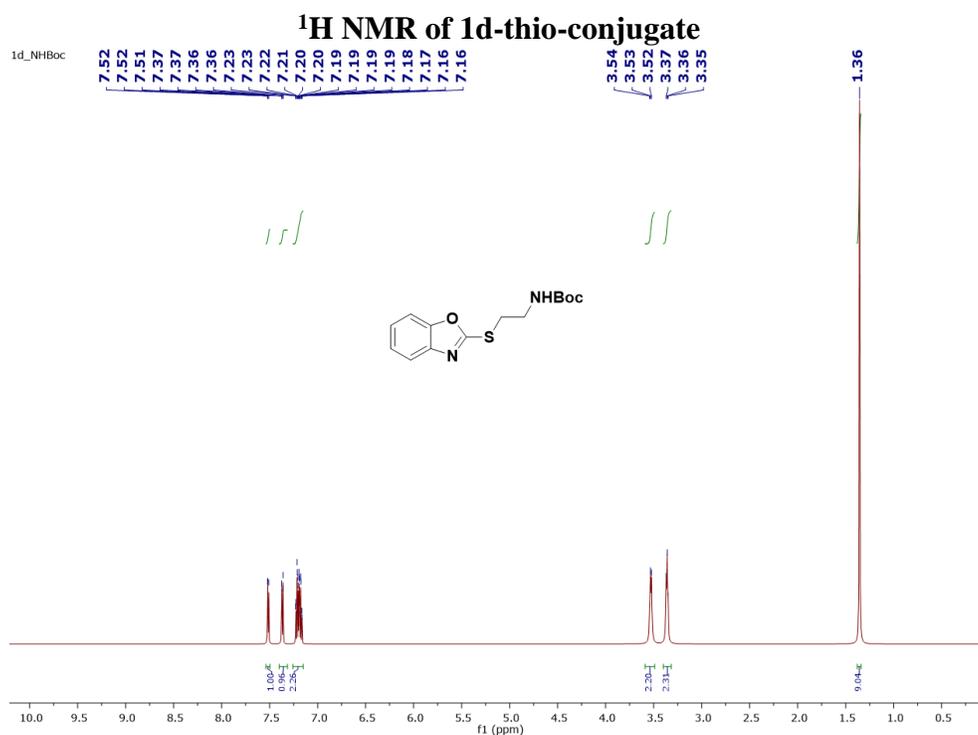


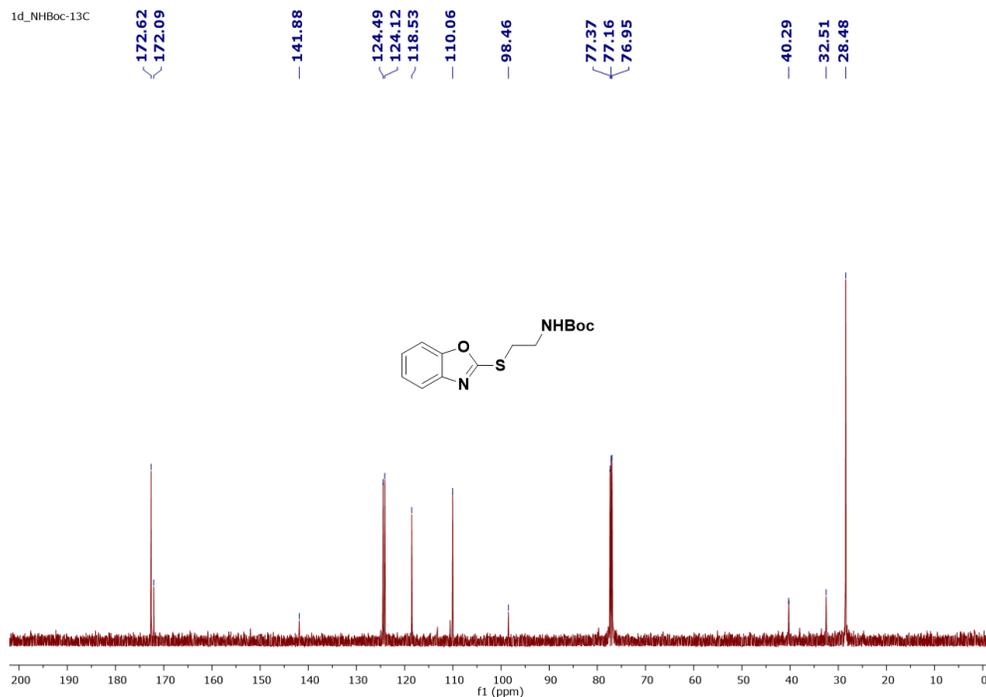
Synthesis of 1d-thio-conjugate

2-Mercaptobenzoxazole **B** (500 mg, 3.31 mmol), potassium carbonate (456 mg, 3.31 mmol), and tert-butyl (2-bromoethyl)carbamate (738.4mg, 3.31 mmol) were sequentially added into dry DMF (10 mL) at 0 °C. The mixture was warmed to room temperature and stirred for 12 hours. The reaction mixture was washed with ethyl acetate and brine for 3 times. The organic layer was collected, dried over anhydrous MgSO_4 , filtered, and concentrated under the reduced pressure.

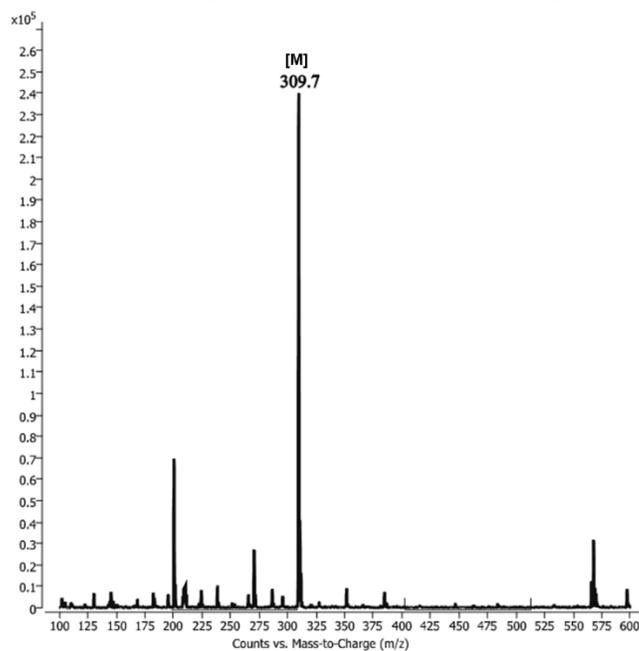
The residue was purified by the column chromatography (hexane: ethyl acetate 3:1) to yield conjugate as colorless crystal (739.6 mg, 76 %). **¹H NMR** (600 MHz, CDCl₃) δ 7.54 – 7.50 (m, 1H), 7.37 (dd, *J* = 7.9, 1.1 Hz, 1H), 7.26 – 7.15 (m, 2H), 3.59 – 3.49 (m, 2H), 3.36 (t, *J* = 6.2 Hz, 2H), 1.36 (s, 9H). **¹³C NMR** (151 MHz, CDCl₃) δ 172.62, 172.09, 141.88, 124.49, 124.12, 118.53, 110.06, 98.46, 40.29, 32.51, 28.48.

To a mixture of **conjugate** (93 mg, 0.316 mmol) in dry DCM (10 mL), MeOTf (41.8 μL, 0.379 mmol) was added at 0 °C. The reaction mixture was warmed to room temperature. After 12 h, solvent was removed by rotary evaporation to afford brown oil product. The **1d-thio-conjugate** was directly used without purification.





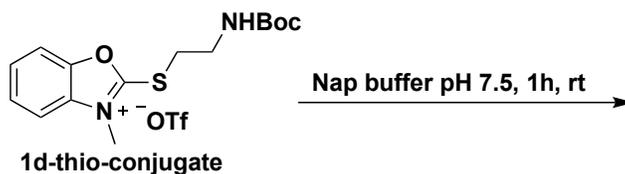
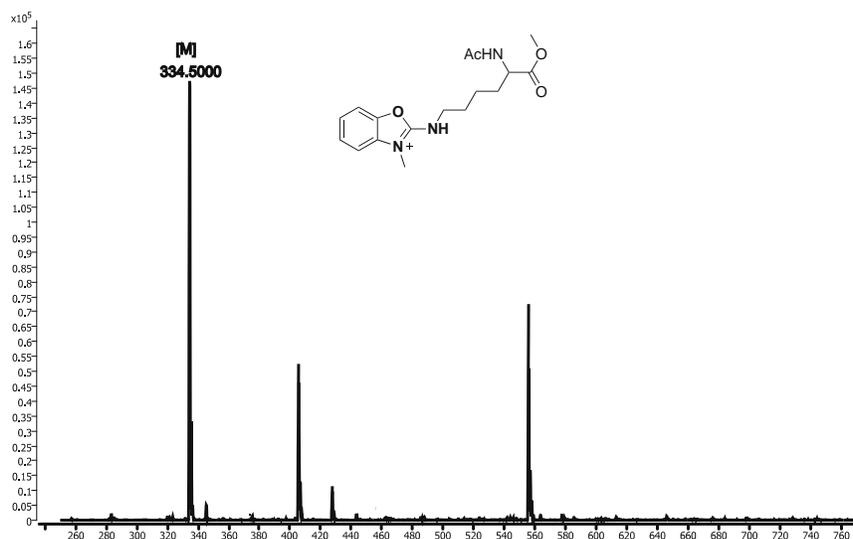
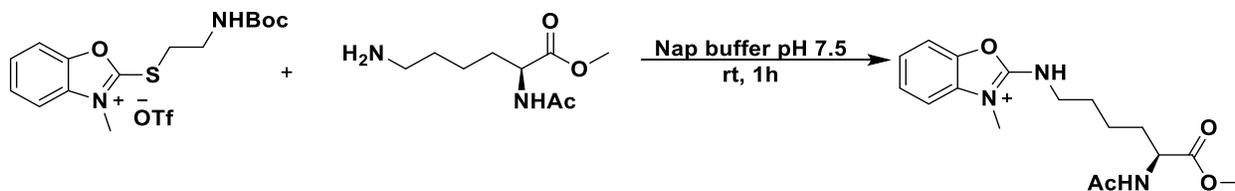
LC-MS spectrum of 1d-thio-conjugate



Reversibility study of 1d-thio-conjugate with lysine methylester.

General procedure for reversibility study of 1d-thio-conjugate: Ac-Lys-OMe (7.23 mg, 0.030 mmol, 10 eq) was dissolved in 0.4 mL of 10 mM phosphate buffer (Nap, pH 7.5), and then

1 equiv. of **1d-thio-conjugate** was added. The mixture was stirred at room temperature for 1 hour. The reaction was analyzed by LC-MS.

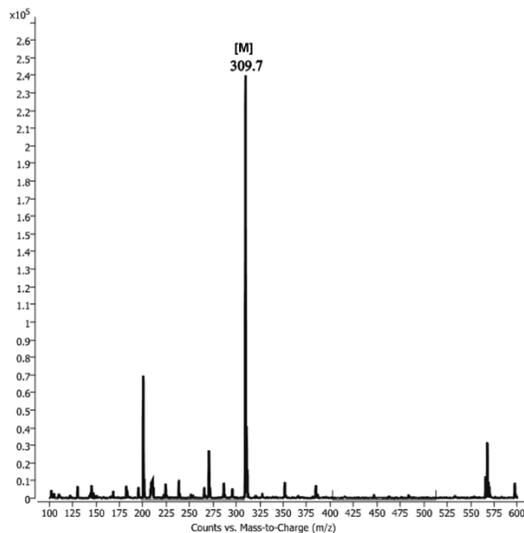


Procedure for stability study of 1d-thio-conjugate in sodium phosphate buffer.

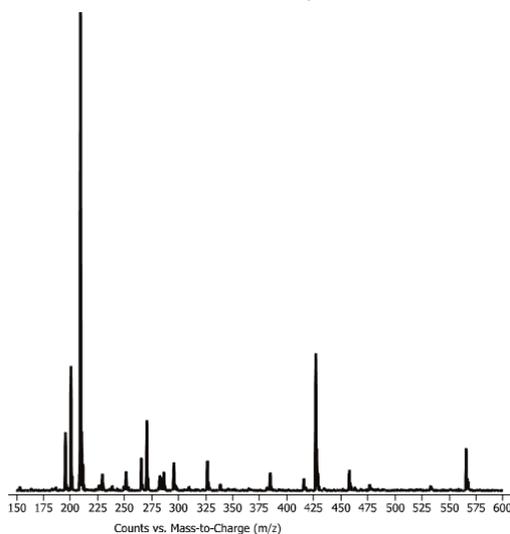
1d-thio-conjugate (0.035 mmol) was incubated in 400 μ L of 10 mM Nap (pH 7.5) at room temperature. A sample (10 μ L) was taken from the mixture and directly injected into LC-MS.

The stability study was monitored by LC-MS.

LC-MS of pure 1d-thio-conjugate at 0 min



LC-MS of pure 1d-thiol conjugate after 1 hour (Compound was completely decomposed in one hour)



Supplementary Figure 27. Chemoproteomic analysis of residue selectivity

Cell culture and preparation of cell lysates.

Cell culture reagents including Dulbecco's phosphate-buffered saline (DPBS), Dulbecco's modified Eagle's medium (DMEM)/high glucose media, trypsin-EDTA and penicillin/streptomycin (Pen/Strep) were purchased from Fisher Scientific. Fetal Bovine Serum (FBS) were purchased from Avantor Seradigm (lot # 214B17).

HEK293T (ATCC: CRL-3216) cells were cultured in DMEM supplemented with 10% FBS and 1% antibiotics (Penn/Strep, 100 U/mL). Media was filtered (0.22 μm) prior to use. Cells were maintained in a humidified incubator at 37 $^{\circ}\text{C}$ with 5% CO_2 . Cell lines were validated prior to use and tested regularly for myoplasma.

HEK293T cells were harvested once cells were grown to 90 – 95% confluence by centrifugation (4500g, 5 min, 4 $^{\circ}\text{C}$), washed twice with cold DPBS, resuspended in 300 μL DPBS, sonicated, and clarified by centrifuging (21,000 g, 10 min, 4 $^{\circ}\text{C}$). The lysates were then transferred to an eppendorf tube. Protein concentrations were determined using a Bio-Rad DC protein assay kit using reagents from Bio-Rad Life Science (Hercules, CA) and the lysate diluted to the working concentrations indicated below.

Gel-based ABPP with 1d. HEK293T proteome (50 μL of 1.5 mg/mL, prepared as described above) was labeled with various concentrations of **1d** (stock solutions in DMSO, final concentration as indicated), STPyne (1 μL of 5 mM stock solution in DMSO, final concentration = 100 μM) or DMSO for vehicle control for 1h at ambient temperature followed by adding 1 μM NHS-Rh or IA-Rh. Samples were allowed to react for another hour at ambient temperature at which point the reactions were quenched with 4 \times Laemmli buffer (20 μL). Samples were then denatured (5 min, 95 $^{\circ}\text{C}$) and then resolved by SDS-PAGE. SDS-PAGE gels were imaged on the Bio-Rad ChemiDocTM Imager using rhodamine channel.

Gel-based ABPP with 1c-yne. HEK293T proteome (50 μL of 1.5 mg/mL) was labeled with different amounts of **1c-yne** (stock solutions in DMSO, final concentration as indicated) for 1h at ambient temperature followed by adding 1 μM NHS-Rh or IA-Rh. Samples were allowed to react for another hour at ambient temperature at which point the reactions were quenched with 4 \times Laemmli buffer (20 μL). Samples were then denatured (5 min, 95 $^{\circ}\text{C}$) and then resolved by

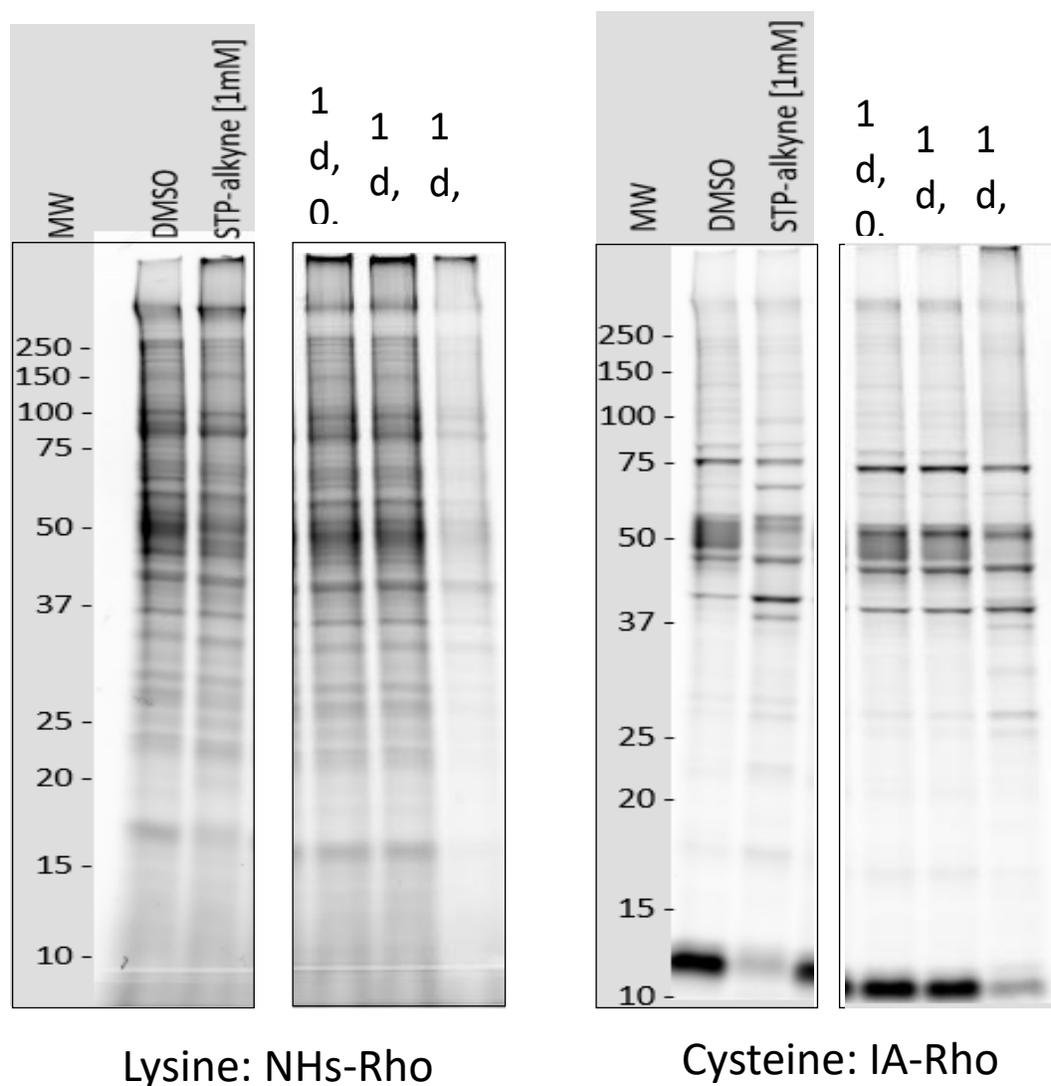
SDS-PAGE. SDS-PAGE gels were imaged on the Bio-Rad ChemiDoc™ Imager using rhodamine channel.

Gel-based ABPP with CuAAC using 1c-yne. HEK293T proteome (50 μ L of 1.5 mg/mL) was labeled with IA-alkyne (**IA-aky**) (1 μ L of 5 mM stock in DMSO, final concentration = 100 μ M) or **1c-yne** (1 μ L of 50 mM stock in DMSO, final concentration = 1 mM) for 1h at ambient temperature followed by copper-mediated azide-alkyne cycloaddition (CuAAC). CuAAC was performed with biotin-azide (200 μ M for **IA-alkyne** treated sample and 2 mM for **1c-yne** treated sample), tris(2-carboxyethyl)phosphine hydrochloride (TCEP; 1 μ L of fresh 50 mM stock in water, final concentration = 1 mM), Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, 3 μ L of 1.7 mM stock in DMSO/t-butanol 1:4, final concentration = 100 μ M), and CuSO₄ (1 μ L of 50 mM stock in water, final concentration = 1 mM). Samples were allowed to react for 1h at ambient temperature at which point the reactions were quenched with 4 \times Laemmli buffer (20 μ L). Samples were then denatured (5 min, 95 °C) and analyzed by SDS-PAGE, using Criterion™ TGX Stain-Free™ gels obtained from Bio-Rad. Loading control images were obtained using the stain-free workflow on a Bio-Rad ChemiDoc™ Imager.

Gel-based ABPP with CuAAC using 1d-yne. HEK293T proteome (50 μ L of 1.5 mg/mL) was labeled with different concentrations of **1d-yne** or **STPyne** (1 μ L of 5 mM stock solution in DMSO, final concentration = 100 μ M) for 1h at ambient temperature followed by copper-mediated azide-alkyne cycloaddition (CuAAC) as described above and the labeling was resolved by SDS-PAGE.

Streptavidin blot. Gels were transferred to either polyvinylidene difluoride (PVDF, Bio-Rad, 1620177) or nitrocellulose (Bio-Rad, 1704271) membranes using a Trans-Blot Turbo transfer system (Bio-Rad) following the manufacturer's instructions. After transfer, the membranes were

then blocked (2% w/v of BSA in TBS-T, 30 min) and probed with streptavidin-IRDye® 800CW (Fisher, NC0883593, 1:4000) in TBS-T. Blots were incubated overnight at 4 °C with rocking and were then washed (3 × 5 min, TBS-T). The membranes were then imaged with a Bio-Rad ChemiDoc™ Imager using the 800 NIR channel.



Supplementary Figure 28. TAREs for gel-based ABPP

Proteomic sample preparation using 1c-yne. HEK293T proteome (200 μ L of 2 mg/mL, prepared as described above) was labeled with **1c-yne** (10 μ L of 10 mM stock solution in DMSO, final concentration = 500 μ M) for 1h at ambient temperature. CuAAC was performed

with biotin-azide (4 μL of 50 mM stock in DMSO, final concentration = 1 mM), TCEP (4 μL of fresh 50 mM stock in water, final concentration = 1 mM), TBTA (12 μL of 1.7 mM stock in DMSO/t-butanol 1:4, final concentration = 100 μM), and CuSO_4 (4 μL of 50 mM stock in water, final concentration = 1 mM). Samples were allowed to react for 1h at ambient temperature. The samples were then subjected to SP3 sample preparation and LC-MS/MS analysis, as described below.

SP3 proteomic sample preparation. After CuAAC labeling, each sample was then treated with 0.5 μL benzonase (Fisher Scientific, 70-664-3) for 30 min at 37 $^\circ\text{C}$. DTT (10 μL of 200 mM stock in water, final concentration = 10 mM) was added into each sample and the sample was incubated at 65 $^\circ\text{C}$ for 15 min. To this iodoacetamide (10 μL of 400 mM stock in water, final concentration = 20 mM) was added and the solution was incubated for 30 min at 37 $^\circ\text{C}$ with shaking. SP3 sample cleanup was performed at a bead/protein ratio of 10:1 (wt/wt). For each 200 μL sample (1 mg/mL protein concentration), 20 μL Sera-Mag SpeedBeads Carboxyl Magnetic Beads, hydrophobic (GE Healthcare, 65152105050250, 50 $\mu\text{g}/\mu\text{L}$, total 2 mg) and 40 μL Sera-Mag SpeedBeads Carboxyl Magnetic Beads, hydrophilic (GE Healthcare, 45152105050250, 50 $\mu\text{g}/\mu\text{L}$, total 1 mg) were aliquoted into a single microcentrifuge tube and gently mixed. Tubes were then placed on a magnetic rack until the beads settled to the tube wall, and the supernatants were removed. The beads were removed from the magnetic rack, reconstituted in 400 μL of MB water, and gently mixed. Tubes were then returned to the magnetic rack, beads allowed to settle, and the supernatants removed. Washes were repeated for two more cycles, and then the beads were reconstituted in 80 μL MB water. The bead slurries were then transferred to CuAAC samples, and incubated for 5 min at RT with shaking. Absolute ethanol (400 μL) was added to each sample, and the samples were incubated for 5 min at RT with shaking. Samples were then

placed on a magnetic rack, and beads allowed to settle. Supernatants were then removed and discarded. Using the magnetic rack as described above, the beads were further washed three times with ethanol (400 μ L of 80% solution in water). Beads were then resuspended in 100 μ L PBS containing 2 M urea followed by addition trypsin solution (Worthington Biochemical, LS003740, 5 μ L, 1 mg/mL in 666 μ L of 50 mM acetic acid and 334 μ L of 100 mM CaCl_2 , final weight = 10 ng). Digest was allowed to proceed overnight at 37 $^\circ\text{C}$ with shaking. The samples were then acidified to a final concentration of 3% (v/v) FA. After incubation for 5 min at RT with shaking, \sim 2 mL acetonitrile ($>$ 95% of the final volume) was added to each sample and the mixtures were then incubated for an additional 10 min at RT with shaking. Supernatants were then removed and discarded using the magnetic rack, and the beads were washed ($3 \times 500 \mu\text{L}$ acetonitrile). Peptides were then eluted from SP3 beads with 100 μ L of 2% DMSO in MB water for 30 min at 37 $^\circ\text{C}$ with shaking. The elution will be used for NeutrAvidin enrichment.

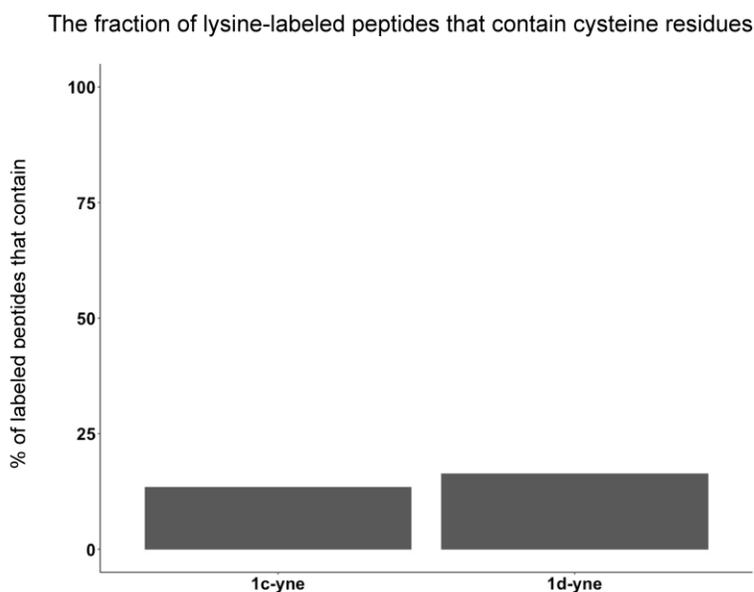
NeutrAvidin enrichment of labelled peptides. For each sample, 50 μ L of NeutrAvidin[®] Agarose resin slurry (Pierce, 29200) was washed twice in 10 mL IAP buffer (50 mM MOPS pH 7.2, 10 mM sodium phosphate, and 50 mM NaCl buffer) and then resuspended in 500 μ L IAP buffer. Peptide solutions eluted from SP3 beads were then transferred to the NeutrAvidin[®] Agarose resin suspension, and the samples were then rotated for 2h at RT. After incubation, the beads were pelleted by centrifugation (21,000 g, 1 min) and washed by centrifugation ($6 \times 700 \mu\text{L}$ water). Bound peptides were eluted with 60 μ L of 80% acetonitrile in MB water containing 0.1% FA (10 min at RT). The samples were then harvested by centrifugation (21,000 g, 1 min) and residual beads separated from supernatants using Micro Bio-Spin columns (Bio-Rad). The remaining peptides were then eluted from pelleted beads with 60 μ L of 80% acetonitrile in water containing 0.1% FA (10 min, 72 $^\circ\text{C}$). Beads were then separated from the eluants using the same

Bio-Spin column. Eluants were then collected by centrifugation (21,000 g, 1 min) and the combined eluants were dried (SpeedVac). The samples were then reconstituted in 40 μ L water containing 5% acetonitrile and 1% FA and analyzed by LC-MS/MS.

Liquid-chromatography tandem mass-spectrometry (LC-MS/MS) analysis. The samples were analyzed by liquid chromatography tandem mass spectrometry using an Q ExactiveTM mass spectrometer (Thermo Scientific) coupled to an Easy-nLCTM 1000 pump. Peptides were resolved on a C18 reversed phase column (3 μ M, 100 \AA pores), packed in-house, with 100 nm internal diameter and 18 cm of packed resin. The peptides were eluted using a 140 min gradient of Buffer B in Buffer A (Buffer A: water with 3% DMSO and 0.1% FA; Buffer B: acetonitrile with 3% DMSO and 0.1% FA) and a flow rate of 220 nL/min with electrospray ionization of 2.2 kV. The regular gradient includes 0 – 5 min from 1% to 5%, 15 – 130 min from 5% to 27%, 15 – 137 min from 27% to 35%, and 137 – 138 min from 35% to 80% buffer B in buffer A. The steep gradient for this study includes 0 – 5 min from 1% to 5%, 5 – 20 min from 5% to 15%, 20 – 130 min from 15% to 35% and from 130 – 135 min from 35 to 95% buffer B in buffer A. The detailed gradient includes 0 – 15 min from 1% to 15%, 15 – 110 min from 15% to 35% and from 130 – 135 min from 35 to 95% buffer B in buffer A. Data was collected in data-dependent acquisition mode with dynamic exclusion (15 s), and charge exclusion (1,7,8,>8) was enabled. Data acquisition consisted of cycles of one full MS scan (400 – 1800 m/z at a resolution of 70,000) followed by 12 MS² scans of the nth most abundant ions at resolution of 17,500.

Peptide and protein identification. The MS² spectra data were extracted from a raw file using RAW Xtractor (version 1.1.0.22; available at <http://fields.scripps.edu/rawconv/>). MS² spectra data were searched using the ProLuCID algorithm (publicly available at http://fields.scripps.edu/yates/wp/?page_id=17 using a reverse concatenated, nonredundant

variant of the Human UniProt database (release-2020_01). Lysine or cysteine residues were searched with a variable modification for carboxyamidomethylation (+57.02146) and an additional variable modifications at either lysine or cysteine residues, which is +526.2231 for probe 1d-yne and +483.19213 for 1c-yne. Peptides were required to have at least one tryptic terminus and to contain the biotin modification. ProLuCID data was filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%. The Xcorr score was used for match confidence criteria. Mass tolerance of the peptide precursor was set to 50 ppm. The built in localization features on IP2 were used predict the PTM index.



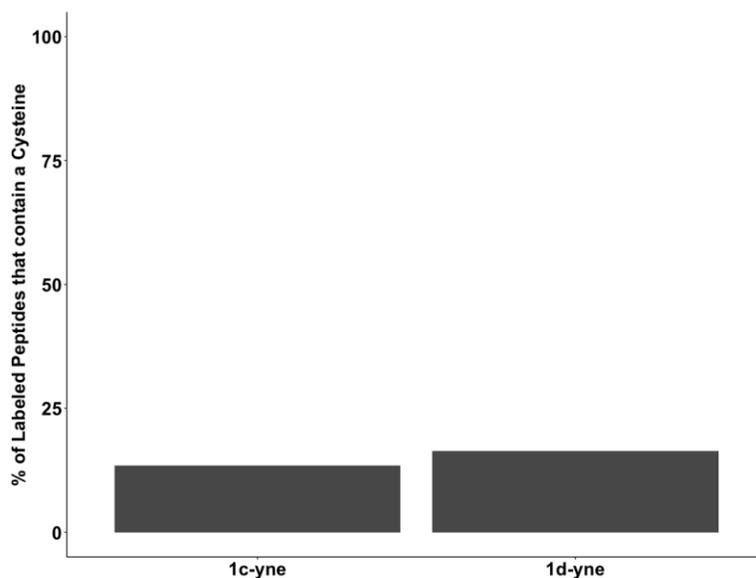
1-cyne datasets were also searched using MSFragger v15 with a precursor mass window of 50 ppm and a fragment mass tolerance of 20 ppm and variable modifications of +483.19213 on K and C as well as a variable modification for carbamidomethylation on cysteine. The options --decoyprobs --ppm --accmass --nonparam --expectscore were used for PeptideProphet and --static --em 1 --nions b --mods C:483.19213, K:483.19213 --minprob 0.5 for PTMProphet.

Distribution of labeling sites identified for 1c-yne by MSFragger, using PTMProphet to score modification sites

Identified	Counts
Cysteine Protein IDs	17
Cysteine Peptide IDs	20
Lysine Protein IDs	1560
Lysine Peptide IDs	4460

Lysine selectivity of 1c-yne and 1d-yne. To further analyze the lysine-selective labeling, we conducted an analysis of the amino acid content of all peptides identified as labeled by the TARE probes. The probability score was used for match confidence criteria. Mass tolerance of the peptide precursor was set to 50 ppm. In addition to using the built in localization features on IP2 to predict the PTM index, to further analyze the lysine-selective labeling, we conducted an analysis of the amino acid content of all peptides identified as labeled by the TARE probes. Gratifyingly <15% of the labeled peptides contained one or more cysteine residues, which again supports preferential labeling for lysine residues.

To further investigate the specificity of the modification, we additionally reprocessed the 1c-yne data using MSFragger using the built in PTMProphet tool to improve our confidence in the localization of the modifications. This analysis revealed that >95% of all labeled residues are lysines, which is consistent with our initial analysis using the ProLuCID/IP2 search.

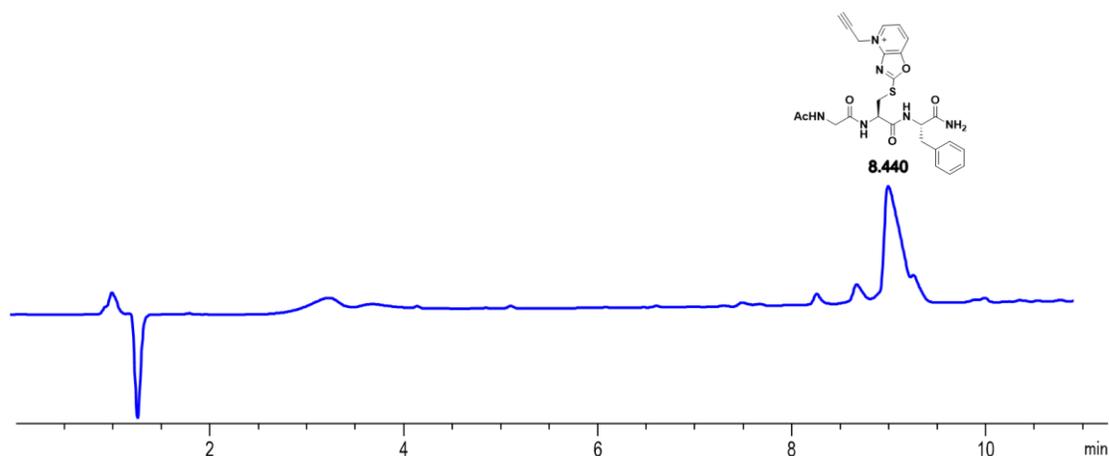


Supplementary Figure 29. Stability study of GCF-1c-yne conjugate in TCEP, and 1d-thio-conjugate under sodium phosphate buffer.

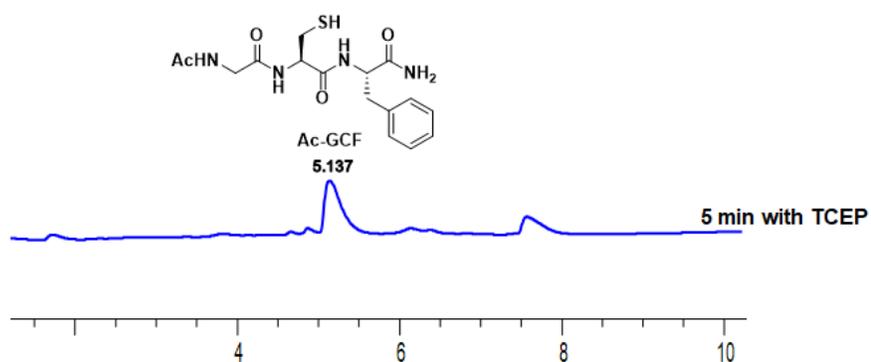
Procedure for stability study of GCF-1c-yne conjugate in TCEP buffer

Ac-GCF-NH₂ (6.3 mM) was dissolved in 0.4 mL of 10 mM phosphate buffer (Nap, pH 7.5), and then 10 equiv. of 1c-yne was added. The mixture was stirred at room temperature for 3 hours. The reaction mixture was purified by HPLC then dried by lyophilization to afford pure modified peptide, GCF-1c-yne conjugate. To a solution of pure GCF-1c-yne conjugate (6.3 mM) in Nap buffer (pH 7.5, 400 μL), TCEP buffer (pH 7.4, 200 μL) was added. The reaction was incubated at 25 °C for 5 min. The reaction was analyzed by HPLC and ESI-MS. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

HPLC trace of pure GCF-1c-yne conjugate



After incubation with TCEP buffer for 5 min



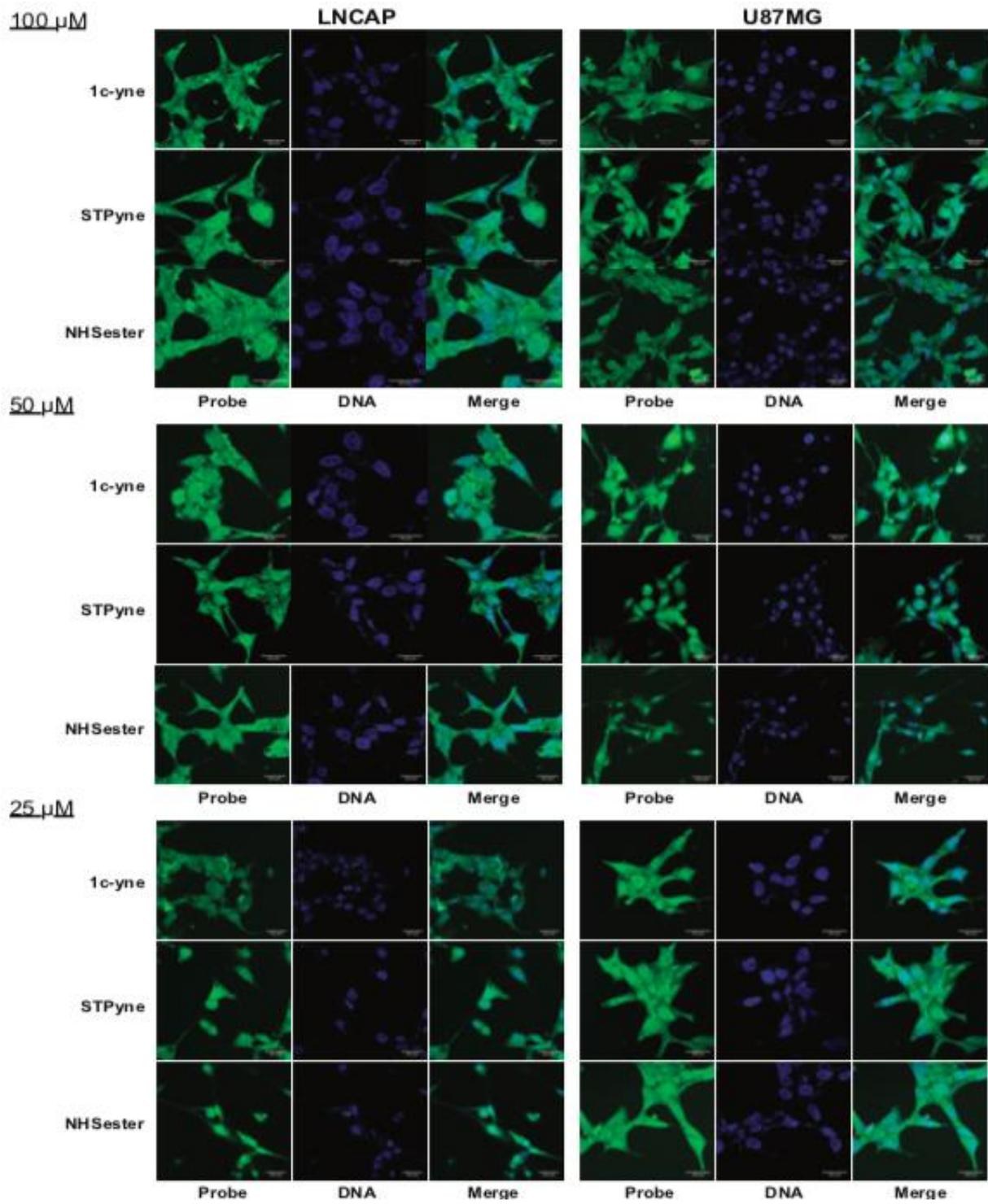
Supplementary Figure 30. Live cell labeling of proteins using different probes.

Cell Culture – Cells were maintained at 37° C and 5% CO₂. T47D and LNCaP cells were cultured in RPMI supplemented with 10% (V/V) fetal bovine serum (FBS) and 1% (V/V) penicillin/streptomycin (100 µg/mL). U87MG cells were cultured in DMEM supplemented with 10% (V/V) fetal bovine serum (FBS) and 1% (V/V) penicillin/streptomycin (100 µg/mL).

Copper Azide-Alkyne Cycloaddition (CuAAC) reaction – CuAAC on fixed cells was performed using 4 mM CuSO₄, 8 mM THPTA, 75 µM picolyl azide-conjugated fluorophore (ClickChemistryTools), and 10 mM sodium ascorbate. For confocal microscopy, cells were fixed

in 3.7% PFA and permeabilized with 0.5% Triton-X. CuAAC reagents were added directly to cells, then incubated, rocking, for 45 min.

Confocal microscopy – LNCaP, T47D, and U87MG cells were plated on glass coverslips (Fisherbrand) in supplemented RPMI or DMEM media. After 16 h, cells were incubated with the indicated concentrations of **1c-yne**, **STPyne**, or **NHS-ester**. Prior to imaging, cells were fixed in 3.7% PFA and permeabilized with 0.5% Triton-X. CuAAC reaction was performed to attach a 488 nM picolyl azide-conjugated fluorophore (ClickChemistryTools) and Hoechst counterstain was used to image nuclei. Cells were imaged on Leica SP8 confocal microscope and images processed and analyzed using ImageJ.



Live cell Labeling using TAREs and other probes at different concentrations. LNCAP and U87MG cells treated with 25-100 μ M 1c-yne, STPyne, or NHSester for 2h show labeling in multiple cellular compartments.

Supplementary Figure 31. Cell viability studies with 1c-yne.

Annexin V/PI staining of T47D cells treated with increasing concentrations of 1c-yne.

To assay for cell death, T47D cells were seeded and treated with the indicated concentration (5 μ M and 20 μ M) of **1c-yne** for 24hrs, after which cells were detached using trypsin (**Biolegend**) and stained using Annexin V/PI following manufacturer's protocol (BioLegend). Cells were analyzed by flow cytometer (BDFACSymphony A3) within 1h. The reaction was done in triplicates. T47D cells treated with indicated concentrations of **1c-yne** for 24h did not show an increase in apoptosis/necrosis compared to DMSO control.

Supplementary Figure 32. General Computation Procedure

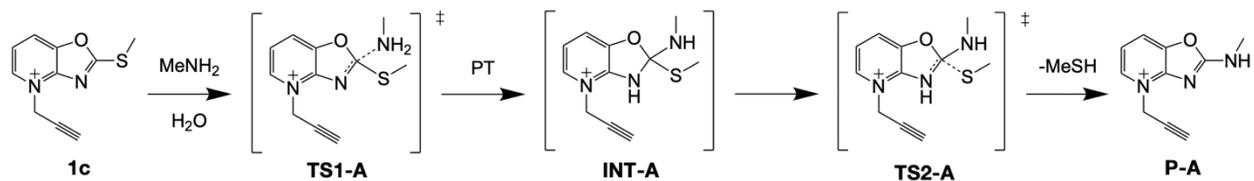
Quantum Mechanical Calculations

All conformational searches were carried out using CREST, Conformer-Rotamer Ensemble Sampling Tool version 2.7.1, of the XTB program version 6.2 RC2 (SAW190805).^{4,5} The RMSD threshold for each conformational search was set to 0.5 Å. Density functional theory calculations were performed using the Gaussian 09 software package.⁶ Ground state and transition state structures were optimized at the

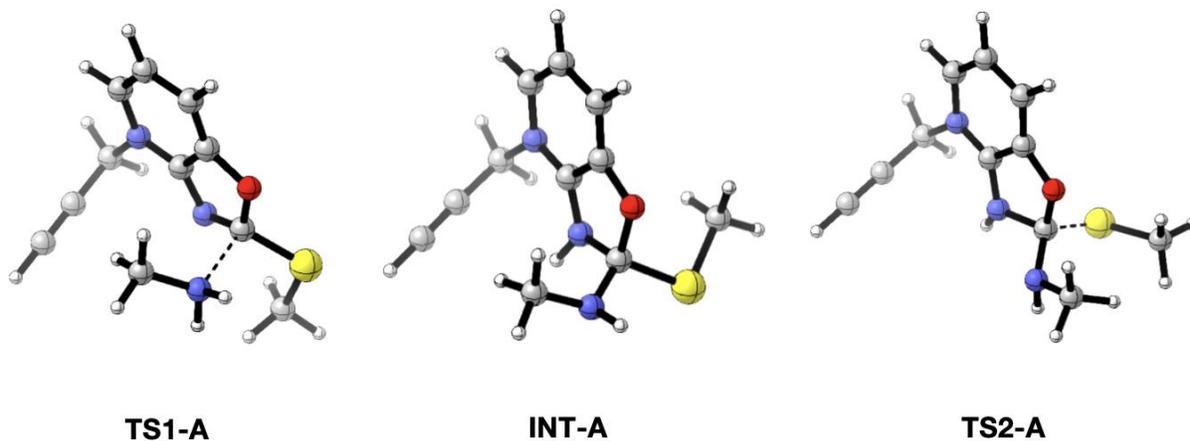
B3LYP/6-31+G(d) level with SMD solvation model for water, the D3 version of Grimme's empirical dispersion correction, and the integration grid set to ultrafine.⁷ Frequency calculations were carried out at the same level of theory. The GoodVibes program was used for quasi-harmonic correction of Gibbs free energies at 298 K.⁸ Single point energy calculations of the

optimized geometries were performed at ω B97X-D/6-311++G(d,p) level with SMD solvation model for water and the integration grid set to ultrafine.

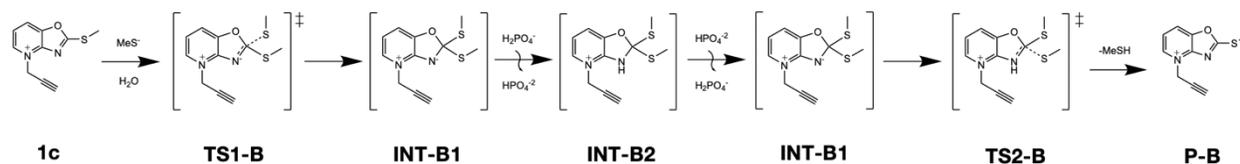
Proposed S_NAr mechanism of TARE probe 1c with methylamine



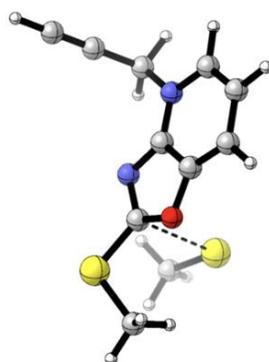
Optimized structures of methylamine calculations at ω B97X-D/6-311++G(d,p) SMD(H_2O)/B3LYP/6-31+G(d) SMD(H_2O) level



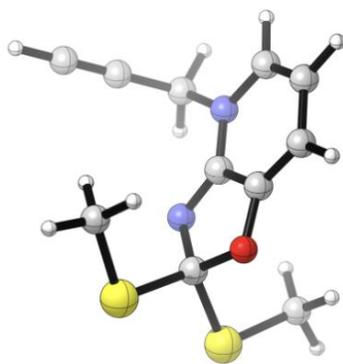
Proposed S_NAr mechanism of TARE probe 1c with methyl thiolate



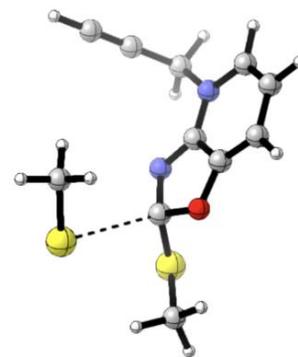
Optimized structures of methyl thiolate calculations at ω B97X-D/6-311++G(d,p) SMD(H_2O)/B3LYP/6-31+G(d) SMD(H_2O) level



TS1-B



INT-B1



TS2-B

Cartesian Coordinates of Calculated Structures

MeNH2

N	-0.05479700	-0.76598500	0.00000000
C	-0.05479700	0.70827600	0.00000000
H	-0.58856500	1.07163500	0.88380800
H	-0.58856500	1.07163500	-0.88380800
H	0.94960600	1.15772000	0.00000000
H	0.46994500	-1.09437700	0.81100100
H	0.46994500	-1.09437700	-0.81100100

TS1-A

C	2.97217400	-0.83487500	-0.49303600
C	2.72513100	-2.08376000	0.02874800
C	1.40295200	-2.50588500	0.32754000
C	0.42469600	-1.59336100	0.05408800
C	0.68089700	-0.31355100	-0.48388500
N	1.95449000	0.05041200	-0.75980500
C	2.23960100	1.42570900	-1.24291900
C	2.14136200	2.40450300	-0.16104600
C	2.04472400	3.20628100	0.73877500
H	1.96146800	3.91667100	1.53662300

H	3.24184800	1.41958600	-1.67733800
H	1.53066100	1.65426400	-2.04295000
N	-0.44887800	0.39351400	-0.66899900
C	-1.40387700	-0.39410400	-0.13528000
S	-3.08928200	-0.43946300	-0.67893900
C	-3.31769700	1.31370400	-1.14750700
H	-2.64477000	1.58606500	-1.96223800
H	-3.16651900	1.97044900	-0.28752500
H	-4.35418300	1.39465700	-1.48616200
O	-0.94210100	-1.66751800	0.22068700
H	1.19314900	-3.48400300	0.74623400
H	3.56393100	-2.74468800	0.21256600
H	3.97001600	-0.48377700	-0.72337000
N	-1.63973400	0.34659000	1.79778600
H	-2.25022400	-0.32875900	2.25716700
H	-2.14672800	1.22679700	1.70368800
C	-0.38118500	0.53548000	2.53105700
H	0.08898200	-0.43629100	2.69857600
H	0.28929200	1.16545400	1.93853300
H	-0.54542300	1.02068000	3.50163200

INT-A

C	2.02374300	-2.44089600	0.87542800
C	0.67675000	-2.27144800	1.29248600
C	0.00718300	-1.19729000	0.77247300
C	0.62976100	-0.31258100	-0.13140000
N	1.90610300	-0.48850800	-0.50907900
C	2.55499500	0.48877300	-1.42331200
C	2.68133900	1.80312400	-0.79640800
C	2.73457900	2.89170400	-0.27334000
H	2.79655100	3.85608900	0.18956400
H	3.53024200	0.08144400	-1.69630000

H	1.95172700	0.54955500	-2.33580400
N	-0.25337400	0.62968900	-0.50547100
C	-1.51624000	0.47217200	0.21523800
S	-2.95502100	0.17965600	-0.89684900
C	-2.43922200	-1.40014300	-1.65396000
H	-2.41084400	-2.20039700	-0.91147300
H	-1.46542600	-1.29675400	-2.14125900
H	-3.19082500	-1.63545700	-2.41300000
O	-1.26848500	-0.77742700	0.99506200
H	0.20701400	-2.95214500	1.99374200
H	2.60947100	-3.27180600	1.25099800
H	3.63764300	-1.64448100	-0.32782800
N	-1.80484100	1.59498400	1.02125600
C	-0.77713900	1.95023700	2.01813300
H	-1.15516200	2.79338000	2.59958000
H	0.13864800	2.26703600	1.51274400
H	-0.54293300	1.12379100	2.70070200
H	-2.71047100	1.46938800	1.47004600
H	-0.02725500	1.51367600	-0.94945800

TS2-A

C	-2.78678500	1.54957400	-0.14267600
C	-2.10354700	2.58989800	0.43095400
C	-0.73383100	2.46178100	0.78428800
C	-0.15591100	1.25079400	0.52855600
C	-0.87679000	0.19168600	-0.05527200
N	-2.16564600	0.34155200	-0.39972800
C	-2.91109200	-0.77868000	-1.03719100
C	-3.18175200	-1.86711700	-0.10072600
C	-3.39687400	-2.77857400	0.66366700
H	-3.59085800	-3.58609100	1.34088200
H	-3.83722500	-0.36398700	-1.43976200

H	-2.31769400	-1.13443000	-1.88687400
N	-0.07601500	-0.88920500	-0.13765400
C	1.30225100	-0.45942100	0.17049500
S	2.10268600	-0.24382200	-1.53447800
C	3.73835700	0.48599300	-1.15300500
H	3.64151900	1.27302200	-0.40084500
H	4.08875700	0.93862000	-2.08570400
H	4.45838900	-0.26864300	-0.83151000
O	1.12042500	0.82638800	0.78638500
H	-0.18055300	3.27814800	1.23474400
H	-2.62730700	3.52203600	0.60739700
H	-3.82742900	1.60968100	-0.43225000
N	1.93243700	-1.33151500	1.10872400
C	3.06470300	-0.79793300	1.89274000
H	2.22190200	-2.17613800	0.61753900
H	-0.26192700	-1.70630200	-0.71229400
H	2.73966300	0.07543800	2.45963400
H	3.35753000	-1.57340900	2.60415000
H	3.94104100	-0.53172900	1.29508000

P-A

C	2.38324100	-0.98451400	-0.31097600
C	2.01612100	-2.24292000	0.11390600
C	0.66373700	-2.55489000	0.39308500
C	-0.21817200	-1.52718400	0.20637200
C	0.16061600	-0.24431400	-0.22854100
N	1.45967500	0.01347000	-0.49161500
C	1.88840300	1.37451000	-0.91096700
C	2.08110600	2.26185100	0.23434300
C	2.24150400	3.00320200	1.17589300
H	2.38420900	3.65899700	2.01125800
H	1.13037300	1.76839800	-1.59194100

H	2.81506700	1.25722400	-1.47724300
N	-0.90395200	0.58862500	-0.32884300
C	-1.92413200	-0.18492100	0.04157700
N	-3.19778700	0.14590400	0.11887500
H	-3.85046200	-0.56649700	0.42847000
C	-3.68776100	1.48229000	-0.22044700
H	-3.46307400	1.71891800	-1.26420500
H	-3.23245600	2.23456700	0.42958200
H	-4.76756200	1.48568400	-0.07368700
O	-1.58775900	-1.48004700	0.38154100
H	0.35666700	-3.53899200	0.72944300
H	2.78647800	-2.99580900	0.23172900
H	3.40929700	-0.71677400	-0.53012300

1c + MeS-

C	-2.11630900	0.82309400	1.43031600
C	-2.37049100	-0.50417600	1.41540100
C	-1.41694900	-1.54810700	0.89305800
C	-0.13352800	-0.85680600	0.70896800
C	0.08743300	0.48189200	0.72359600
N	-0.87470900	1.39964500	1.12403100
C	-0.84975900	2.76492600	0.57782500
C	-1.26628200	2.82382800	-0.83111600
C	-1.61182900	2.81765500	-1.99148300
H	-1.92044100	2.81451500	-3.01713200
H	0.15994800	3.16692100	0.69149700
H	-1.51215000	3.37996500	1.19476400
N	1.40577600	0.77136700	0.39366200
C	1.93255400	-0.40568100	0.18503100
S	3.56431900	-0.84711400	-0.25598500
C	4.32669700	0.81197100	-0.35339000

H	3.85352300	1.40723400	-1.13717400
H	5.37608200	0.63840900	-0.60662200
H	4.26266600	1.32370200	0.60901500
O	1.07393200	-1.45675100	0.35306300
H	-1.35277200	-2.42371300	1.54824300
H	-3.34728000	-0.84497800	1.74324500
H	-2.86911800	1.53849700	1.74567800
S	-2.06422500	-2.33424500	-0.72892000
C	-2.04885800	-0.86147200	-1.81227600
H	-2.41167200	-1.17937000	-2.79452800
H	-1.03515400	-0.46345600	-1.91958000
H	-2.71012900	-0.08182100	-1.42284800

INT-B1

C	2.75673000	-1.41929900	0.33489000
C	2.23463500	-2.07970500	1.40865800
C	0.85346200	-1.92324400	1.76865500
C	0.11949900	-1.09305800	0.98213400
C	0.65966100	-0.40015600	-0.14970900
N	-0.21513600	0.36737300	-0.76066900
C	-1.43600700	0.22568800	-0.03668000
O	-1.19047500	-0.72665900	1.07741000
S	-2.06532500	1.81624800	0.68435800
C	-0.57385100	2.30825000	1.62157200
H	-0.80262900	3.27120000	2.08776900
H	0.28586400	2.42591800	0.95672700
H	-0.35000600	1.57692800	2.40292900
S	-2.85017600	-0.42520000	-1.04865600
C	-2.12454100	-2.03246700	-1.52712700
H	-1.99200500	-2.67581000	-0.65384200

H	-1.16856600	-1.89080900	-2.03857700
H	-2.83167300	-2.50198000	-2.21730100
N	1.97529200	-0.59280600	-0.46187700
C	2.58664900	0.18297300	-1.55647700
C	2.90558800	1.55463100	-1.14897600
C	3.15770300	2.68661000	-0.80427600
H	3.38189400	3.68934300	-0.50091500
H	1.89629300	0.18658500	-2.40466800
H	3.49051400	-0.34418500	-1.87278300
H	0.42958800	-2.43620800	2.62538400
H	2.88045400	-2.72583300	1.99230000
H	3.79382700	-1.50194600	0.03430300

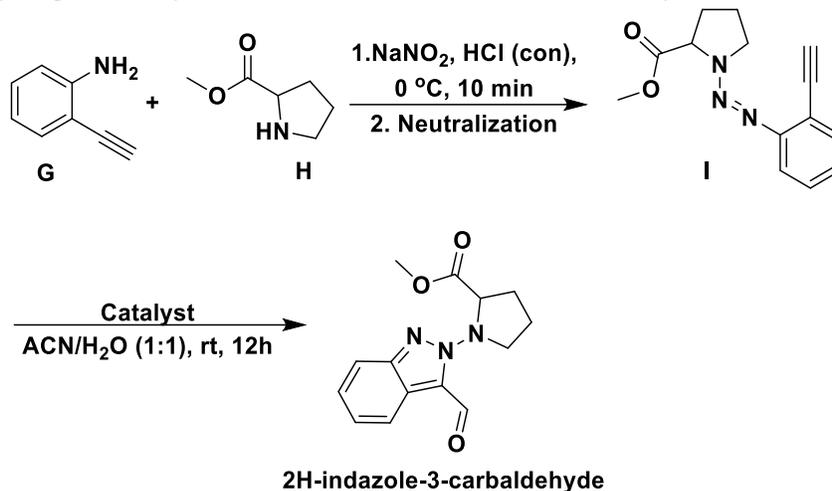
TS2-B

C	3.22998000	-1.04218700	0.43793600
C	2.74585900	-1.85221500	1.42320400
C	1.33165500	-2.00662100	1.62137300
C	0.52826500	-1.31337200	0.77287400
C	1.02937500	-0.46646700	-0.26773300
N	2.38144500	-0.35891300	-0.42391200
C	2.92294800	0.60026800	-1.40421800
C	2.87265000	1.97849400	-0.90722600
C	2.82048000	3.11238000	-0.48831900
H	2.77535700	4.11680600	-0.11832600
H	3.95258900	0.30803600	-1.62432900
H	2.34998800	0.50733800	-2.33115000
N	0.08582500	0.12855100	-0.96302500
C	-1.14631800	-0.29159300	-0.37470300
S	-2.16912700	-1.00048900	-1.53116400
C	-3.55708200	-1.58435600	-0.46048900

H	-4.17625600	-0.75771000	-0.10478000
H	-3.17577900	-2.16169900	0.38539700
H	-4.16534600	-2.23923300	-1.09218500
O	-0.83323900	-1.23044400	0.72695200
H	0.93538600	-2.63643600	2.41078400
H	3.44561500	-2.37820500	2.06264700
H	4.28742800	-0.89013800	0.26037800
C	-1.53726100	2.35999400	1.81127100
H	-1.35093200	1.59043500	2.56476500
H	-1.94644100	3.24934300	2.30096800
H	-0.60342800	2.62880300	1.30888600
S	-2.85440200	1.82253800	0.58654300

Supporting Information of Chapter Two:

Supplementary Figure 1: Synthesis of 2H-indazole-3-carbaldehyde and benzotriazole



Entry	Catalyst	Yield (%)
1	IPrAuCl , AgSbF_6 , DPSO	20 % (isolated yield two steps) ^a
2	CuCl , DPSO	No reaction ^b
3	CuCl	No reaction ^c 52 % (isolated yield two steps) ^d

Condition: a) 4 eq DPSO, 0.1 eq AgSbF_6 , 0.1 eq IPrAuCl b) 15 eq CuCl , 4 eq DPSO c) 15 eq CuCl d) $60\text{ }^\circ\text{C}$.

General procedure for the synthesis of compound I

2-ethynylaniline compound G (0.40 mmol, 1.0 equiv) was dissolved in 1 mL H_2O at room temperature, and then 1 mL of concentrated hydrochloric acid was added. The solution was stirred and cooled to $0\text{ }^\circ\text{C}$. After 15 min, a solution of NaNO_2 (0.48 mmol, 1.2 equiv, in 1 mL water) was added dropwise. The resulting solution of the diazonium salt was stirred for 10 min and then added to a solution of proline methyl ester H (1.2 mmol, 3 equiv) and K_2CO_3 (0.6 mmol, 1.5 equiv) in 1 mL of H_2O under $0\text{ }^\circ\text{C}$. The reaction mixture was warmed to room

temperature and stirred for 30 min. After completion, monitored by thin layer chromatography (TLC), the reaction mixture was extracted with ethyl acetate. The organic layer was dried over anhydrous Na₂SO₄, concentrated under reduced pressure to obtain the triazene intermediate I. The intermediate was directly used for next step without further purification.

Condition a:

Methyl ((2-ethynylphenyl)diazenyl)prolinate I (1eq, 1.2 mmol) was dissolved in 5 mL of ACN and 5 mL of H₂O. Diphenyl sulfoxide (4 eq, 4.8 mmol) was added followed by premixed IPrAuCl (10 mol%, 0.12 mol) and AgSbF₆ (10 mol%, 0.12 mmol). The reaction mixture was maintained room temperature for 12 h. The crude reaction mixture was purified by flash chromatography (hexane: ethyl acetate 10:1) to obtain **2H-indazole-3-carbaldehyde** as a white solid (57.1 mg, 20%).

Condition b:

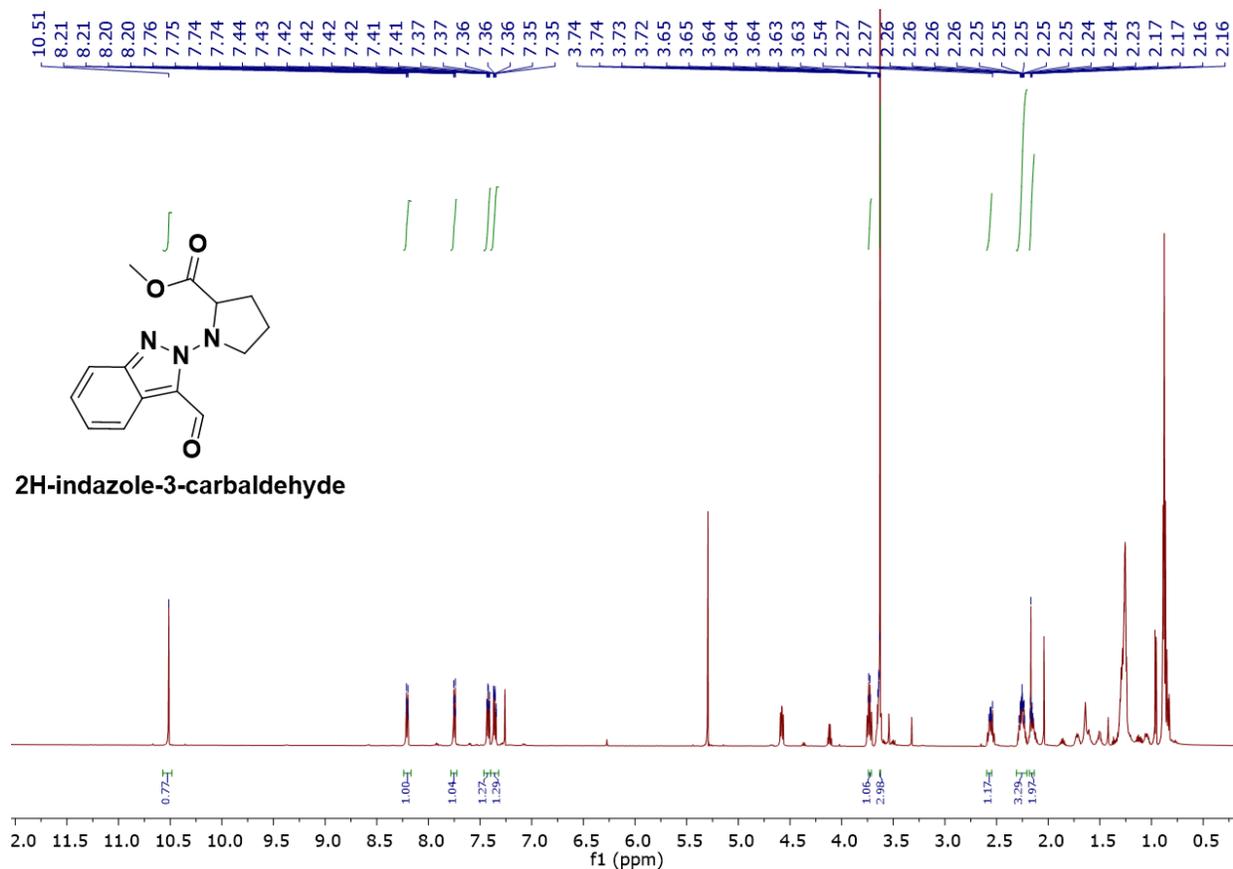
Methyl ((2-ethynylphenyl)diazenyl)prolinate I (1eq, 1.2 mmol) was dissolved in 5 mL of ACN and 5 mL of H₂O. Diphenyl sulfoxide (4 eq, 4.8 mmol) and CuCl (15 eq, 18 mmol) were added to the solution. The reaction mixture was stirred at room temperature for 12 h. The reaction was monitored by thin layer chromatography (TLC) to confirm no **2H-indazole-3-carbaldehyde** was forming.

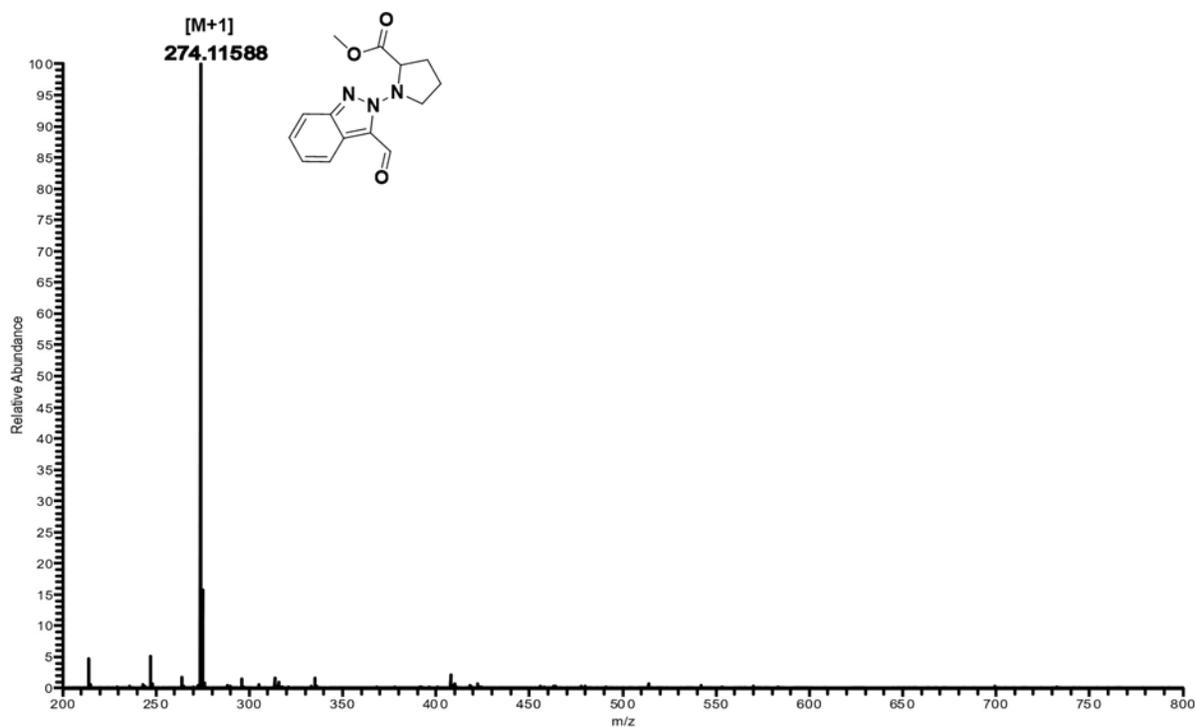
Condition c:

To a solution of methyl ((2-ethynylphenyl)diazenyl)prolinate I (1eq, 1.2 mmol) in 5 mL of ACN and 5 mL of H₂O, CuCl (15 eq, 18 mmol) was added to the mixture. The reaction mixture was stirred at room temperature for 12 h. The reaction was monitored by thin layer chromatography (TLC) to confirm no **2H-indazole-3-carbaldehyde** was forming.

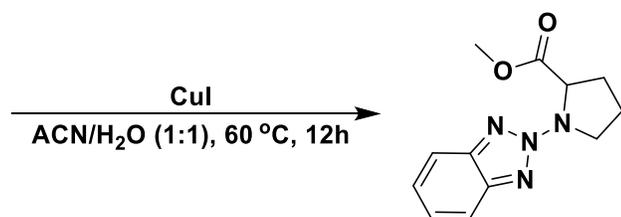
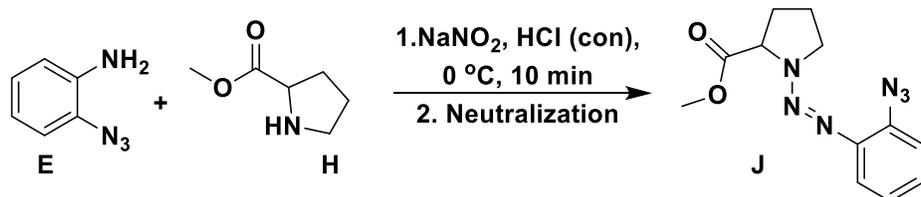
Condition d:

To a solution of methyl ((2-ethynylphenyl)diazenyl)prolinate I (1eq, 1.2 mmol) in 5 mL of ACN and 5 mL of H₂O, CuCl (15 eq, 18 mmol) was added to the mixture. The reaction mixture was allowed to stir at 60 °C for 12 h. The crude reaction mixture was purified by flash chromatography (hexane: ethyl acetate 10:1) to obtain **2H-indazole-3-carbaldehyde** as a white solid (148 mg, 52%). ¹H NMR (600 MHz, CDCl₃) δ 10.51 (s, 1H), 8.21 (dt, *J* = 8.4, 1.1 Hz, 1H), 7.75 (dt, *J* = 8.6, 1.0 Hz, 1H), 7.42 (ddd, *J* = 8.6, 6.7, 1.2 Hz, 1H), 7.36 (ddd, *J* = 8.3, 6.8, 1.0 Hz, 1H), 3.74 – 3.71 (m, 1H), 3.63 (s, 3H), 2.60 – 2.55 (m, 1H), 2.31 – 2.21 (m, 3H), 2.18 – 2.14 (m, 2H). ¹³C





General procedure for the synthesis of benzotriazole

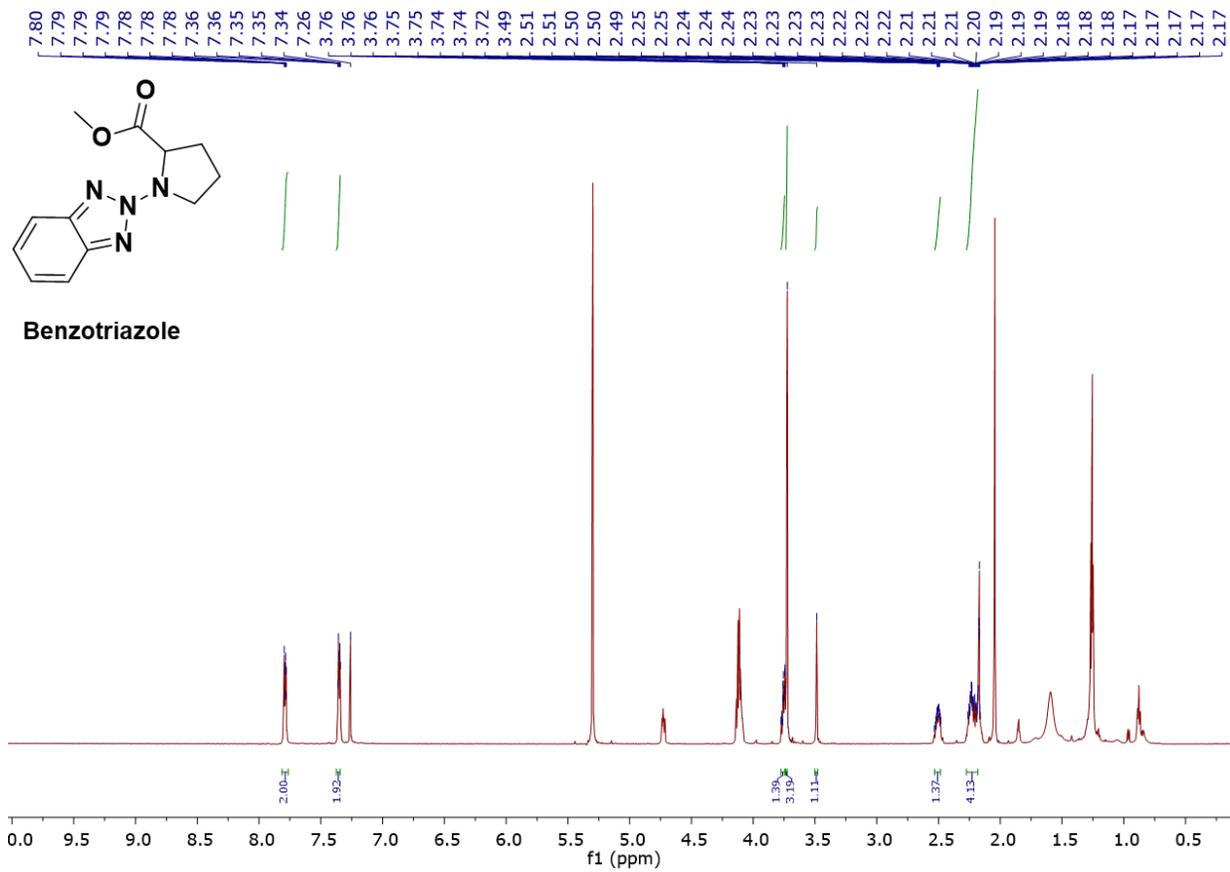


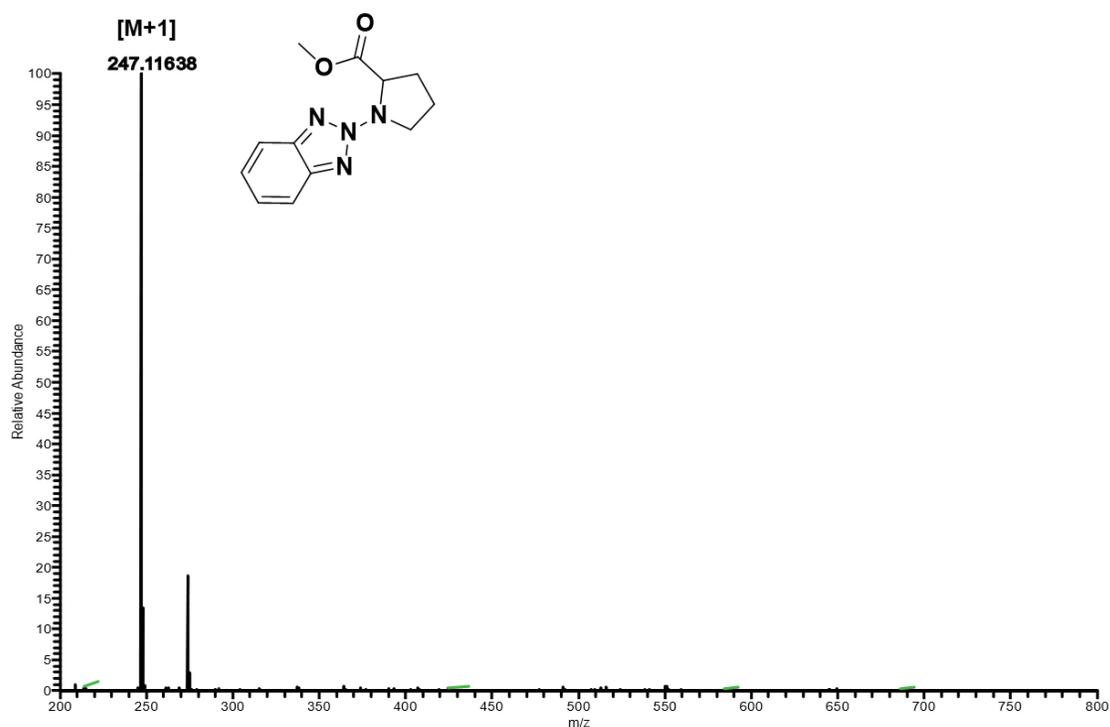
Benzotriazole, 43 % (two step isolated yield)

2-azidoaniline E (0.40 mmol, 1.0 equiv) was dissolved in 1 mL H_2O at room temperature, and then 1 mL of concentrated hydrochloric acid was added. The solution was stirred and cooled to $0\text{ }^\circ\text{C}$. After 15 min, a solution of NaNO_2 (0.48 mmol, 1.2 equiv, in 1 mL water) was added dropwise. The resulting solution of the diazonium salt was stirred for 10 min and then added to a solution of proline methyl ester H (1.2 mmol, 3 equiv) and K_2CO_3 (0.6 mmol, 1.5 equiv) in 1

mL of H₂O under 0 °C. The reaction mixture was warmed to room temperature and stirred for 30 min. After completion, monitored by thin layer chromatography (TLC), the reaction mixture was extracted with ethyl acetate. The organic layer was dried over anhydrous Na₂SO₄, concentrated under reduced pressure to obtain the triazene intermediate J. The intermediate J was directly used for next step without further purification. To a solution of methyl ((2-azidophenyl)diazenyl)prolinate compound J (1eq, 0.37 mmol) in 2 mL of ACN and 2 mL of H₂O, CuI (15 eq, 5.6 mmol) was added to the mixture. The reaction mixture was allowed to stir at 60 °C for 12 h. The crude reaction mixture was purified by flash chromatography (hexane: ethyl acetate 10:1) to obtain compound **benzotriazole**

as a white solid (39 mg, 43%). ¹H NMR (600 MHz, CDCl₃) δ 7.81 – 7.76 (m, 2H), 7.35 (dd, *J* = 6.2, 3.1 Hz, 2H), 3.77 – 3.75 (m, 1H), 3.73 – 3.72 (m, 3H), 3.49 (s, 1H), 2.51 (ddt, *J* = 11.8, 6.3, 2.8 Hz, 1H), 2.27 – 2.18 (m, 4H). ¹³C

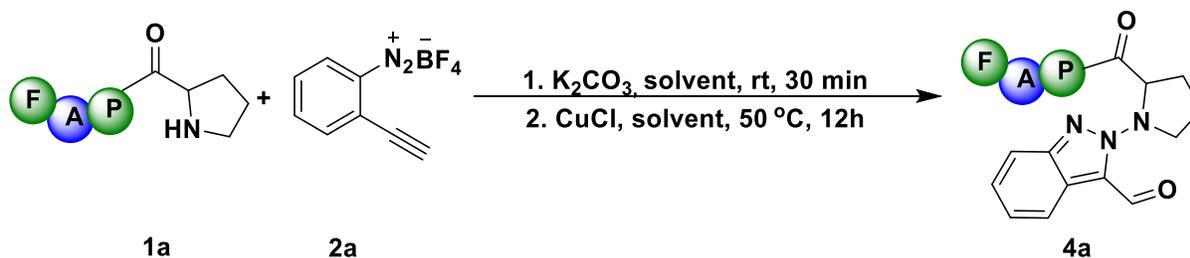




Supplementary Figure 2: General procedure for the modification of FAP with probes 2a and 3a.

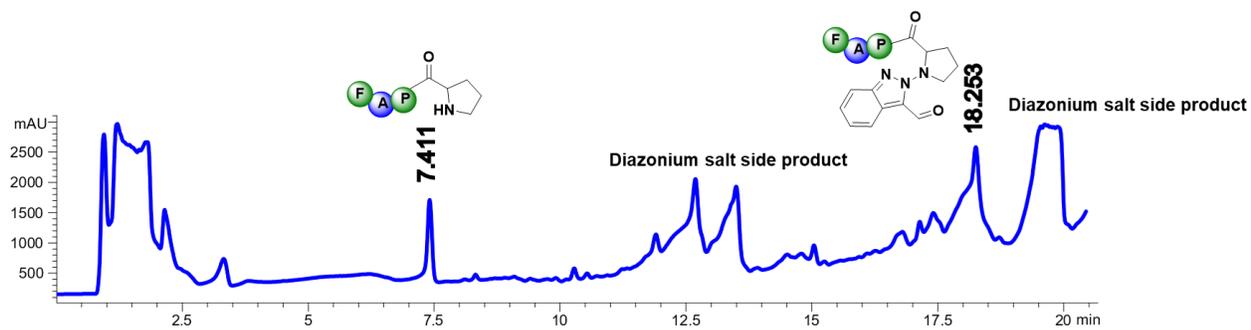
General procedure for the synthesis of compound 4a

To a 1 mg of PAF **1a** (7.5 mM) in 0.4 mL of solvent was added K_2CO_3 (3 eq or 10 eq) and diazonium salt **2a** (3 eq or 10 eq). The reaction was incubated at room temperature after 30 mins $CuCl$ or CuI (15 eq) in 100 μ l ACN was added. The reaction mixture was allowed to incubate at 50 °C for 12 h. The reaction was analyzed by HPLC and MS/MS. HPLC was carried out with 1% formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

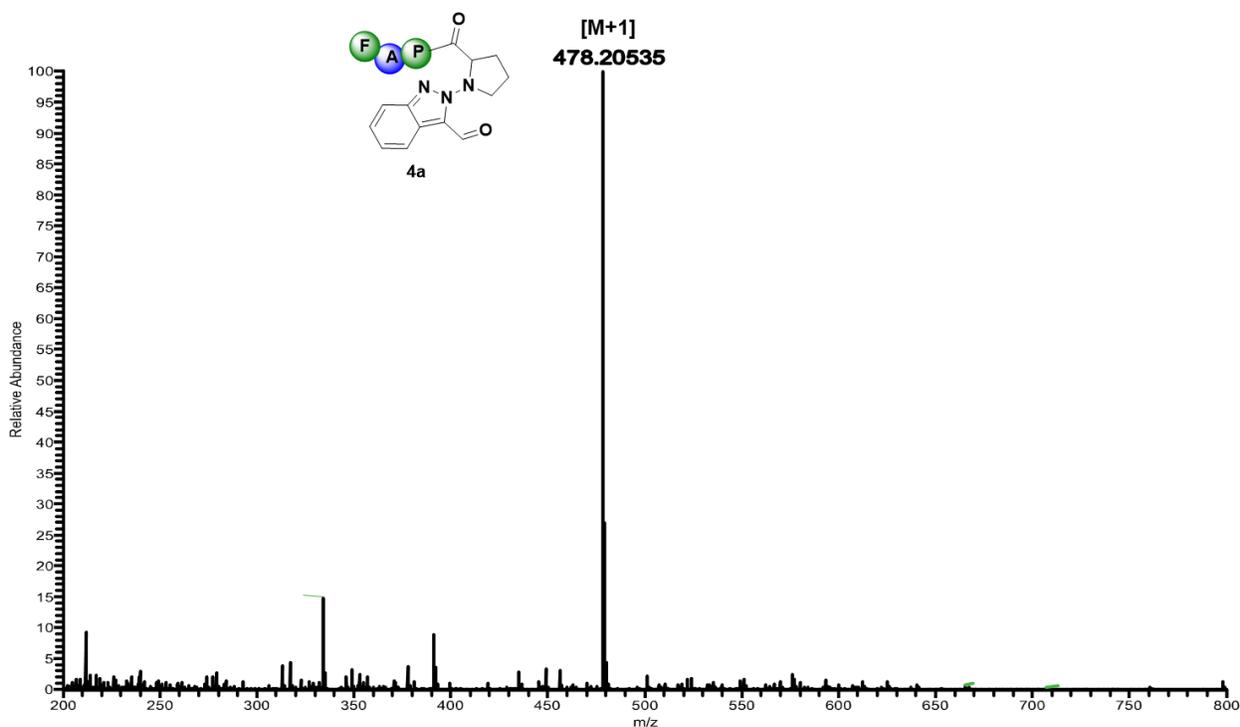


Entry	Catalyst	Eq of 2a	Temperature °C	Solvent	Conversion (%)
1	CuCl	3	50 °C	ACN/Nap (1:1)	25 %
2	CuCl	3	50 °C	ACN/Nap (9:1)	46 %
3	CuCl	10	50 °C	ACN/Nap (9:1)	76 %
4	CuI	3	50 °C	ACN/Nap (9:1)	0 %
5	CuCl	3	50 °C	ACN/Nap (1:9)	3 %
6	CuCl	3	50 °C	DMF/Nap (1:9)	14.2 %
7	CuCl	3	rt	ACN/Nap (1:1)	0 %

HPLC trace of 4a



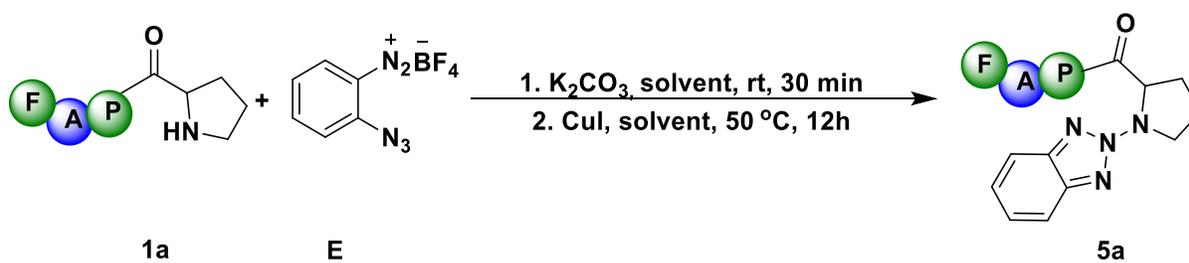
HR-MS of 4a



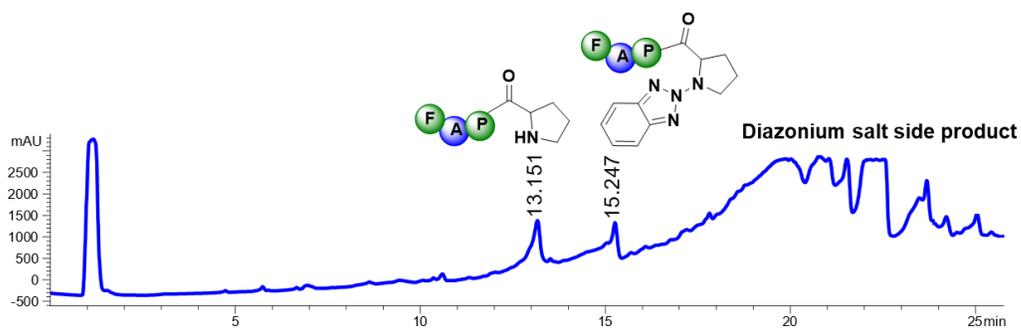
General procedure for the synthesis of compound 5a

To a 1 mg of PAF **1a** (7.5 mM) in 0.4 mL of solvent was added K_2CO_3 (3 eq or 10 eq) and diazonium salt E (3 eq or 10 eq). The reaction was incubated at room temperature after 30 mins CuI or $CuCl$ (15 eq) in 100 μ L ACN was added. The reaction mixture was allowed to incubate at 50 °C for 12 h. The reaction was analyzed by HPLC and MS/MS. HPLC was carried out with 1% formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

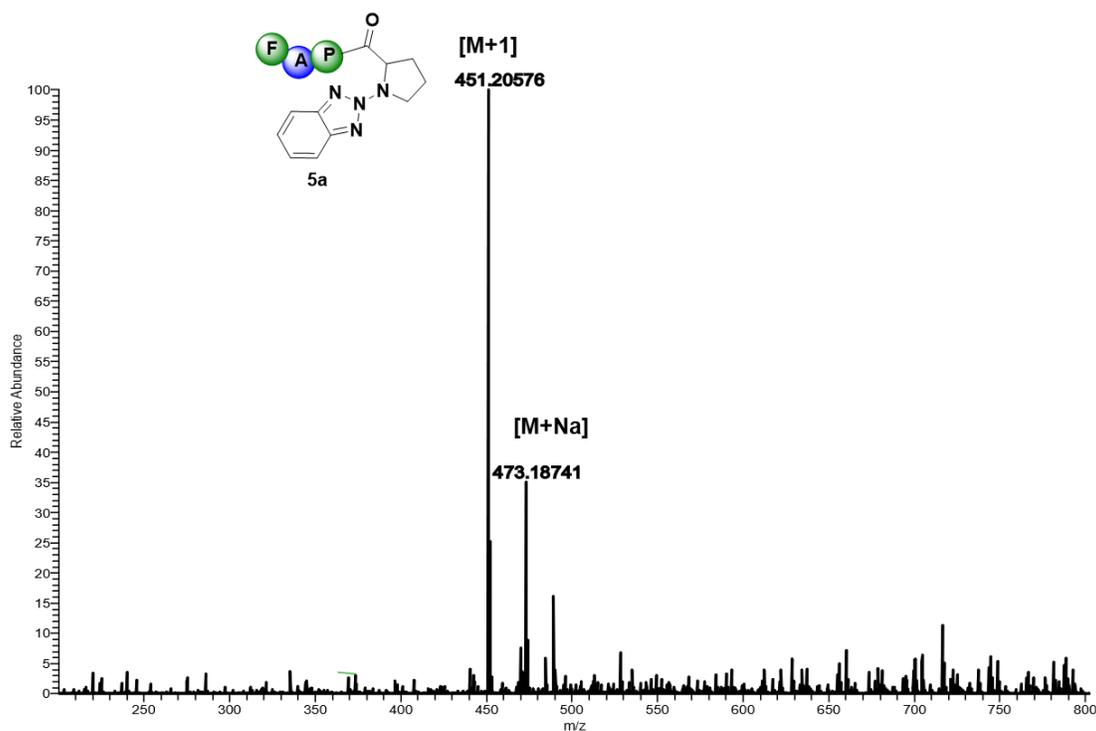
Entry	Catalyst	Eq of E	Temperature °C	Solvent	Conversion (%)
1	CuI	3	50 °C	ACN/Nap (1:1)	35 %
2	CuCl	3	50 °C	ACN/Nap (1:1)	15 %
3	CuI	3	50 °C	ACN/Nap (9:1)	17 %
4	CuI	10	50 °C	ACN/Nap (9:1)	57 %
5	CuI	3	50 °C	ACN/Nap (1:9)	30 %
6	CuI	3	50 °C	DMF/Nap (1:9)	12.5 %
7	CuI	3	rt	ACN/Nap (1:1)	10 %



HPLC trace of 5a



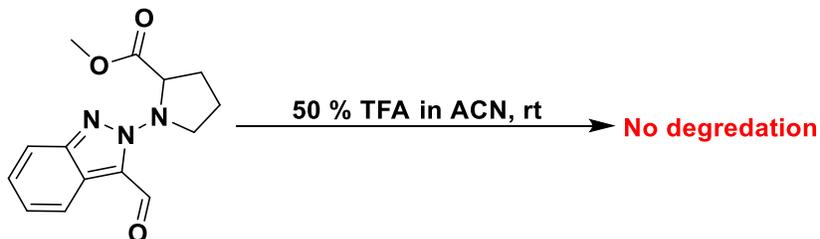
HR-MS of 5a

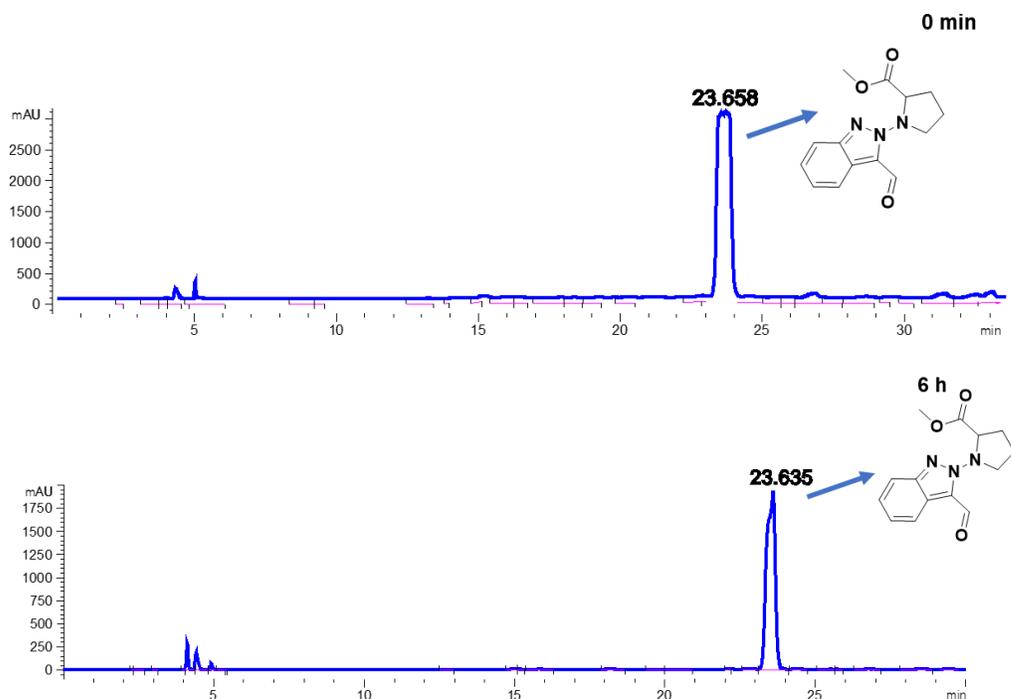


Supplementary Figure 3: Stability study of 2H-indazole-3-carbaldehyde group in TFA and pyridine.

General procedure for stability study in TFA

Methyl (3-formyl-2H-indazol-2-yl) proline (4.5 mM) was incubated in 400 μ L (TFA : ACN 1:1) at room temperature for 6 h. The samples (50 μ L) were taken from the reaction mixtures and directly injected into the HPLC. The reactions were monitored by injecting samples in HPLC after regular intervals of time 0 h, and 6 h. The bioconjugate product showed high stability in TFA and pyridine for 6 h.

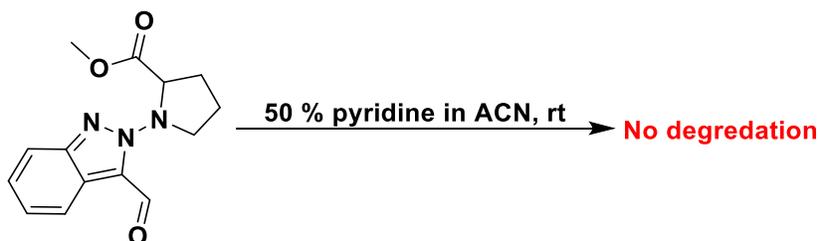


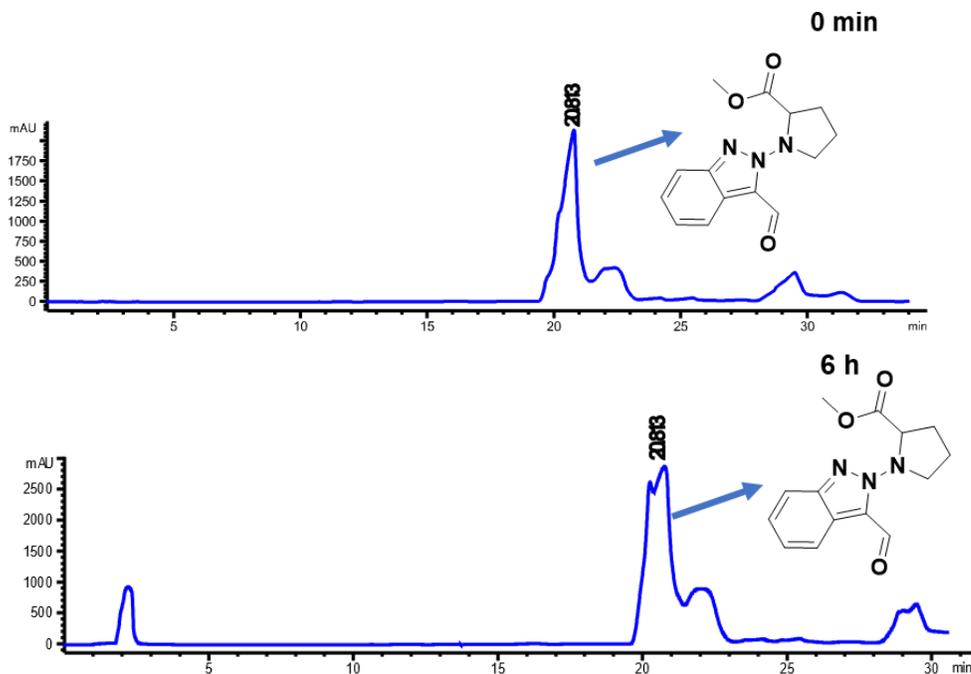


Stability study of 2H-indazole-3-carbaldehyde group in pyridine.

General procedure for stability study in pyridine

Methyl (3-formyl-2H-indazol-2-yl) proline (4.5 mM) was incubated in 400 μ L (pyridine ACN 1:1) at room temperature for 6 h. The samples (50 μ L) were taken from the reaction mixtures and directly injected into the HPLC. The reactions were monitored by injecting samples in HPLC after regular intervals of time 0 h, and 6 h. The bioconjugate product showed high stability in and pyridine for 6 h.





Supplementary Figure 4: Chemoselectivity studies for the TCC reaction

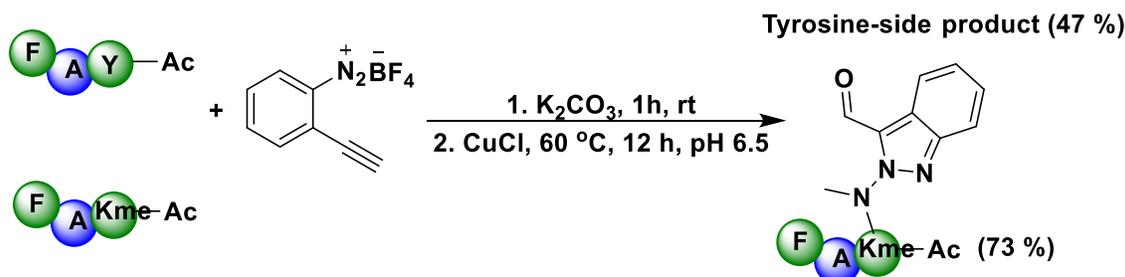
Procedure for chemo-selectivity studies for the TCC reaction

To peptides Ac-XAF (0.0027 mmol) with 10 eq of K_2CO_3 in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of 2a in 40 μ L ACN was added, the reaction mixture was incubated at 25 $^\circ$ C in the incubator for 1 h. The reaction mixture was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

Supplementary Figure 5: Restriction of TCC side reaction with tyrosine

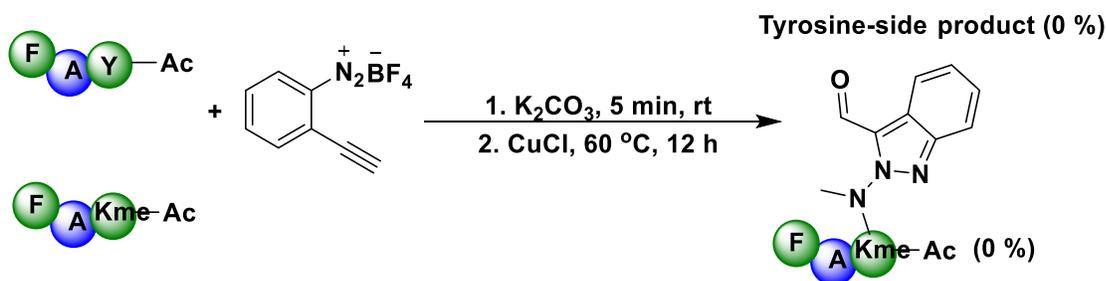
General procedure for optimization of triazene cyclization:

pH adjustment:



To peptides Ac-KmeAF or Ac-YAF (0.0027 mmol) with 10 eq of K_2CO_3 in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of 2a in 40 μ L ACN was added, the pH of the reaction mixture was adjusted to 6.5 by 1.0 M HCl, then the reaction mixture was incubated at 25 $^{\circ}$ C in the incubator for 1 h. Next, 15 eq of CuCl (0.0405 mmol) in 100 μ L of ACN was added into this mixture, then the mixture was allowed to stir at 40 $^{\circ}$ C for 12 h. The reaction mixture was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

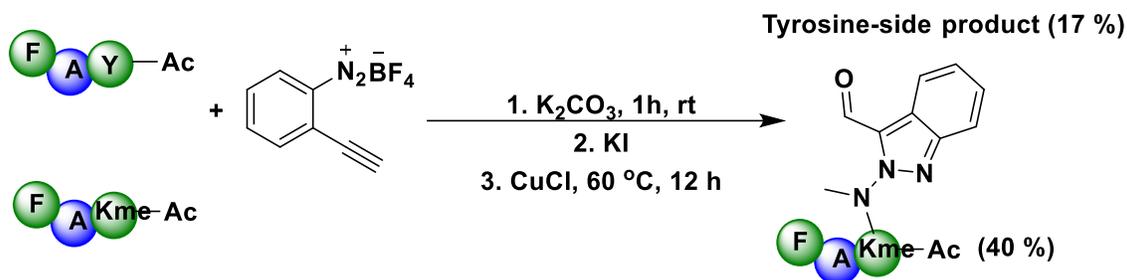
5 min:



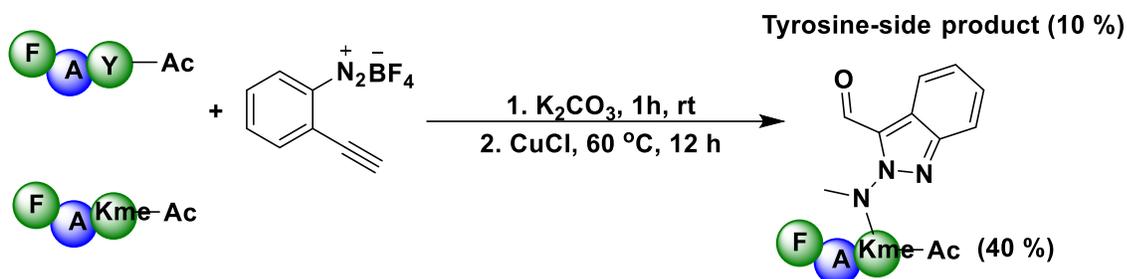
To peptides Ac-KmeAF or Ac-YAF (0.0027 mmol) with 10 eq of K_2CO_3 in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of 2a in 40 μ L ACN was added, the reaction mixture was incubated at 25 $^{\circ}$ C in the incubator for 5 min. Next, 15 eq of CuCl (0.0405 mmol) in

100 μL of ACN was added into this mixture, then the mixture was allowed to stir at 40 $^{\circ}\text{C}$ for 12 h. The reaction mixture was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

KI quench:

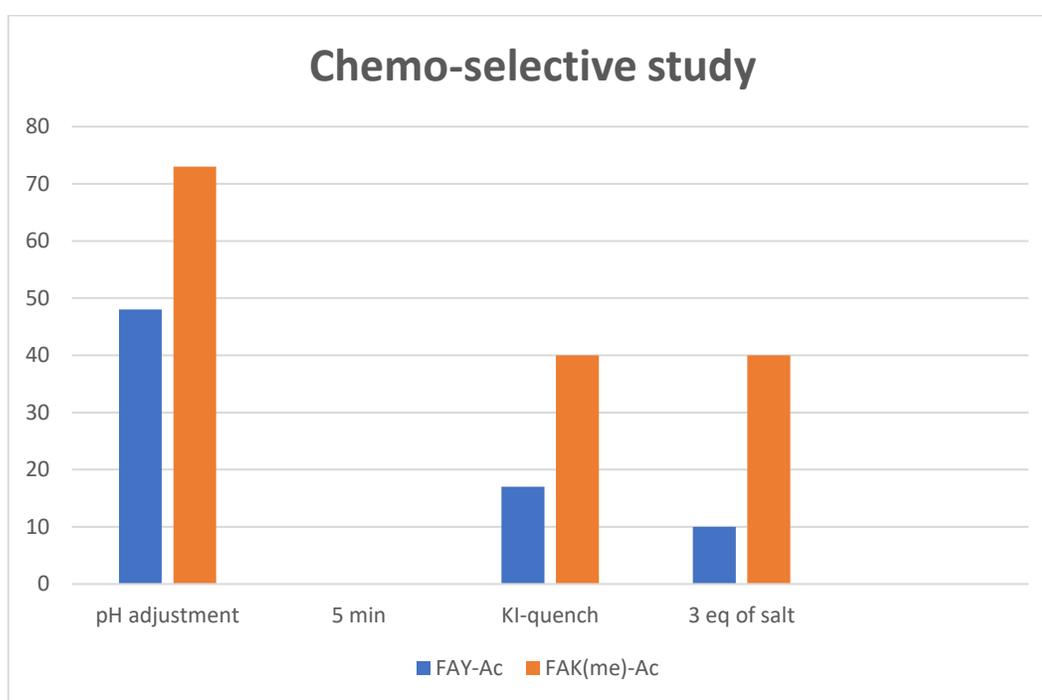


To peptides Ac-KmeAF or Ac-YAF (0.0027 mmol) with 10 eq of K_2CO_3 in 360 μL of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of 2a in 40 μL ACN was added, the reaction mixture was incubated at 25 $^{\circ}\text{C}$ in the incubator for 1 h, then the KI (0.027 mmol) was added to quench the excess diazonium salt. Next, 15 eq of CuCl (0.0405 mmol) in 100 μL of ACN was added into this mixture, then the mixture was allowed to stir at 40 $^{\circ}\text{C}$ for 12 h. The reaction mixture was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



3 eq of diazonium salt.

To peptides Ac-KmeAF or Ac-YAF (0.0027 mmol) with 10 eq of K_2CO_3 in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of 2a in 40 μ L ACN was added, then the reaction mixture was incubated at 25 $^{\circ}C$ in the incubator for 1 h. Next, 15 eq of CuCl (0.0405 mmol) in 100 μ L of ACN was added into this mixture, then the mixture was allowed to stir at 40 $^{\circ}C$ for 12 h. The reaction mixture was analyzed by LC/MS. LC: water (solvent A): acetonitrile (solvent B); gradient 0-80 %, acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



Supplementary Figure 6: Selective blocking of tyrosine by 1,3 diphenyl propynone

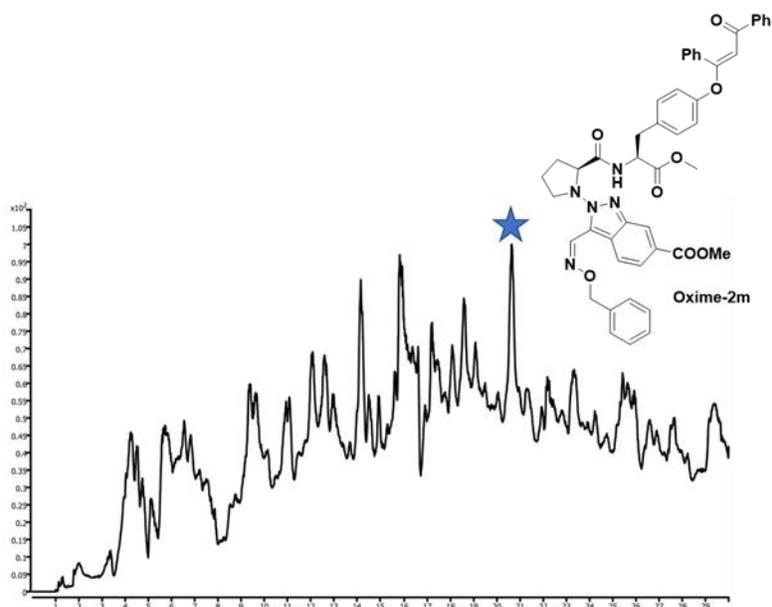
General procedure for tyrosine labeling by 1,3 diphenyl propynone

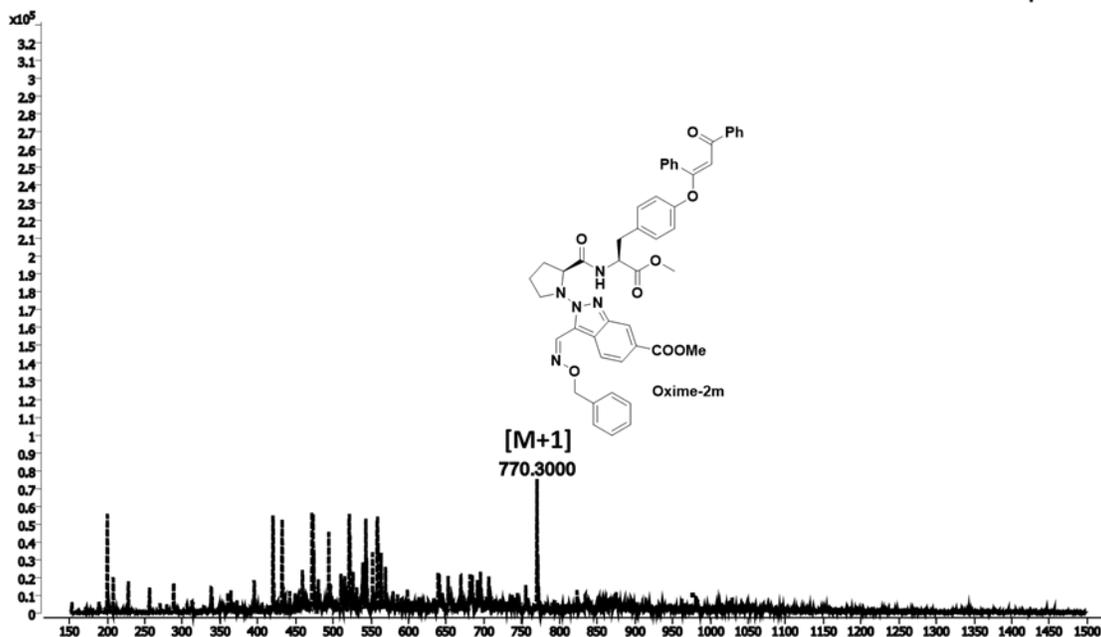
To the peptide 1m (1 eq., 0.0034 mmol) with K_2CO_3 (2 eq., 0.0068 mmol) in 200 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), 3.5 mg of 1,3 diphenyl propynone (5 eq, 0.017 mmol) in 200 μ L of ACN was added, the reaction was allowed to stir at 40 $^{\circ}C$ for 12 h. After 12 h, the

reaction mixture was incubated with 50 μL of 1 M HCl at 25 $^{\circ}\text{C}$ for 4h to decouple the proline modified product, the mixture was dried by lyophilization for next step.

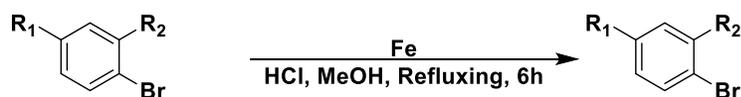
Procedure for oxime enrichment:

10 eq of K_2CO_3 (0.034 mmol) in 100 μL of 10 mM sodium phosphate buffer (Nap, pH 7.5) and 10 eq of 2d (0.034 mmol) in 100 μL ACN were mixed with 1m' (0.0034 mmol), the reaction mixture was incubated at 25 $^{\circ}\text{C}$ for 1 h. The mixture was directly used for next step without further purification. 15 eq of CuCl (0.051 mmol) in 50 μL ACN was mixed with triazene intermediate (0.0034 mmol) in 100 μL of 10 mM sodium phosphate buffer (Nap, pH 7.5) and 100 μL ACN, the reaction mixture was allowed to stir in incubator at 60 $^{\circ}\text{C}$ for 12 h. The reaction mixture was directly used for next step without further purification. O-benzylhydroxylamine (5 eq, 0.017 mmol) was mixed with 2m, the mixture was incubated at 25 $^{\circ}\text{C}$ in the incubator for 5 h. The reaction mixture was filtered and purified by HPLC to obtain the oxime-2m. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm



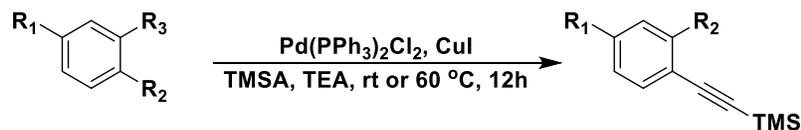


Supplementary Figure 7: Optimization of triazination coarctate cyclization with different diazonium salts analogs.



1. R₁ = COOMe, R₂ = NO₂

4. R₁ = COOMe, R₂ = NH₂ (55%)



2. R₁ = F, R₂ = I, R₃ = NH₂

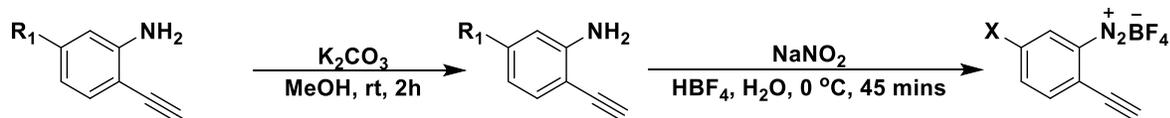
3. R₁ = CF₃, R₂ = Br, R₃ = NH₂

4. R₁ = COOMe, R₂ = Br, R₃ = NH₂

5. R₁ = F, R₃ = NH₂ (81%)

6. R₁ = CF₃, R₃ = NH₂ (82%)

7. R₁ = COOMe, R₃ = NH₂ (84%)



5. R₁ = F

6. R₁ = CF₃

7. R₁ = COOMe

8. R₁ = F (85%)

9. R₁ = CF₃ (65%)

10. R₁ = COOMe (90%)

2b. R₁ = F (90%)

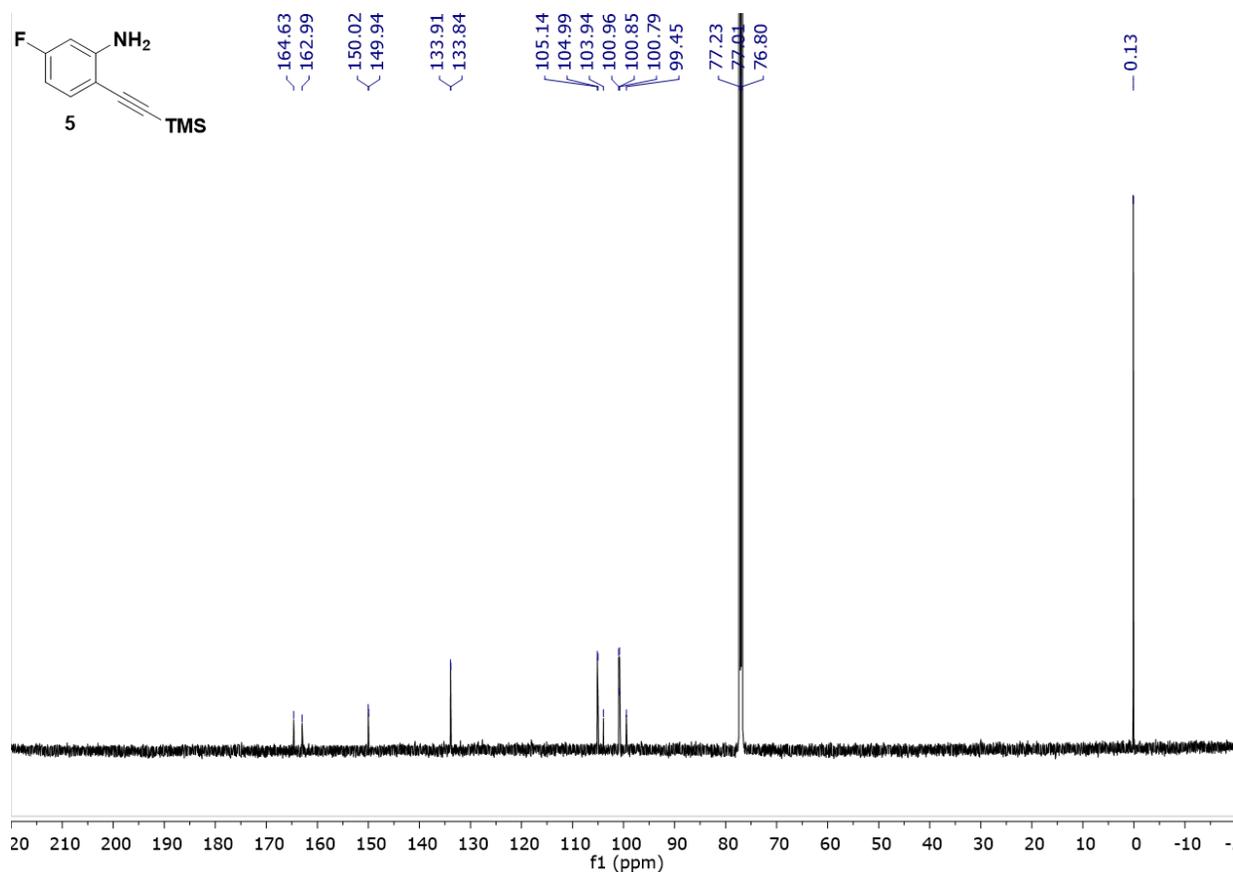
2c. R₁ = CF₃ (92%)

2d. R₁ = COOMe (88%)

Compound 5.

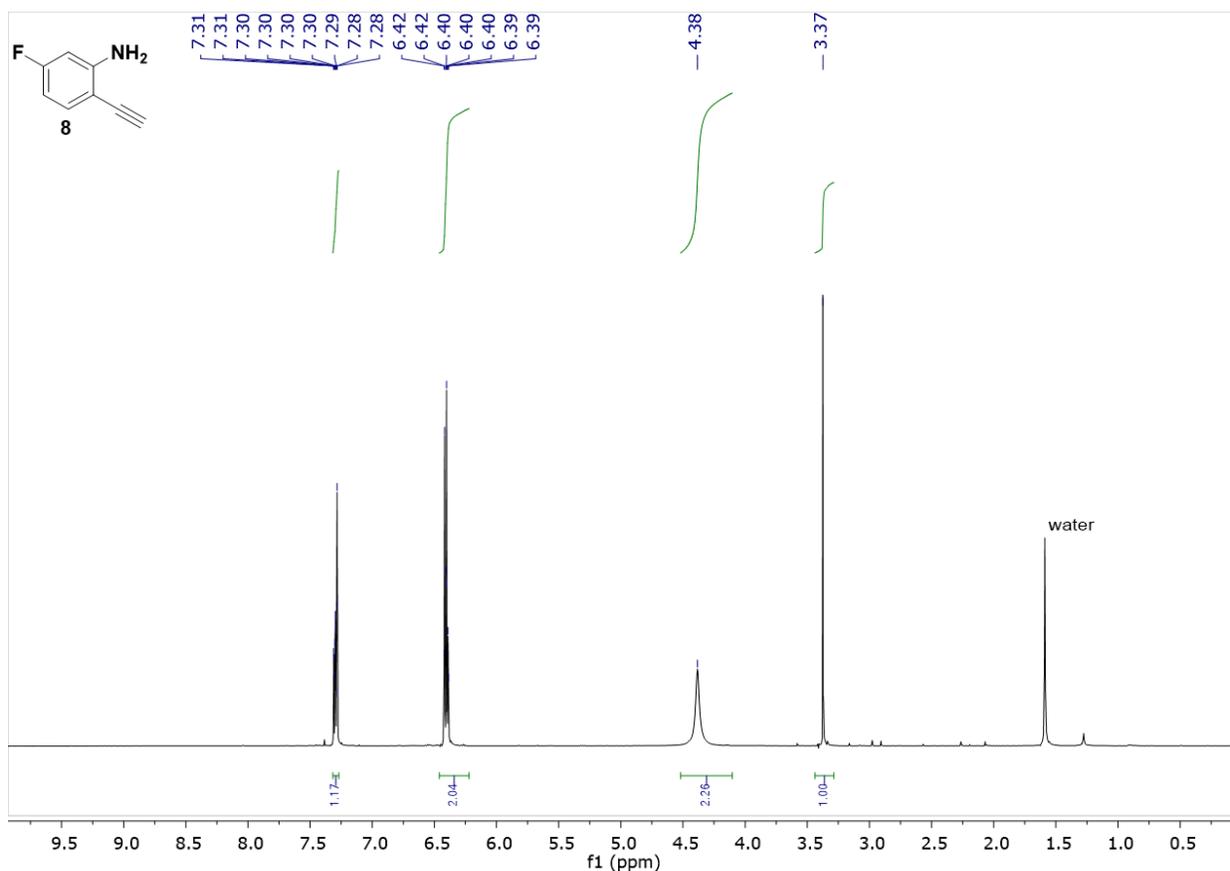
A mixture of compound 3 (948 mg, 4 mmol), Pd(PPh₃)₂Cl₂ (56.2 mg, 0.08 mmol), CuI (30.4mg, 0.08 mmol), and TMS-acetylene (620 ul, 4.8 mmol), in TEA (16 mL) was allowed to stirred at rt 12 h under nitrogen, when the reaction was completed, TEA was removed under pressure, and the residues were purified by column chromatography (hexane: ethyl acetate 10:1) to obtain compound 6 as brown oil (786 mg, 81 %). ¹H NMR (600 MHz, CDCl₃)

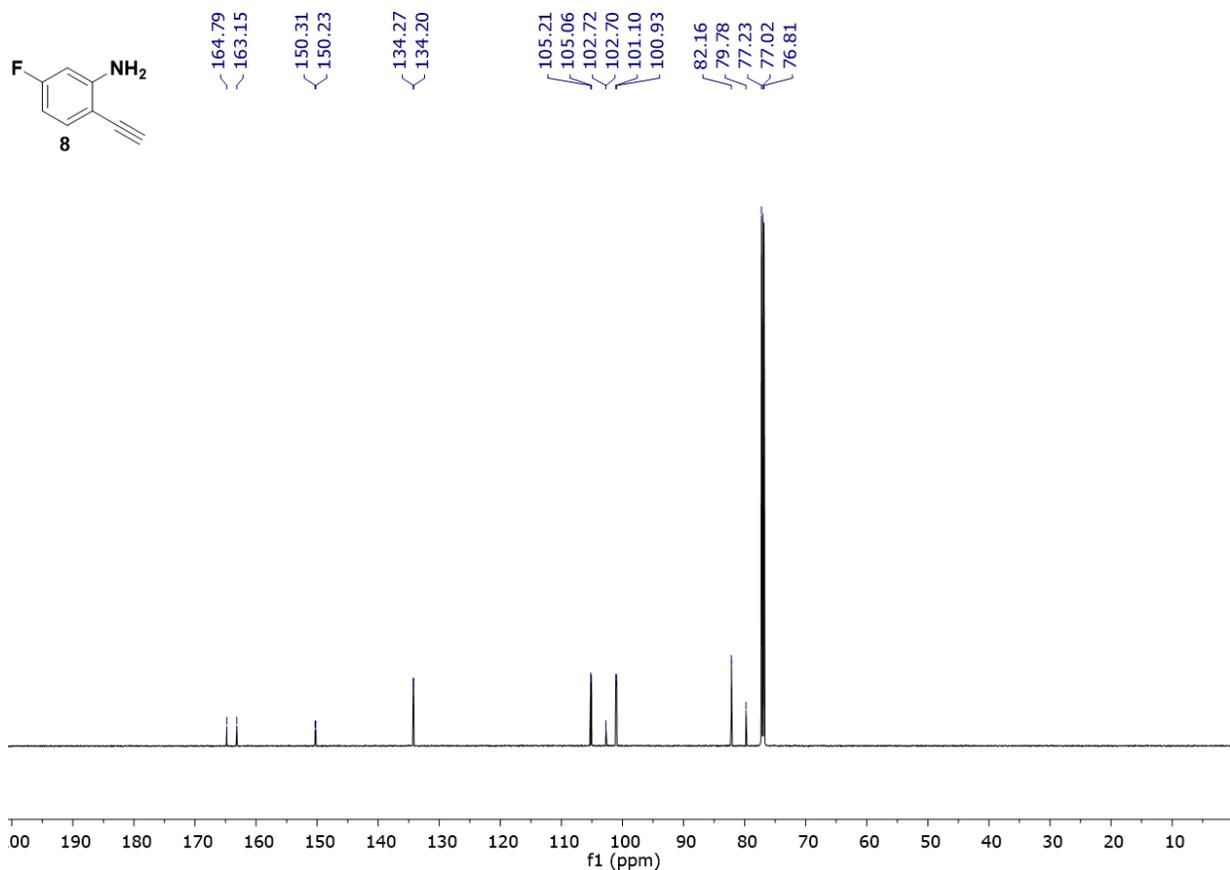
δ 7.25 (ddd, *J* = 9.2, 6.1, 3.3 Hz, 1H), 6.36 (ddt, *J* = 10.7, 8.3, 5.7 Hz, 2H), 4.34 (s, 2H), 0.25 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 163.81 (d, JC-F = 246.0 Hz), 149.98 (d, JC-F = 12.0 Hz), 133.87 (d, JC-F = 10.5 Hz), 105.06 (d, JC-F = 22.5 Hz), 103.94, 100.87 (d, JC-F = 25.5 Hz), 100.85, 99.45, 0.13.



Compound 8.

To a solution of 6 (825 mg, 4 mmol), MeOH (20 mL), K₂CO₃ (1.1 g, 8 mmol) was added. The resulting mixture was allowed to stir at room temperature for 2 h then concentrated in vacuo. The residue was extracted with ethyl acetate and the combined organic layers were washed with brine, dried over Na₂SO₄, and concentrated under reduced pressure. The crude material was purified by column chromatography (hexane: ethyl acetate 10:1) to obtain the pure compound 8 as brown oil (459 mg, 85 %). ¹H NMR (600 MHz, CDCl₃) δ 7.32 – 7.27 (m, 1H), 6.46 – 6.22 (m, 2H), 4.38 (s, 2H), 3.37 (s, 1H). ¹³C NMR (151 MHz, CDCl₃) δ 163.97 (d, JC-F = 246.0 Hz), 150.27 (d, JC-F = 12.0 Hz), 134.23 (d, JC-F = 10.5 Hz), 105.13 (d, JC-F = 22.5 Hz), 102.71 (d, JC-F = 3.0 Hz), 100.01 (d, JC-F = 25.5 Hz), 82.16, 79.78.

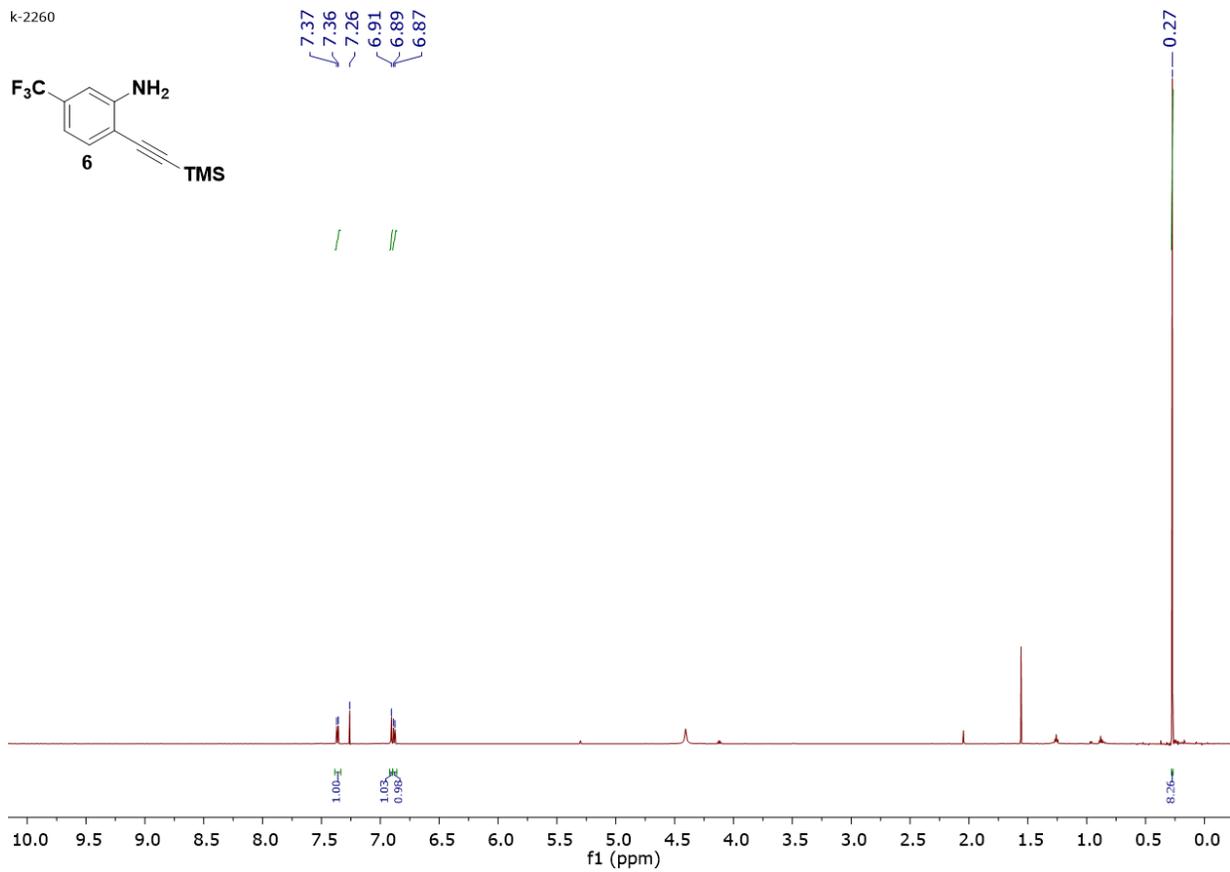
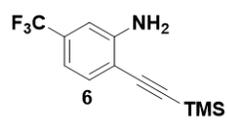


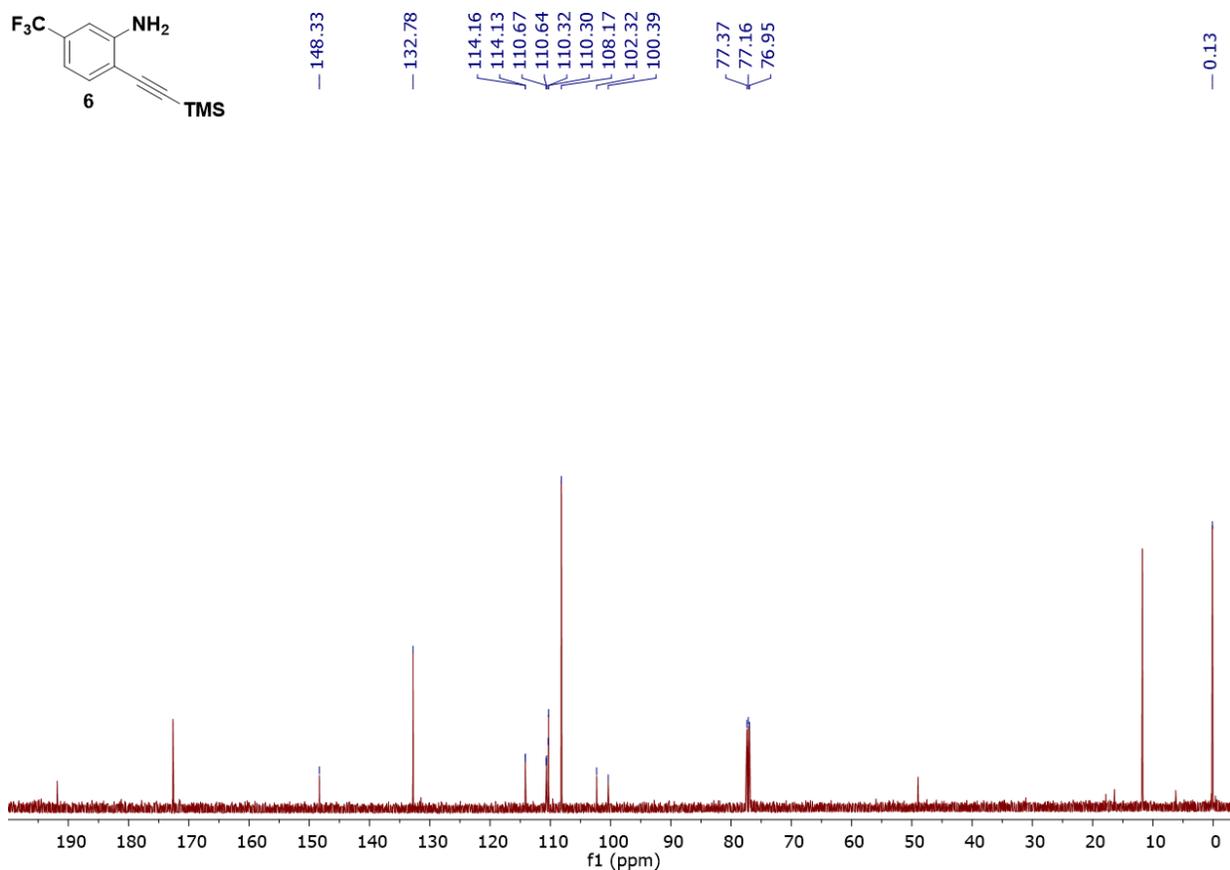


Compound 6.

A mixture of compound 4 (1.25 g, 5.23 mmol), Pd(PPh₃)₂Cl₂ (91.2 mg, 0.13 mmol), CuI (25 mg, 0.13 mmol), and TMS-acetylene (900 ul, 6.3 mmol), in TEA (20 mL) and DMF (1 mL) was allowed to stirred at 60 °C 12 h under nitrogen, when the reaction was completed, TEA was removed under pressure, and the residues were purified by column chromatography (hexane: ethyl acetate 10:1) to obtain compound 7 as brown oil (1.1 g, 82 %). ¹H NMR (600 MHz, CDCl₃) δ 7.36 (d, *J* = 8.1 Hz, 1H), 6.91 (s, 1H), 6.88 (d, *J* = 8.0 Hz, 1H), 0.27 (s, 8H). ¹³C NMR (151 MHz, CDCl₃) δ 148.33, 132.78, 114.14 (d, *J*C-F = 4.5 Hz), 110.65 (d, *J*C-F = 4.5 Hz), 110.31 (d, *J*C-F = 3.0 Hz), 108.17, 102.32, 100.39, 0.13.

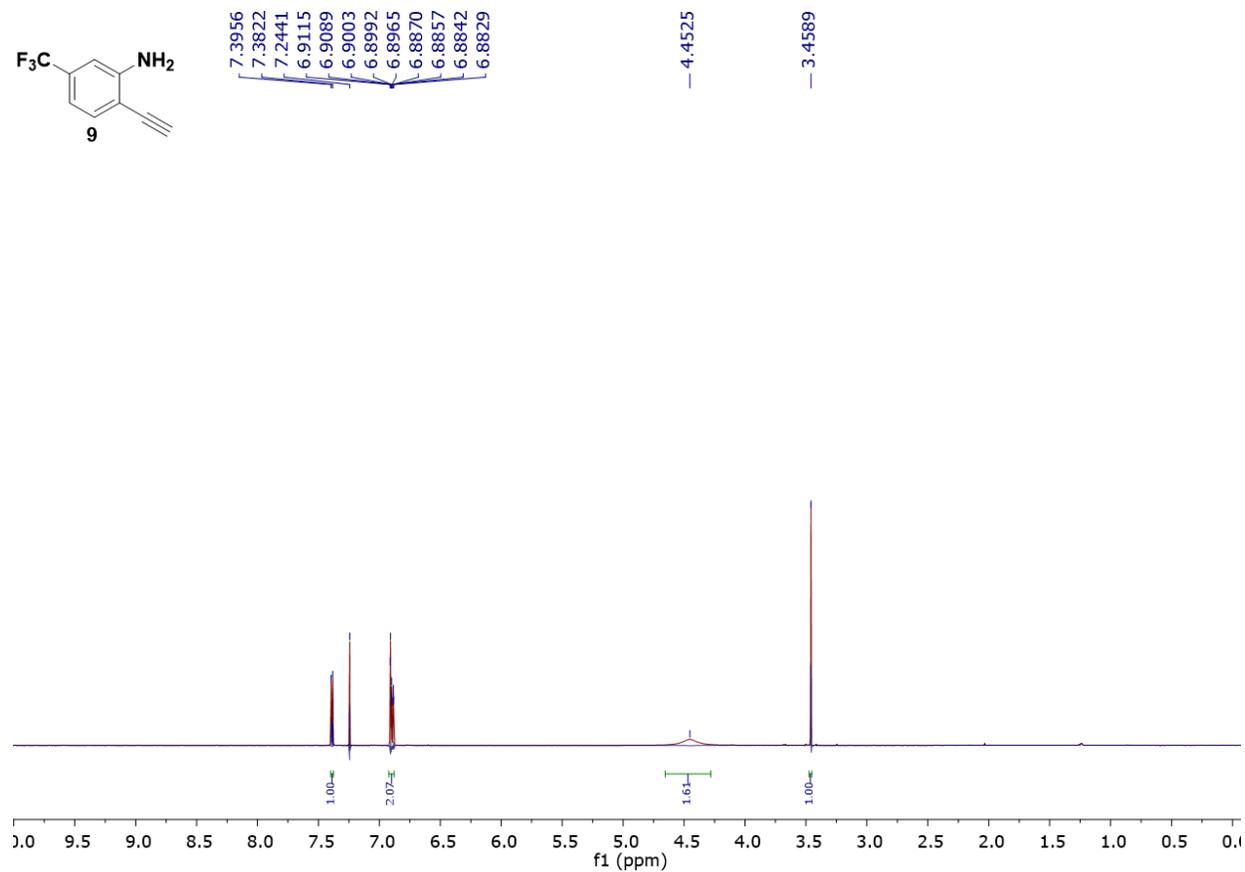
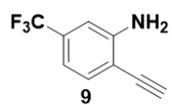
K-2260

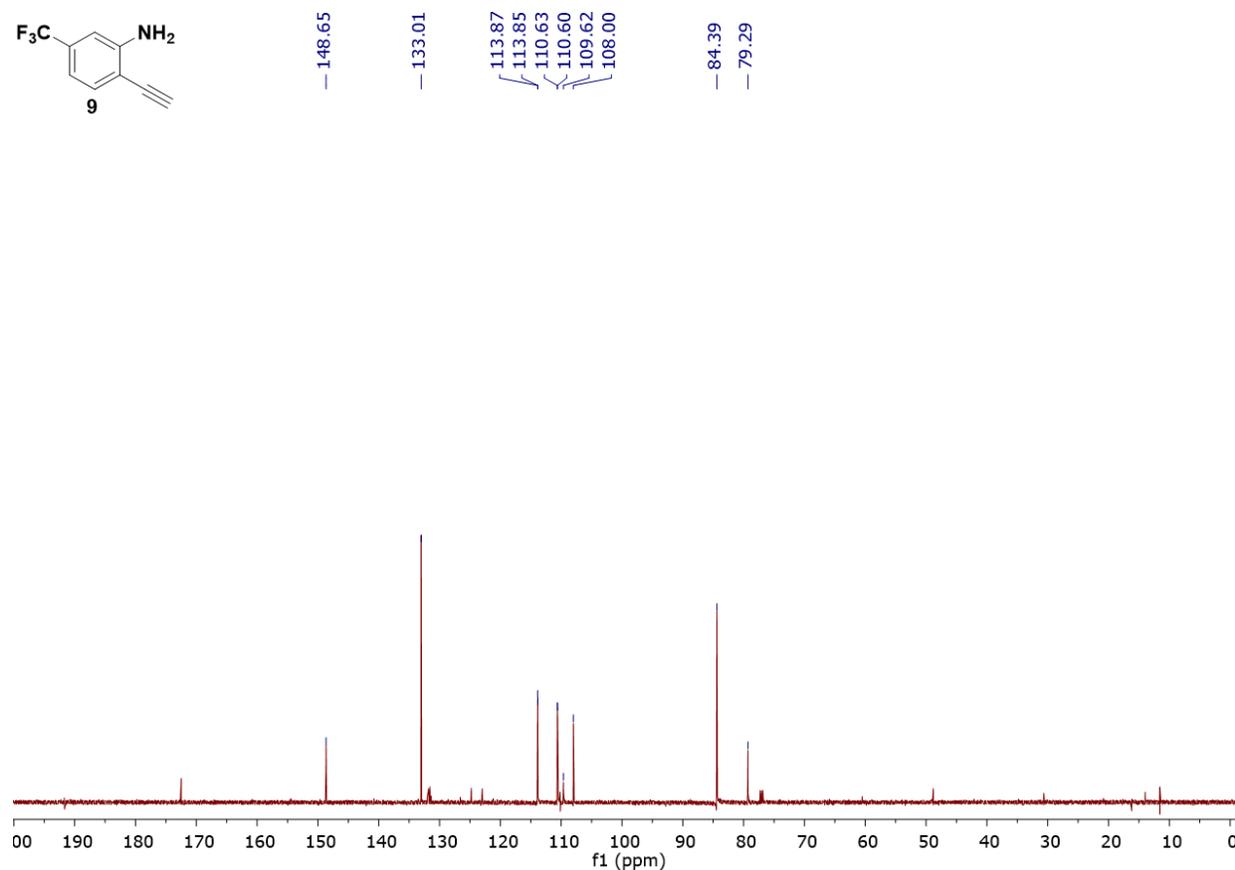
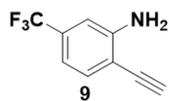




Compound 9.

To a solution of compound 7 (624.8 mg, 2.43 mmol) in CH₃OH (12 ml). Potassium carbonate (671 mg, 4.86 mmol) was added, and the resulting mixture was allowed to stir at room temperature for 2 h then concentrated in vacuo. The residue was extracted with ethyl acetate and the combined organic layers were washed with brine, dried over Na₂SO₄, and concentrated under reduced pressure. The crude material was purified by column chromatography (hexane: ethyl acetate 30:1) to obtain the pure compound 9 as brown oil (292 mg, 65%). ¹H NMR (600 MHz, CDCl₃) δ 7.39 (d, *J* = 8.0 Hz, 1H), 6.92 – 6.88 (m, 2H), 4.45 (s, 2H), 3.46 (s, 1H). ¹³C NMR (151 MHz, CDCl₃) δ 148.65, 133.01, 113.86 (d, *J*C-F = 4.5 Hz), 110.62 (d, *J*C-F = 4.5 Hz), 109.62, 108.00, 84.39, 79.29.





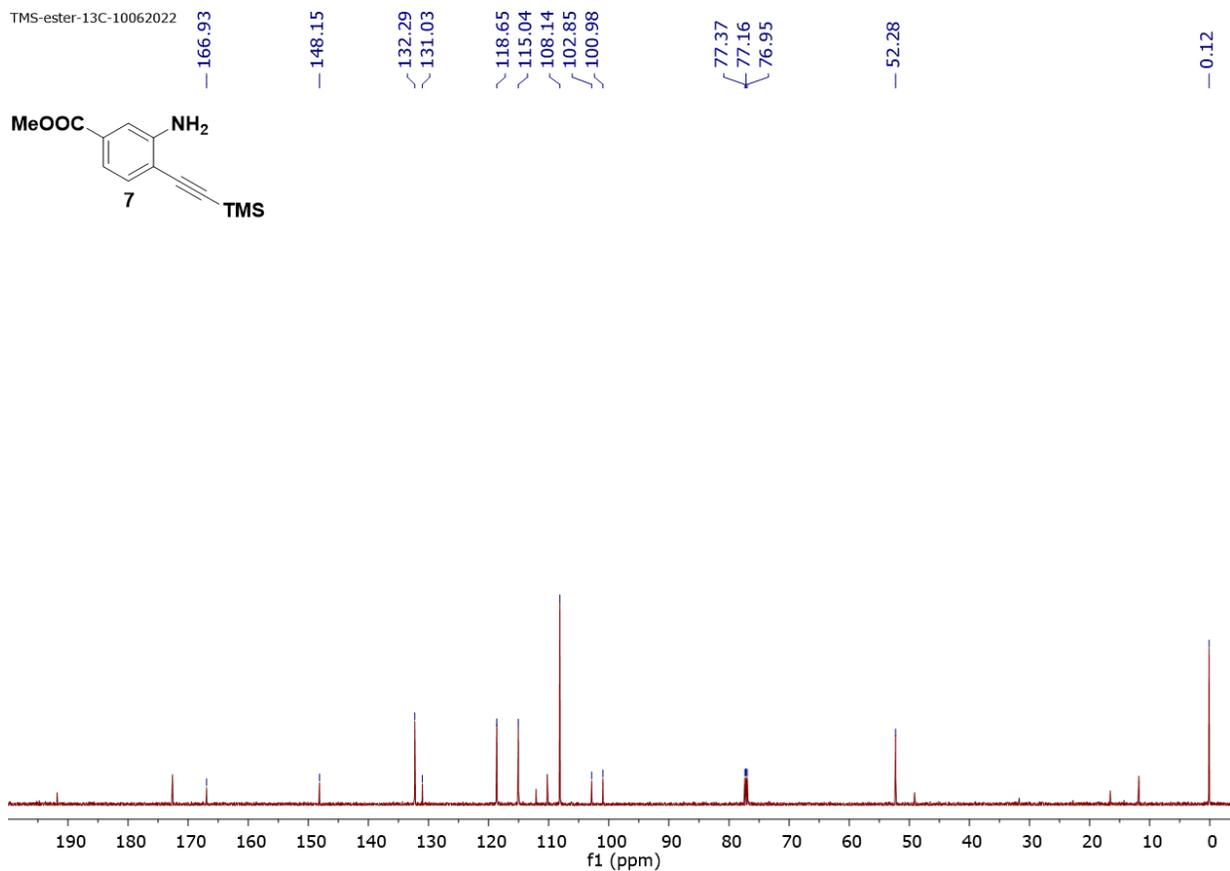
Compound 2.

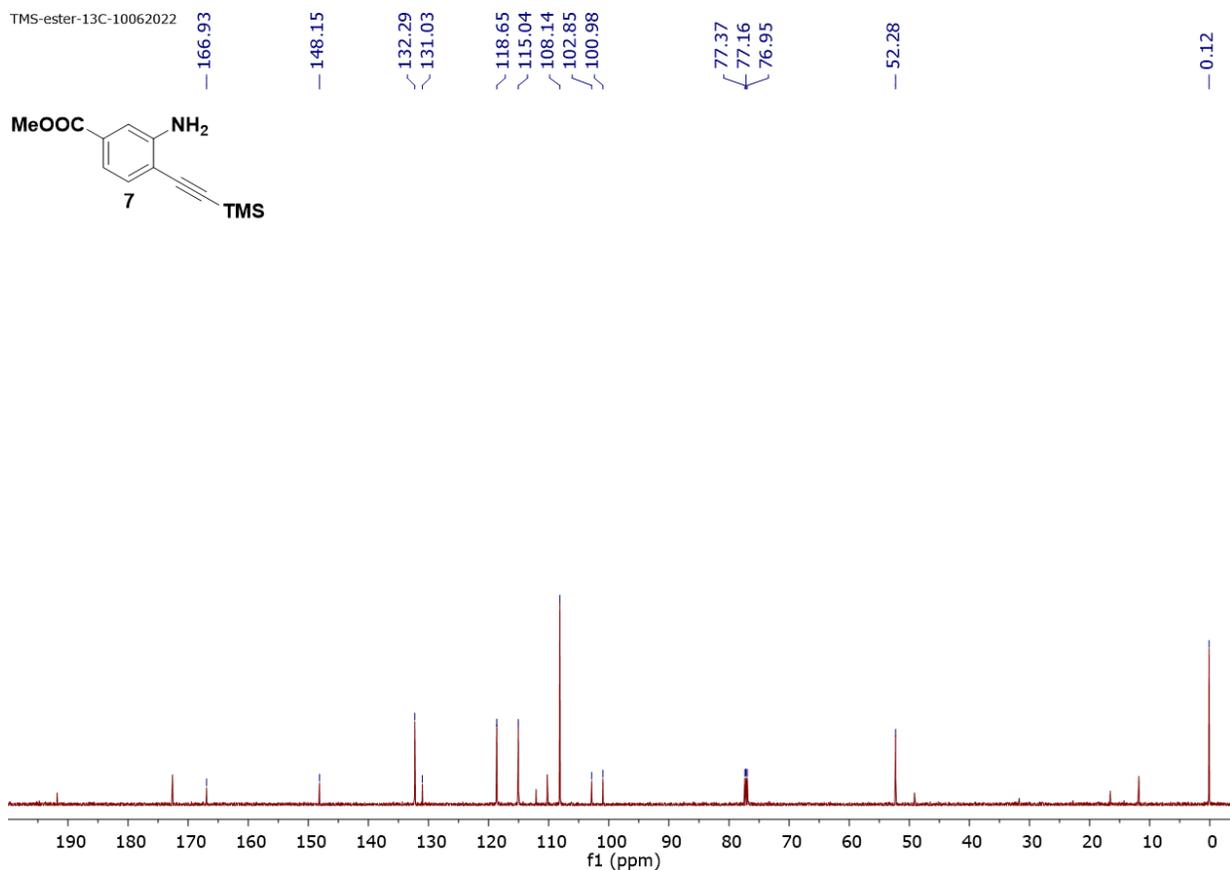
To a solution of compound 1 (1 g, 3.8 mmol) in MeOH (10 mL), iron (646 mg, 11.5 mmol) and HCl (3.8 mL) were added sequentially. The reaction mixture was stirred at 80 °C for a hours. The reaction solution was filtered through Celite and washed with ethyl acetate and sodium bicarbonate for 3 times. The organic layer was dried over MgSO₄, and the solvent was removed under reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 9:1) to obtain compound 2 as white solid (483 mg, 55 %). **NMR**

Compound 7.

A mixture of compound 2 (485 mg, 2.1 mmol), Pd(PPh₃)₂Cl₂ (73.7 mg, 0.1 mmol), CuI (20 mg, 0.1 mmol), and TMS-acetylene (450 ul, 3.1 mmol), in TEA (3.2 mL) was allowed to stir at room

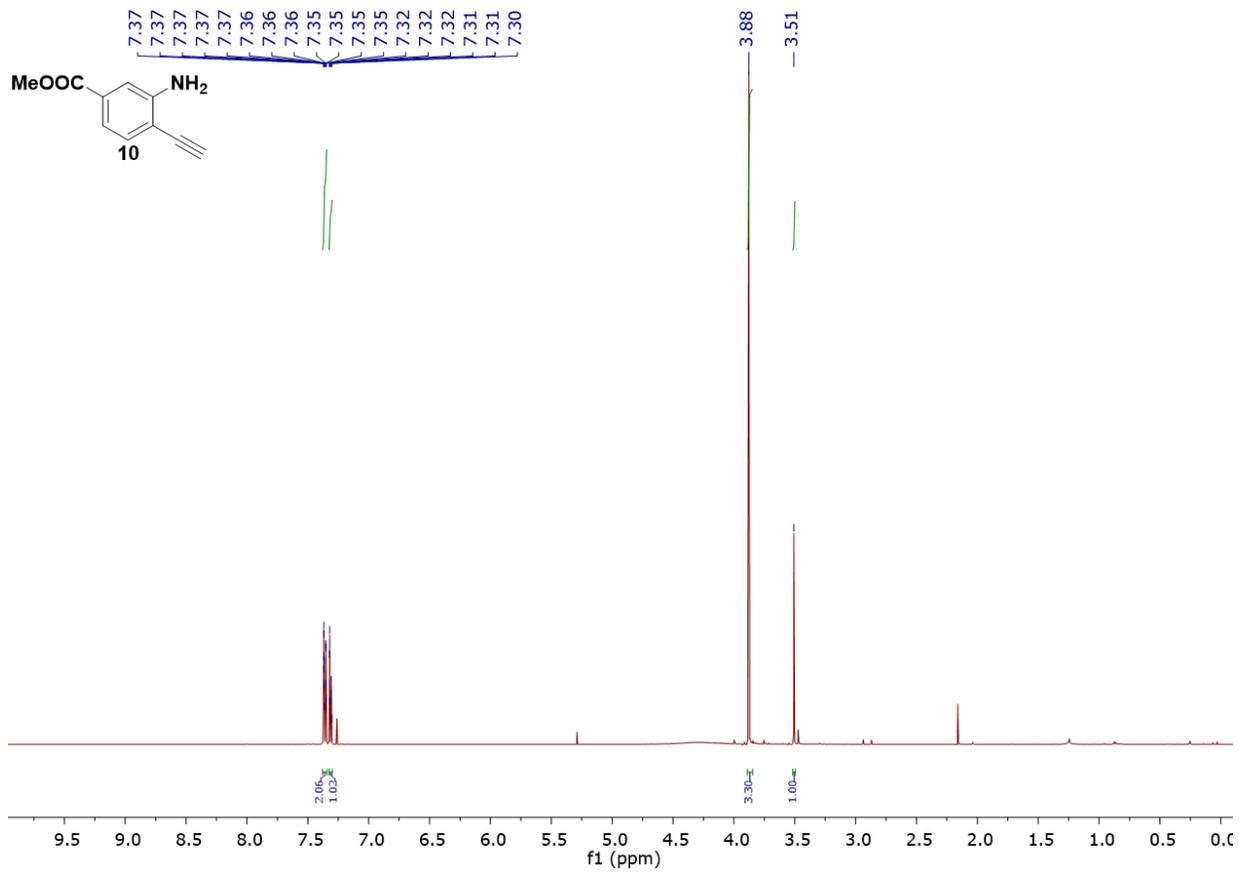
temperature 12 h under nitrogen. when the reaction was completed, triethylamine was removed under pressure, and the residues were purified by column chromatography (hexane: ethyl acetate 9:1) to obtain compound 5 as brown oil (435 mg, 84 %). ^1H NMR (600 MHz, CDCl_3) δ 7.37 – 7.36 (m, 1H), 7.34 – 7.29 (m, 2H), 4.37 (s, 2H), 3.87 (s, 3H), 0.26 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 166.93, 148.15, 132.29, 131.03, 118.65, 115.04, 108.14, 102.85, 100.98, 52.28, 0.12.

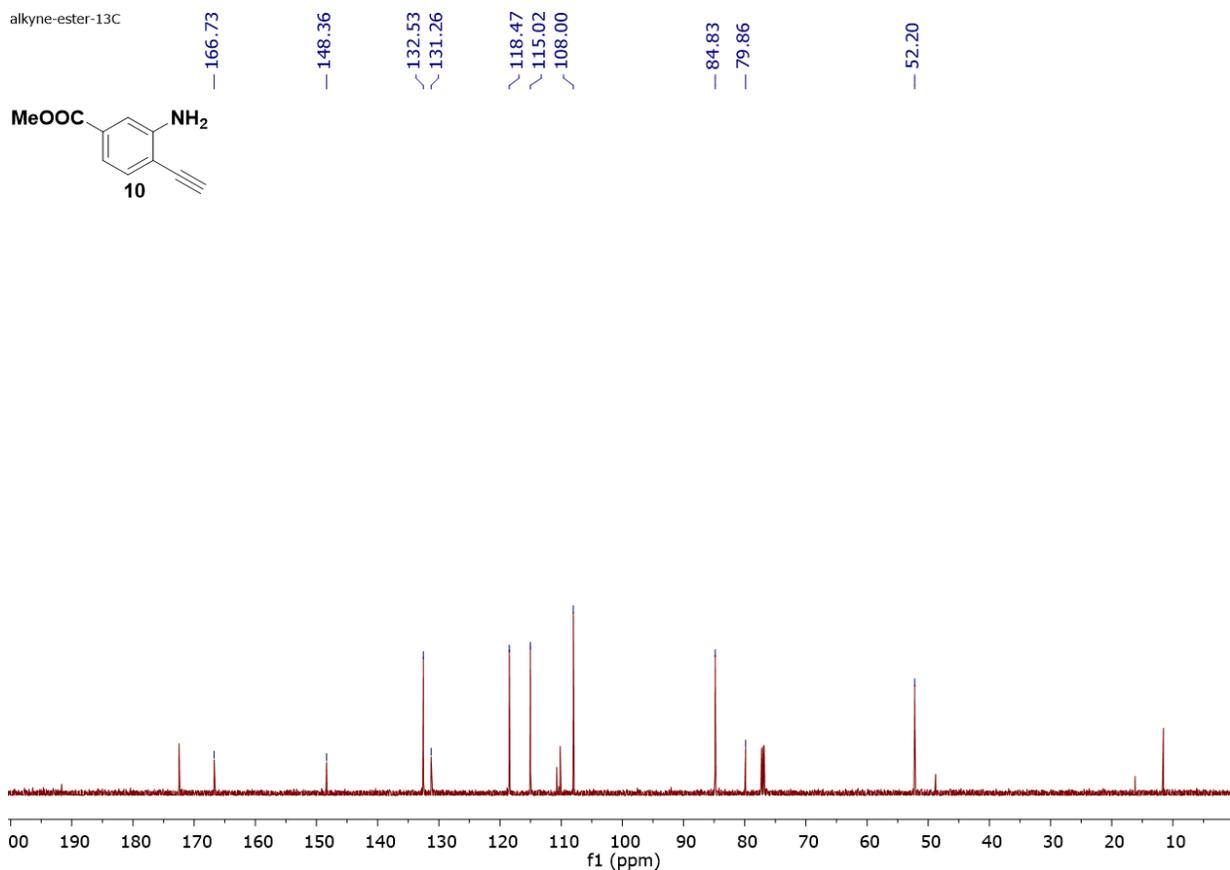




Compound 11.

To a solution of compound 5 (100 mg, 0.404 mmol) in CH_3OH (2 mL). Potassium carbonate (111.5 mg, 0.808 mmol) was added, and the resulting mixture was allowed to stir at room temperature for 2 h then concentrated in vacuo. The residue was extracted with ethyl acetate and the combined organic layers were washed with brine, dried over Na_2SO_4 , and concentrated under reduced pressure. The crude material was purified by column chromatography (hexane: ethyl acetate 50:1) to obtain the pure compound 11 as yellow solid (63.7 mg, 90 %). ^1H NMR (600 MHz, CDCl_3) δ 7.38 – 7.34 (m, 2H), 7.31 (dt, $J = 8.0, 1.5$ Hz, 1H), 3.88 (s, 3H), 3.51 (s, 1H). ^{13}C NMR (151 MHz, CDCl_3) δ 166.73, 148.36, 132.53, 131.26, 118.47, 115.02, 108.00, 84.83, 79.86, 52.20.



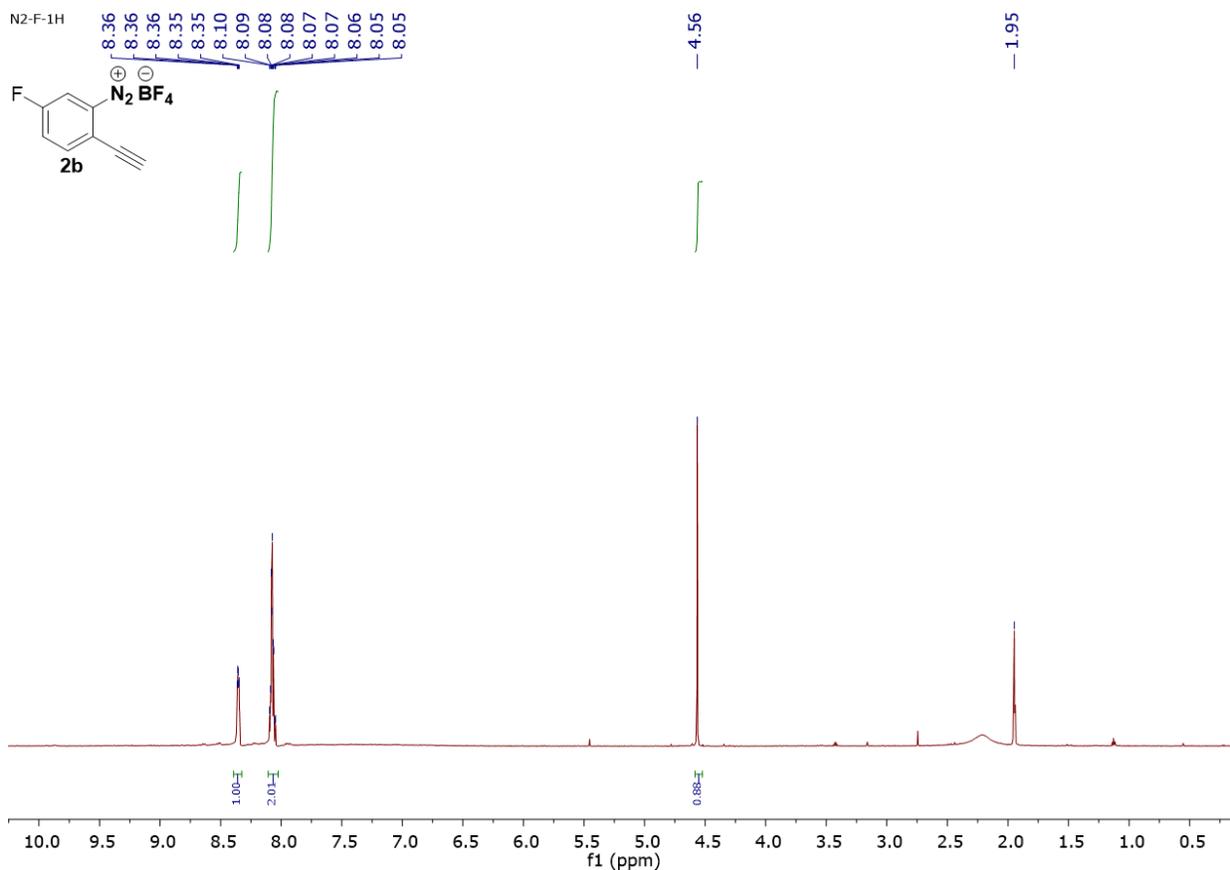


General procedure for synthesis of different diazonium salts

To the aniline derivative (1 eq., 4.27 mmol) in a mixture of 48% HBF₄ (1.5 mL), EtOH (1 mL) and H₂O (1.7 mL), the reaction mixture was cooled down to 0 °C, the NaNO₂ (1.2 equiv., 5.12 mmol) in distilled H₂O (1 mL) was added dropwise to the mixture over a period of 5 min. The reaction was allowed to stir at 0 °C for 45 min. The precipitate was collected, washed with cold Et₂O, and dried under vacuum to afford diazonium salt (2b-2d) as yellow or white powder (80%-92%).

Compound 2b

¹H NMR (600 MHz, Acetonitrile-*d*₃) δ 8.39 – 8.33 (m, 1H), 8.11 – 8.03 (m, 2H), 4.56 (s, 1H).

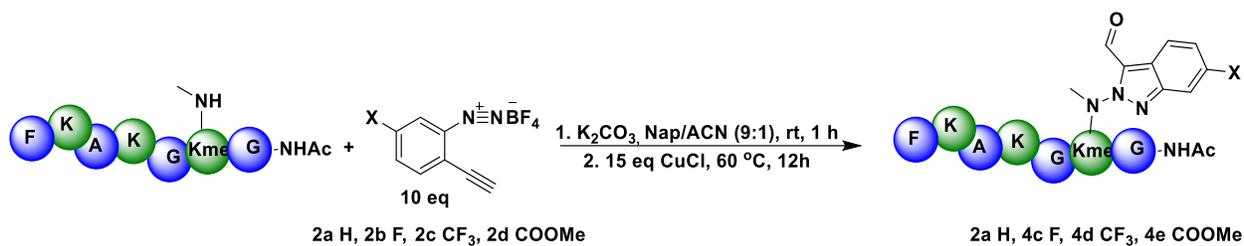


Procedure for triazination of kme peptide:

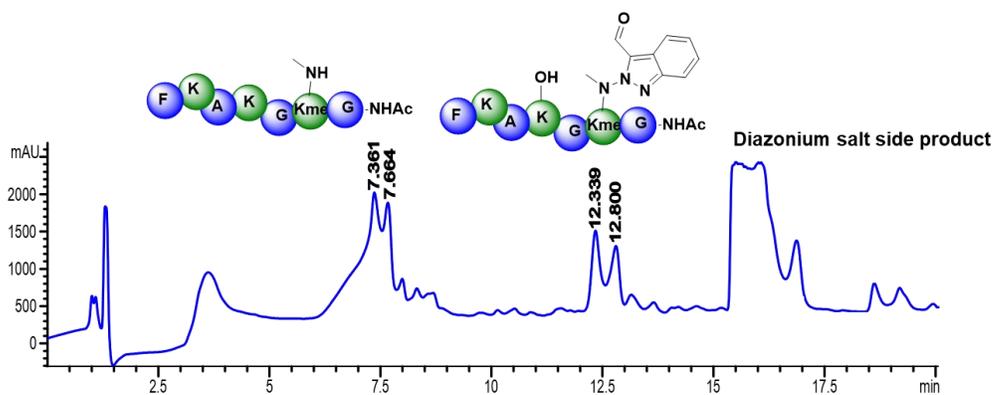
1 mg (1eq., 1.26 mmol) of Kme peptide with 10 eq of K_2CO_3 in 360 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of diazonium salt derivative in 40 μ L ACN were mixed together, the reaction mixture was incubated at 25 °C in the incubator for 1 h. The reaction mixture was directly used for next step without further purification.

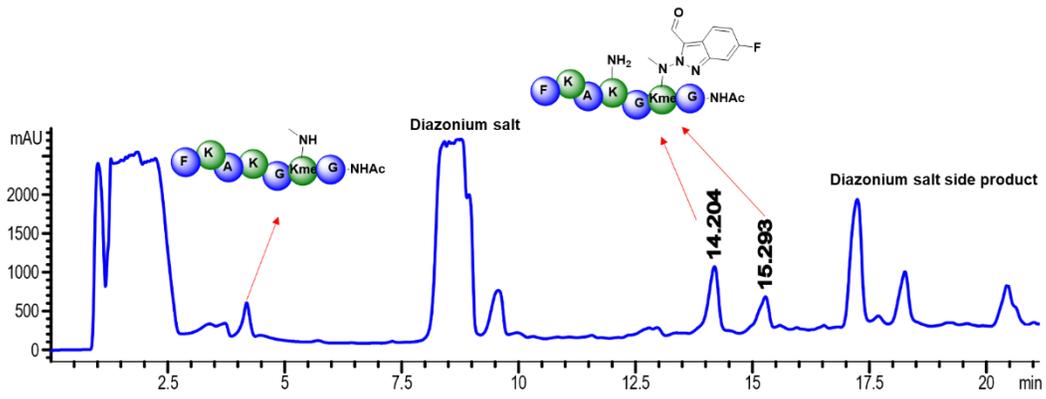
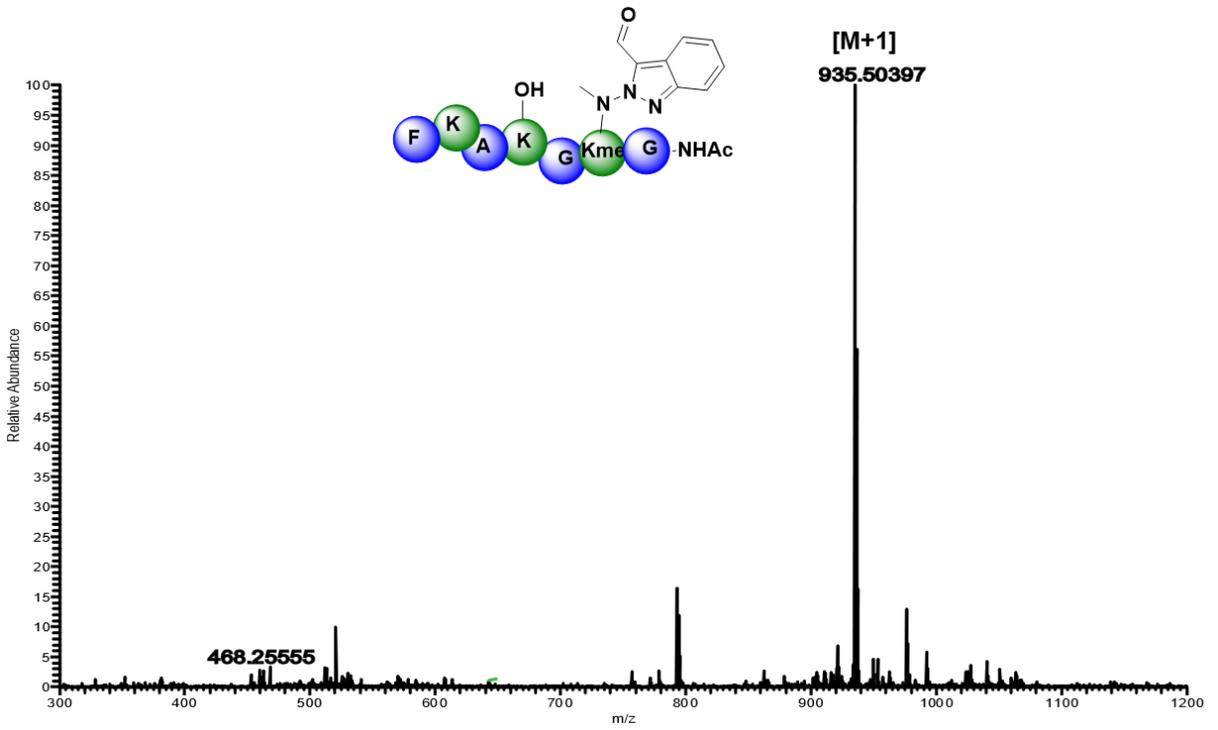
General procedure for cyclization of modified peptide:

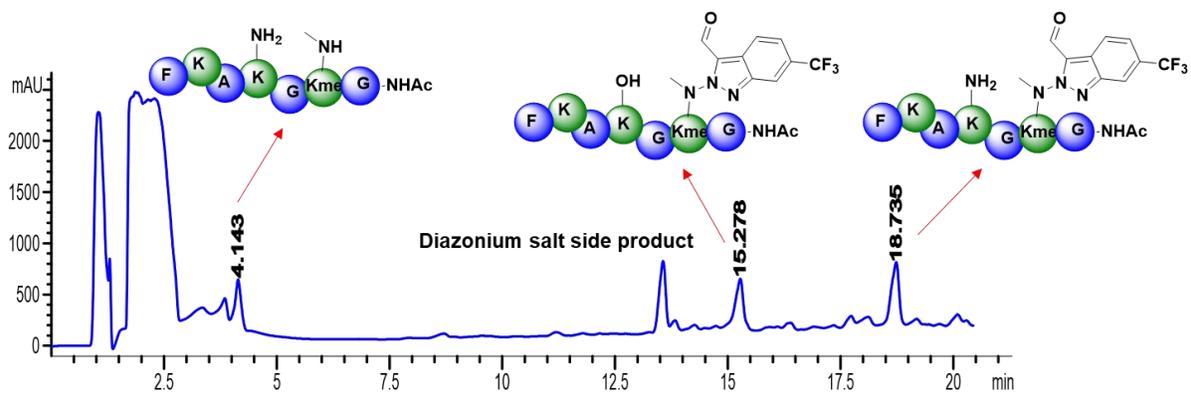
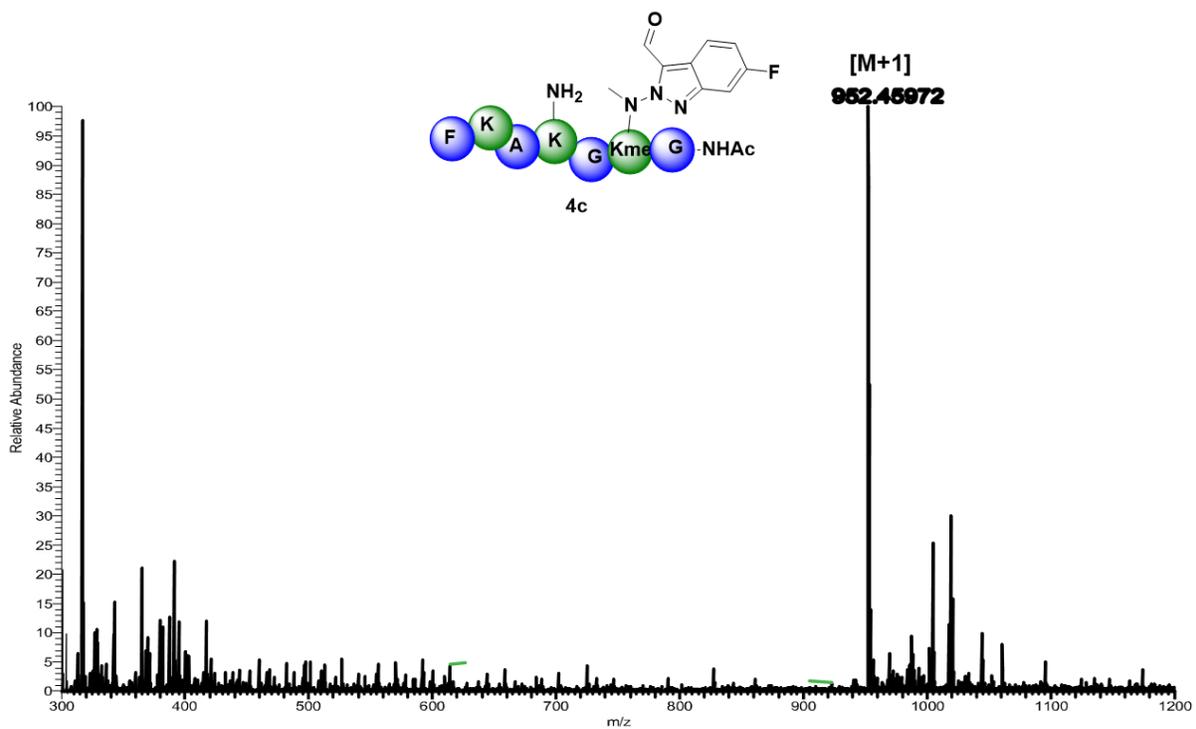
15 eq of CuCl in 100 μ l ACN was added into the reaction mixture, the reaction mixture was allowed to stir in incubator at 60 °C for 12 h. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

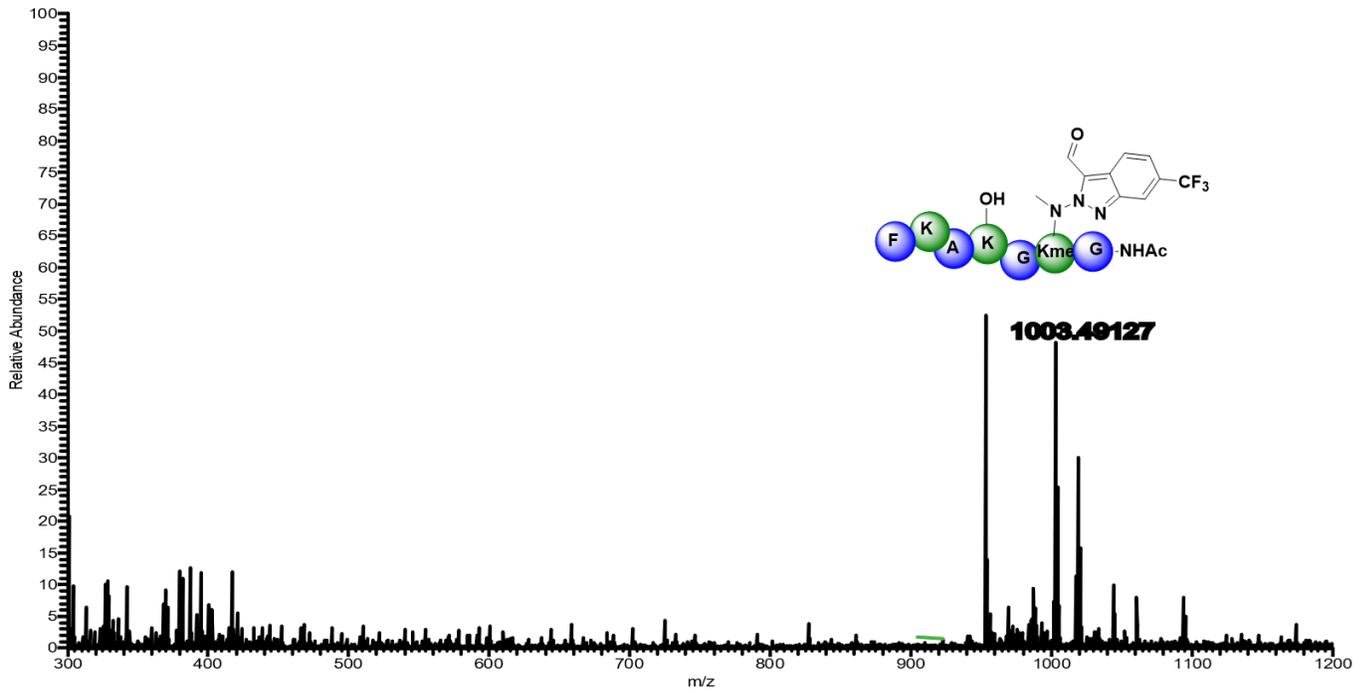
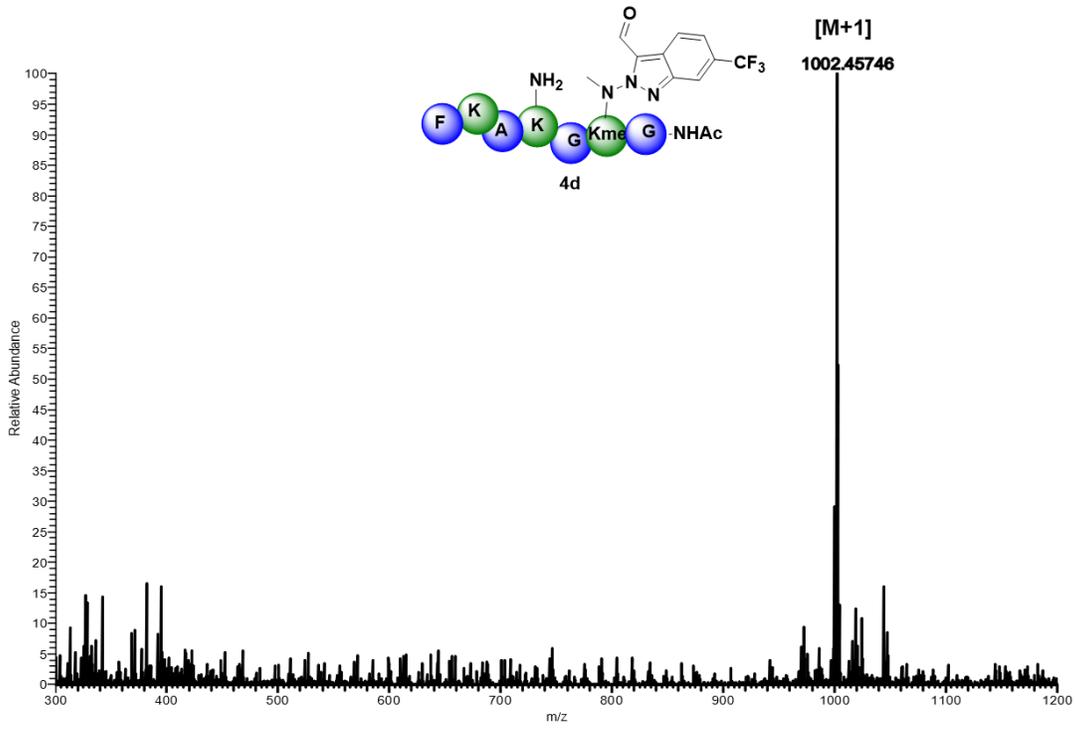


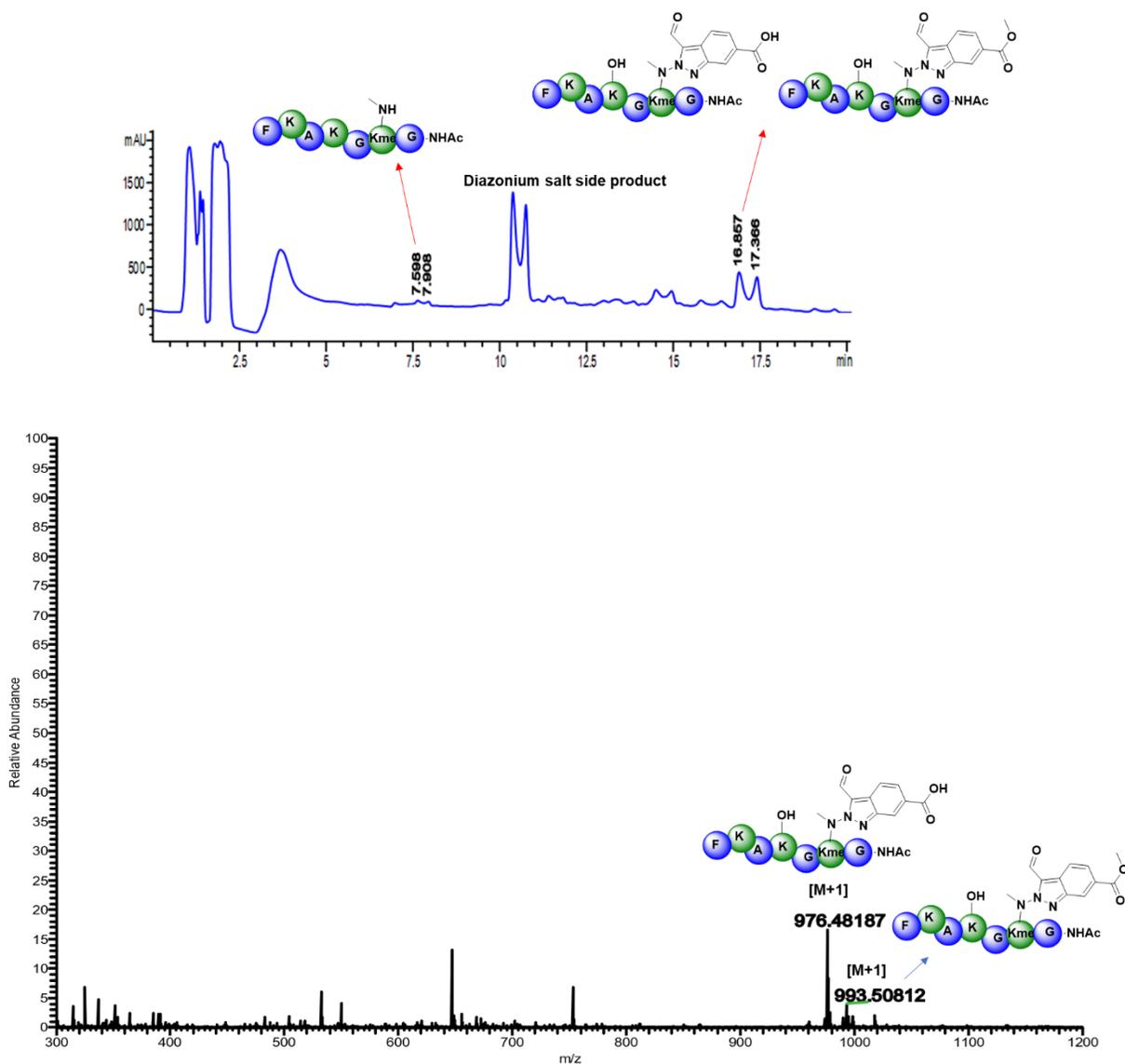
Analog	Conversion (%)
H	30%
F	71%
CF ₃	53%
COOMe	83%









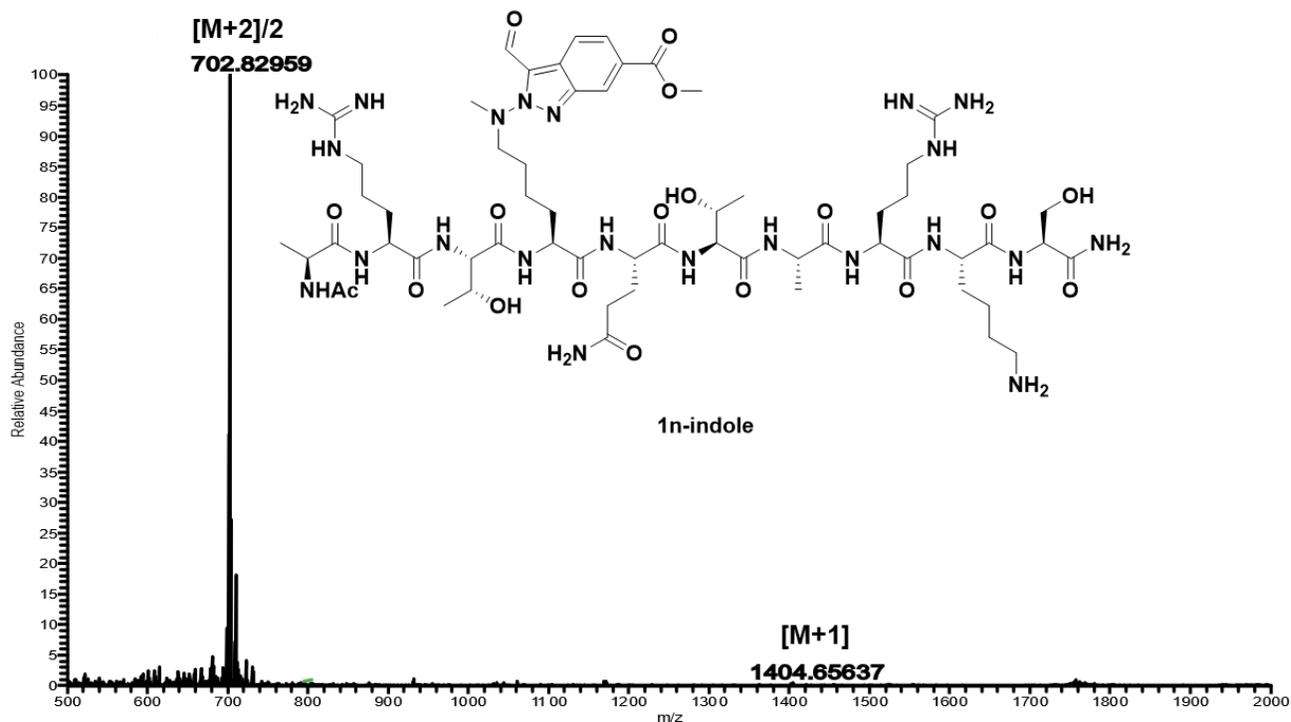
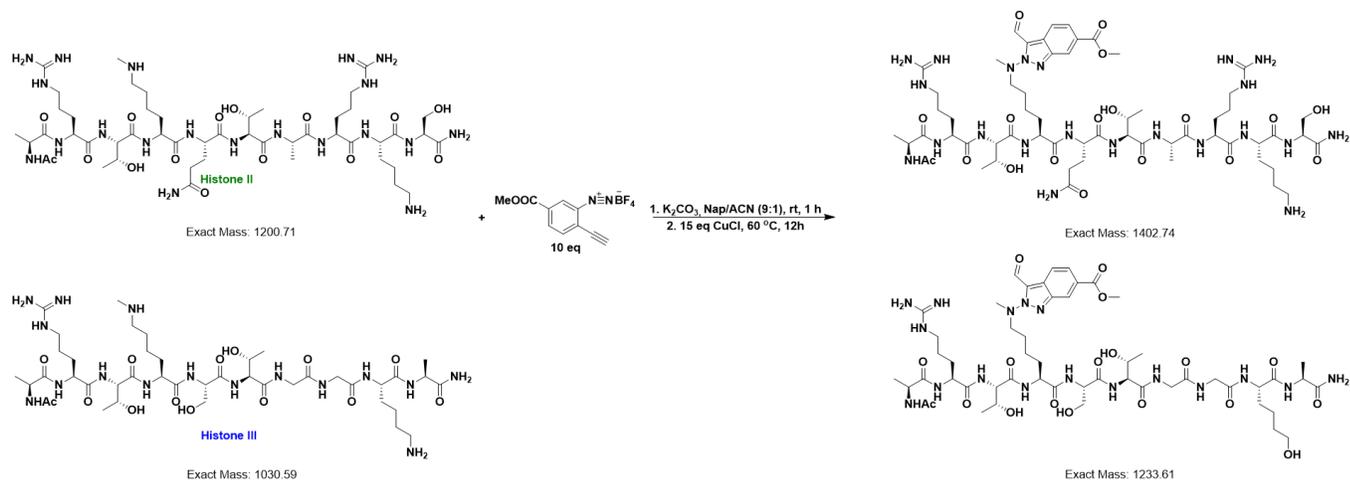


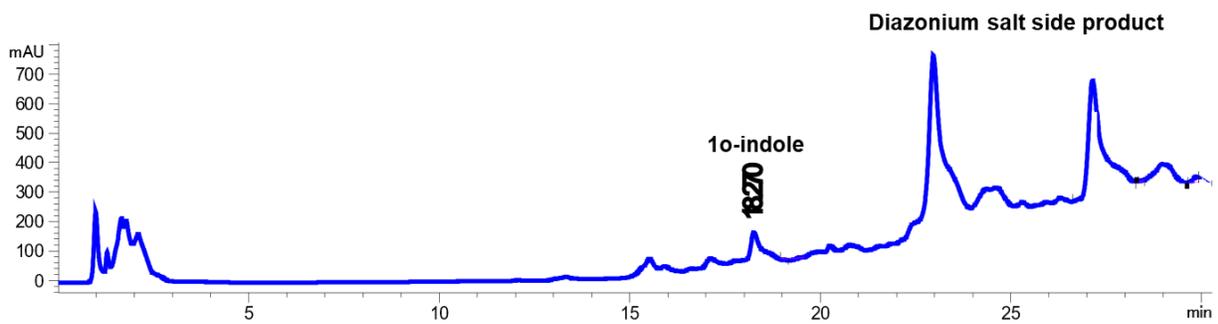
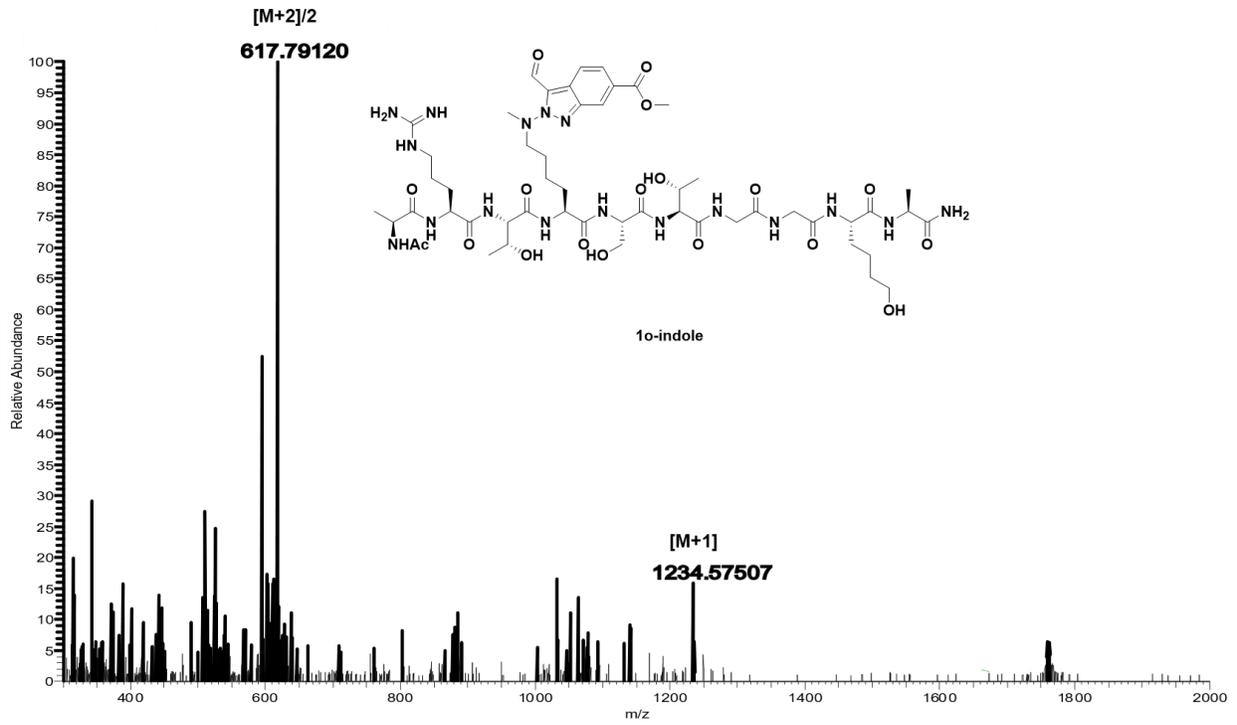
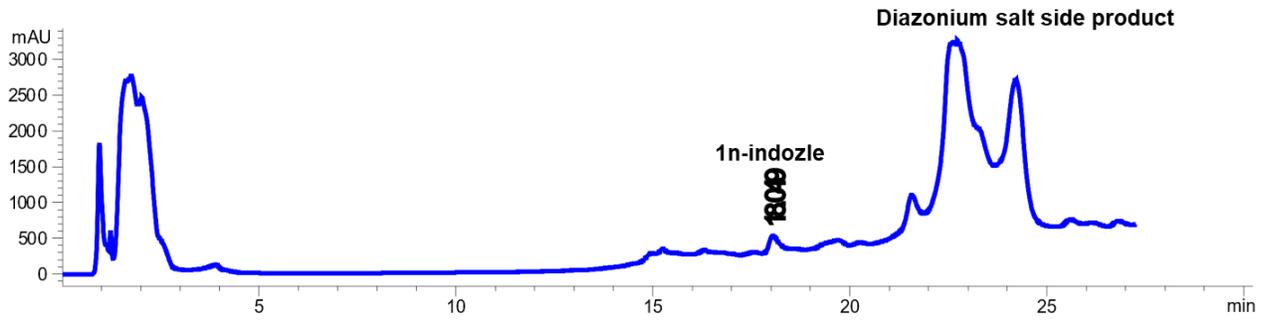
Supplementary Figure 8: Pan-specificity of triazening coarctate cyclization with truncated histone peptides.

General procedure for cyclization of modified histone peptide:

(1eq., 0.0027 mmol) of truncated histone peptides with 10 eq of K_2CO_3 (0.027 mmol) in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 10 eq of 2d (0.027 mmol) in 40 μ L ACN were mixed together, the reaction mixture was incubated at 25 $^{\circ}$ C for 1 h. The reaction mixture was directly used for next step without further purification. 15 eq of CuCl (0.04 mmol) in 100 μ L

ACN was added into the reaction mixture, the reaction mixture was allowed to stir in incubator at 60 °C for 12 h. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

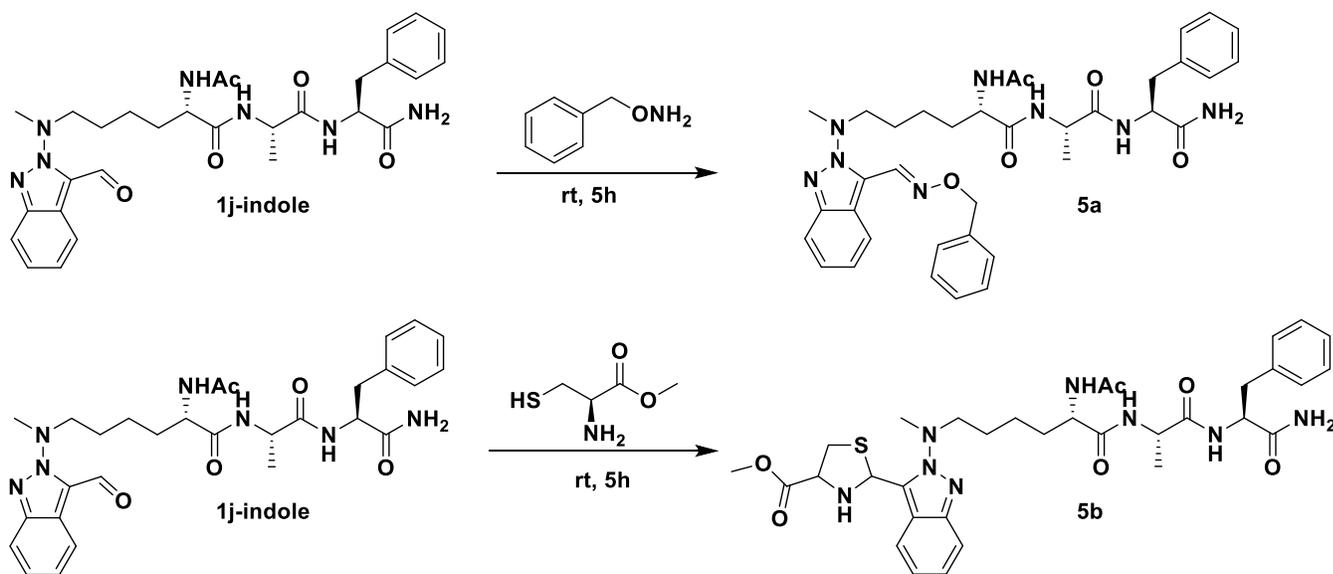


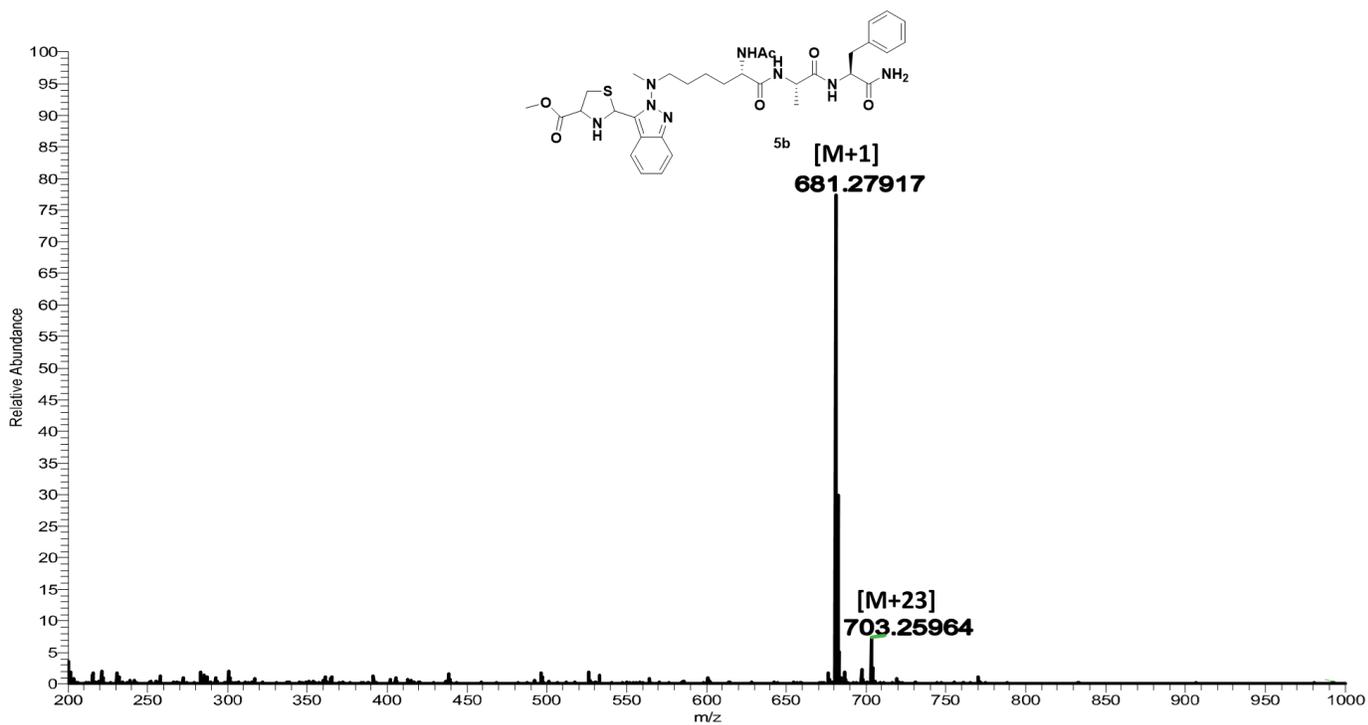
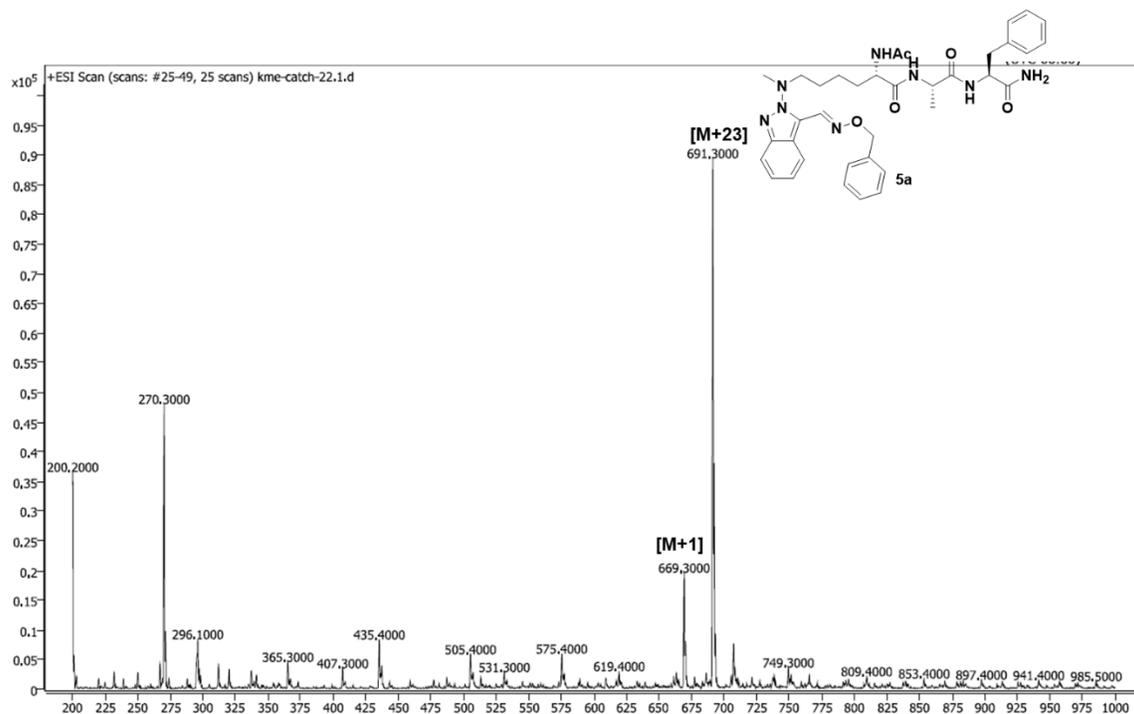


Supplementary Figure 9: Enrichment of modified peptides by cysteine condensation and oxime chemistry.

Procedure for enrichment of modified kme peptide with cysteine methyl ester and O-benzylhydroxylamine:

Modified Kme peptide (1eq., 0.027 mmol) in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 40 μ L of ACN, 5 eq of cysteine methyl ester or O-benzylhydroxylamine (5eq, 0.135 mmol) was added. The reaction mixture was incubated at 25 $^{\circ}$ C in the incubator for 5 h. The reaction mixture was filtered and purified by HPLC to obtain the product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

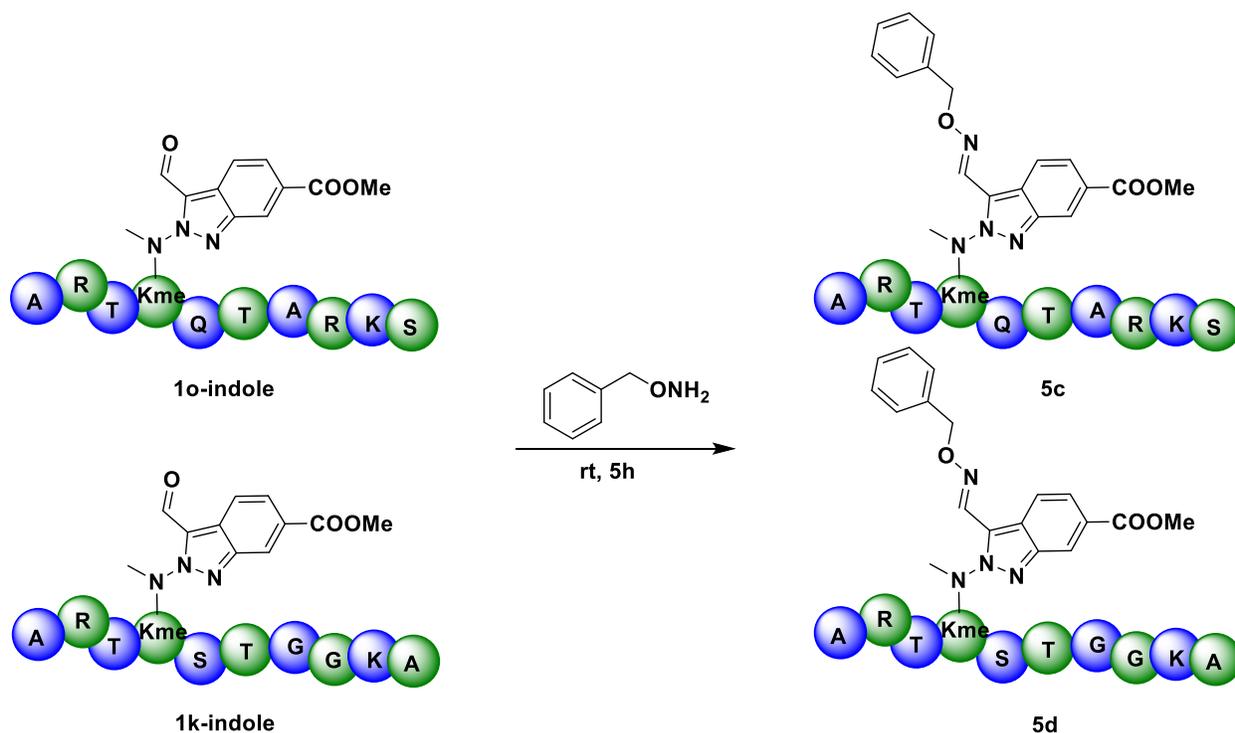


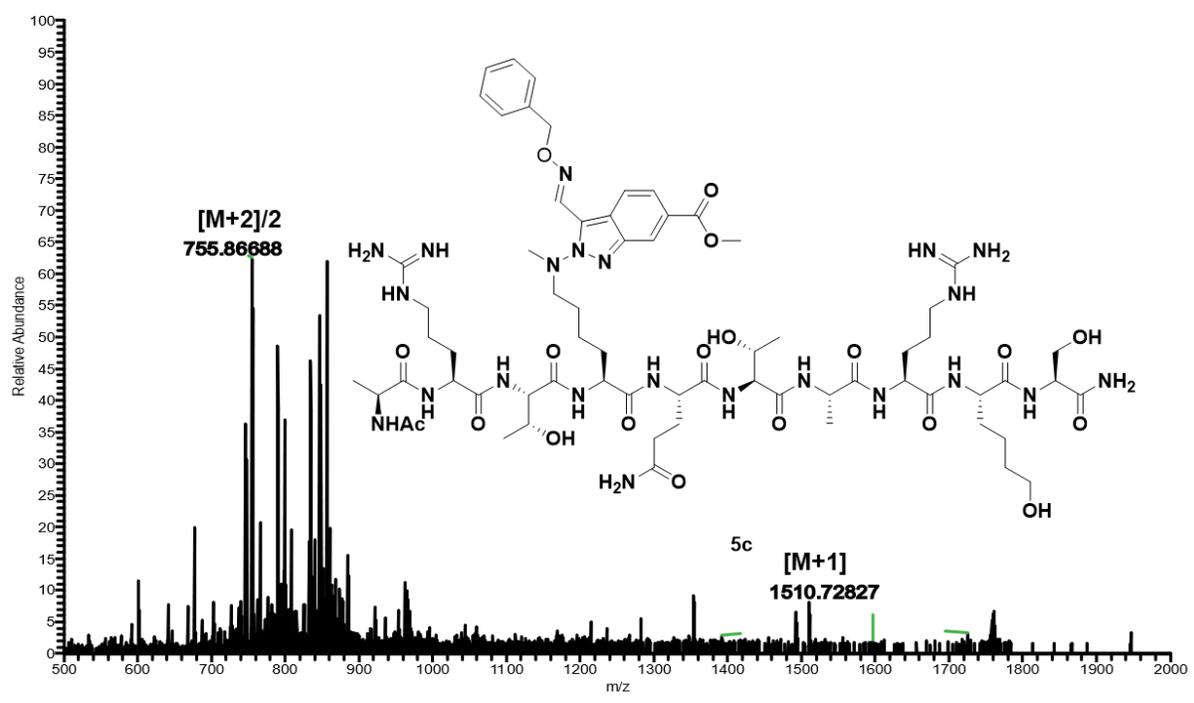
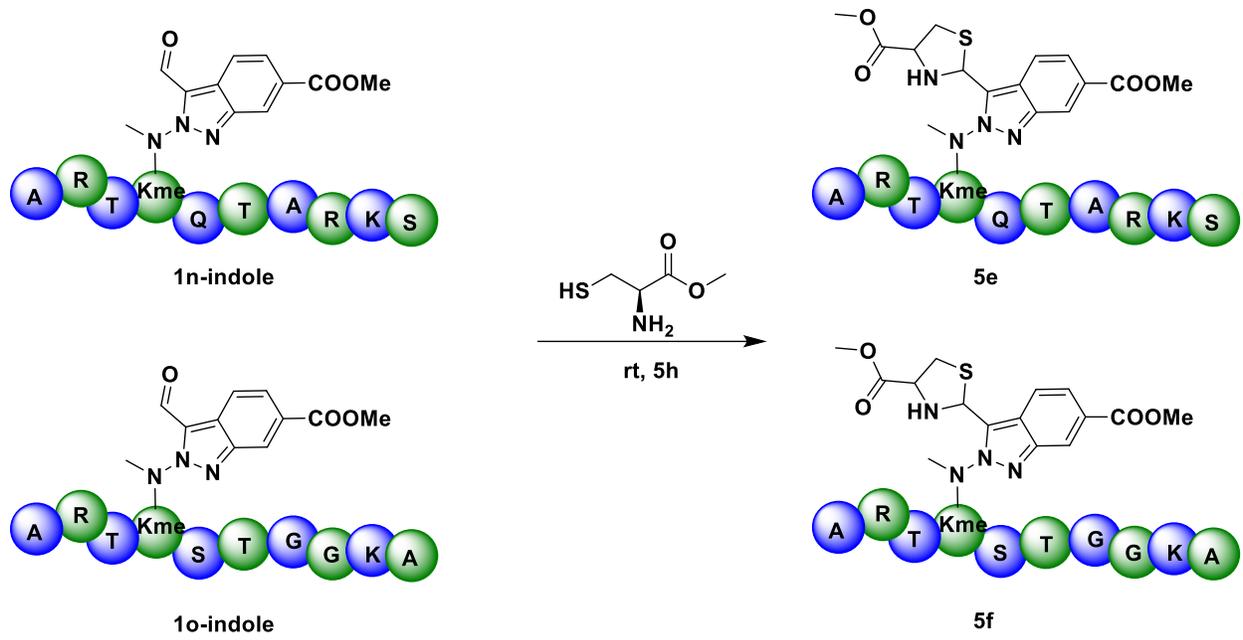


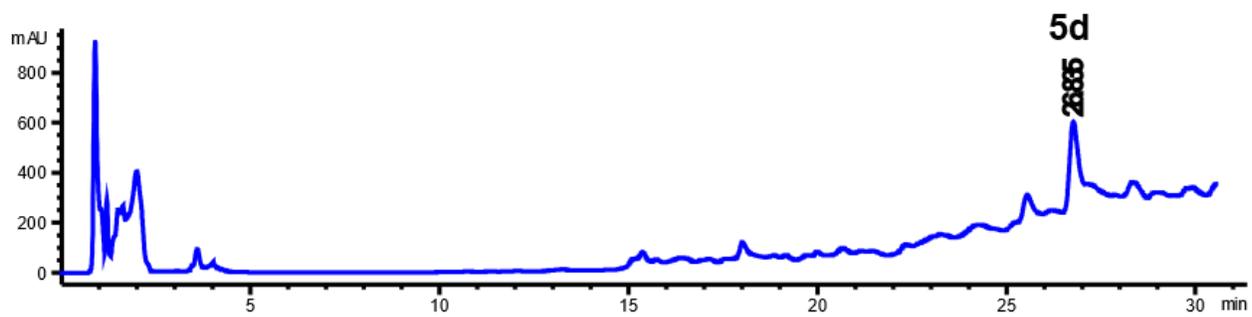
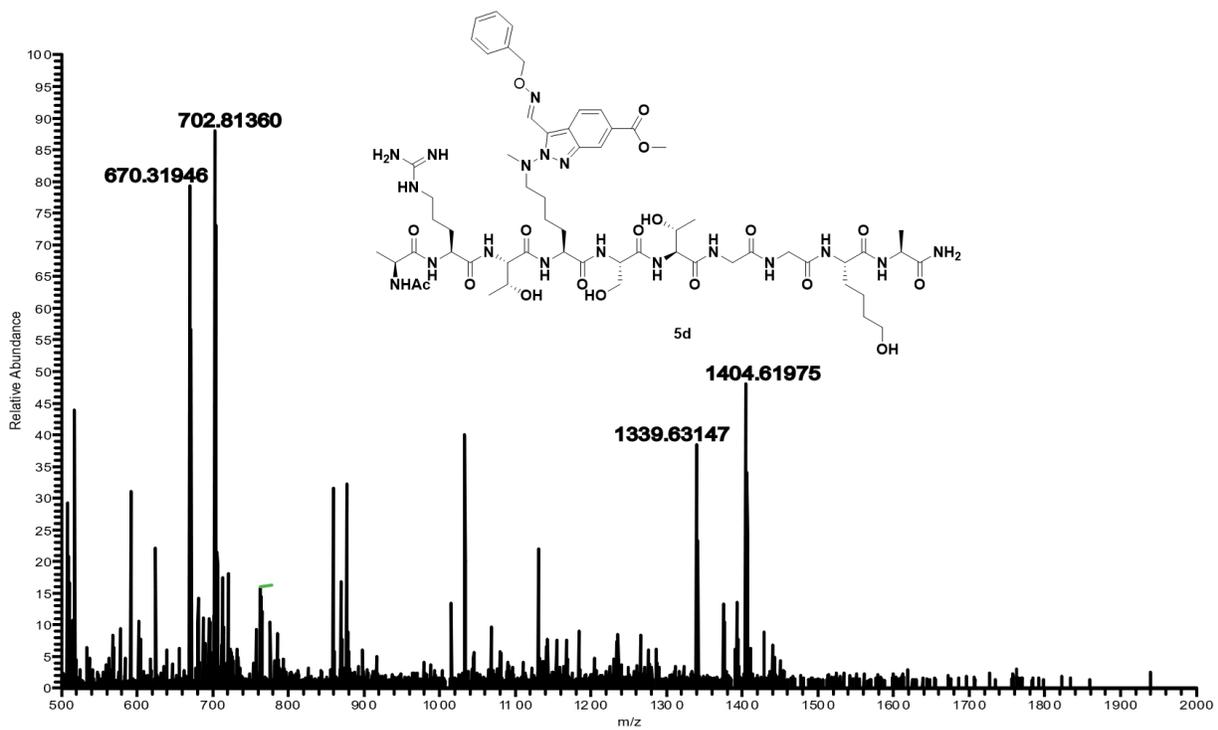
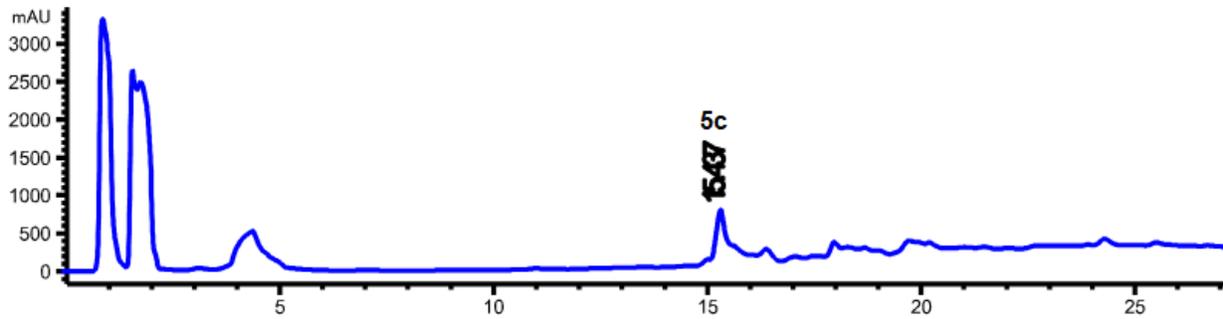
Supplementary Figure 10: Enrichment of modified truncated peptides by cysteine condensation and oxime chemistry.

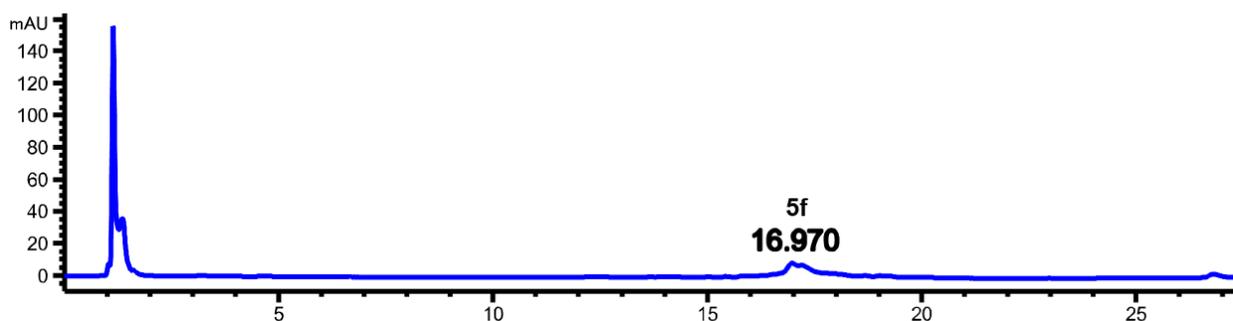
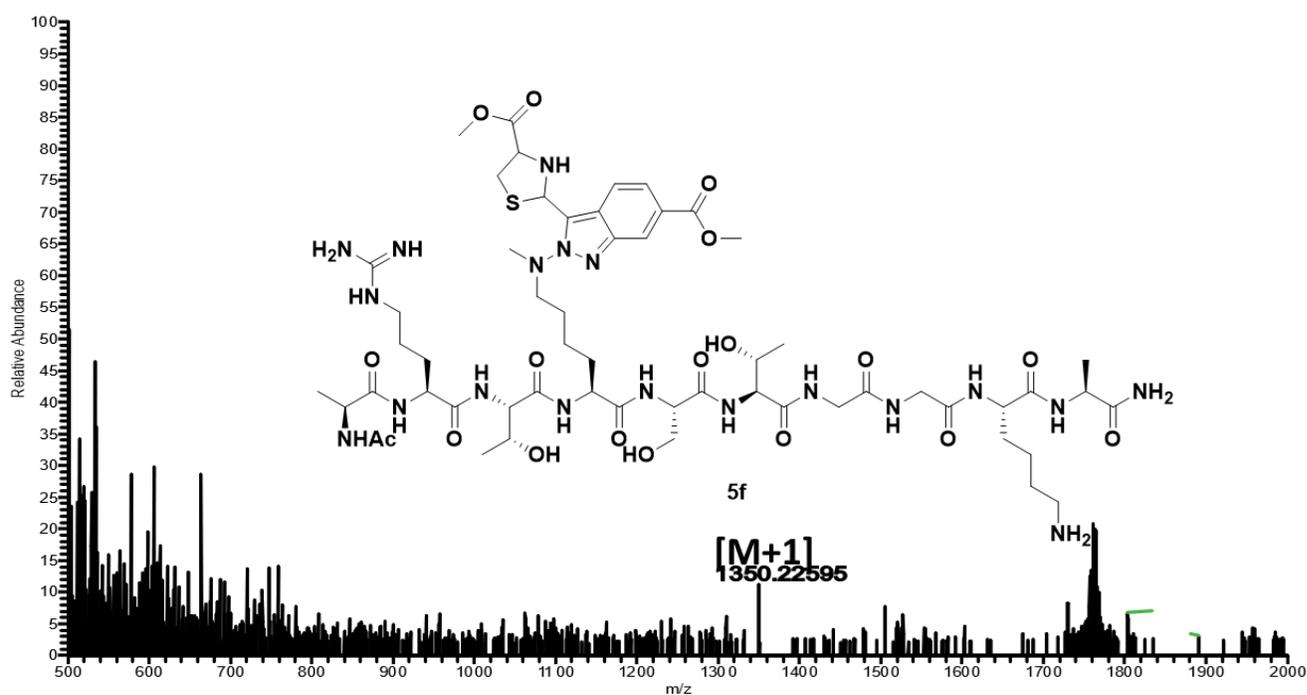
Procedure for enrichment of modified truncated peptide with cysteine methyl ester and O-benzylhydroxylamine:

Modified Kme peptide (1eq., 0.027 mmol) in 360 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 40 μ L of ACN, 5 eq of cysteine methyl ester or O-benzylhydroxylamine (5eq, 0.135 mmol) was added. The reaction mixture was incubated at 25 $^{\circ}$ C in the incubator for 5 h. The reaction mixture was filtered and purified by HPLC to obtain the product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.









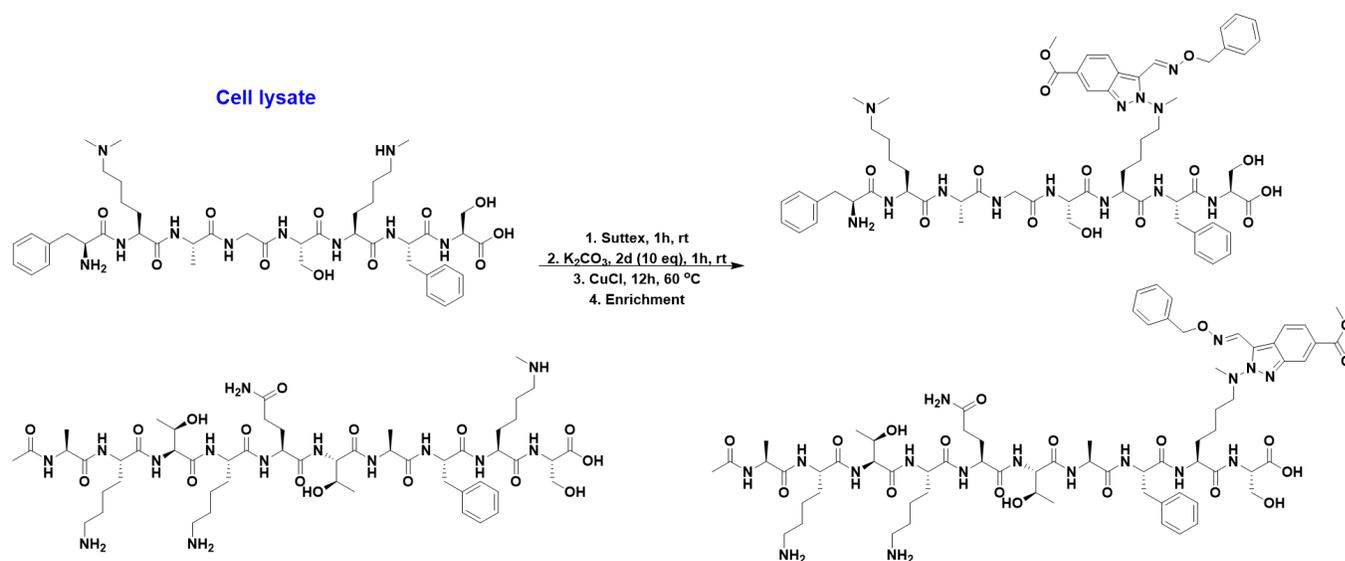
Supplementary Figure 11: Selective enrichment of Kme1 containing peptides in a complex cell lysate mixture by TCC

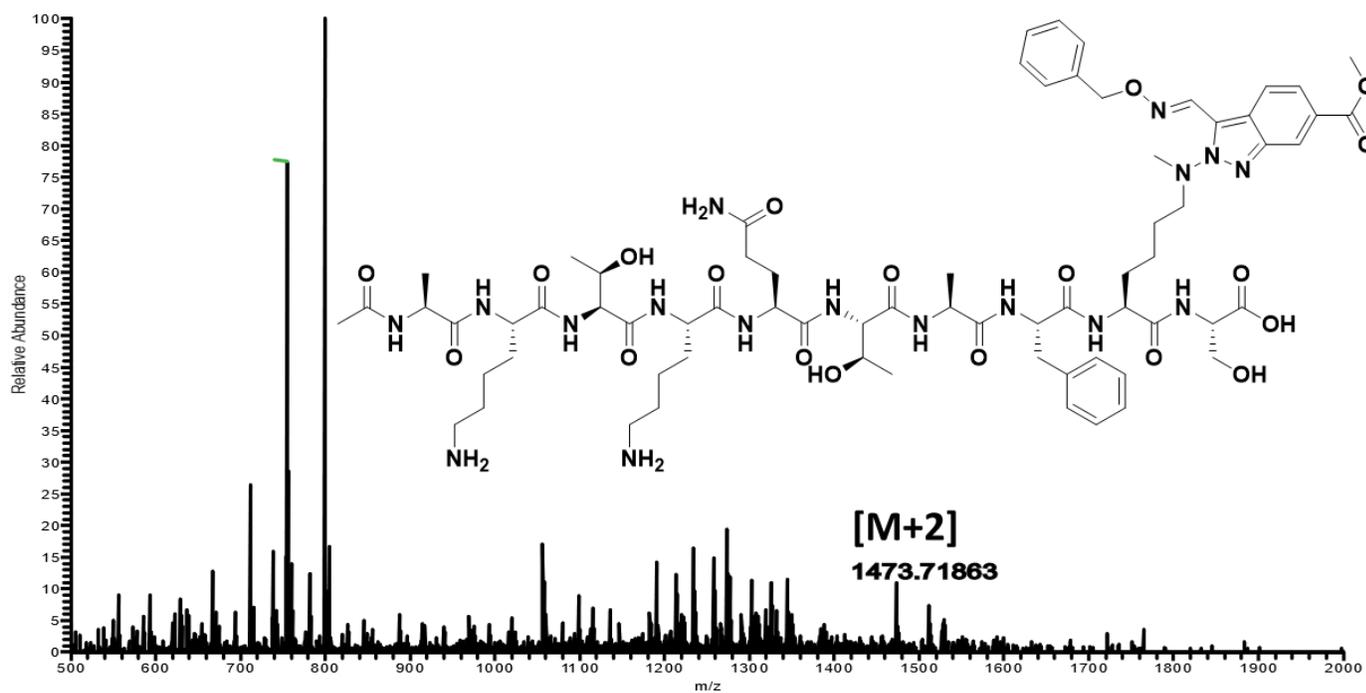
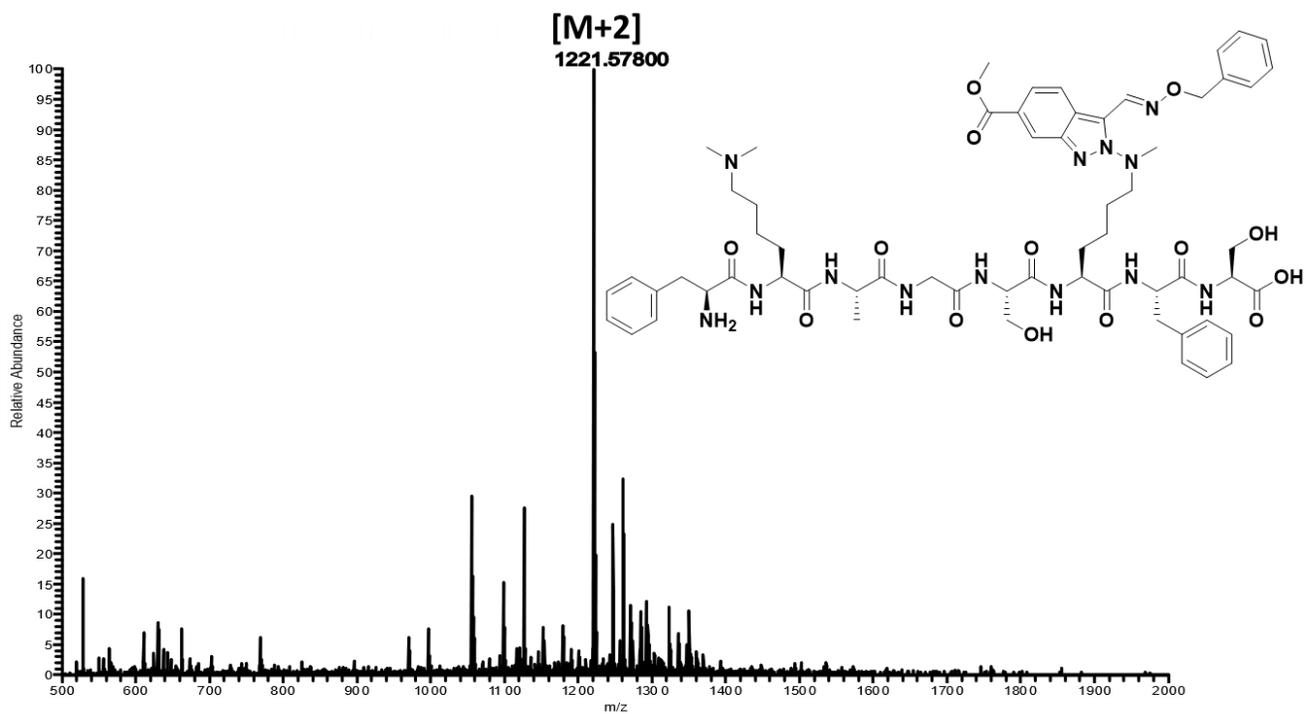
General procedure for tyrosine blocking bySuttex

To the peptide NH₂-FK(me)₂AGSSKmeFS (1 eq., 0.00088 mmol), Ac-AKTK(me)QTAFKS (1 eq., 0.00088 mmol) and cell lysate (100 ul) in 200 μL of 10 mM sodium phosphate buffer (Nap, pH 7.5), 2mg of Suttex (10 eq, 0.0088 mmol) was added, the reaction was allowed to stir at 25 °C for 60 min. The reaction mixture was directly used for next step without further purification.

Procedure for oxime enrichment of Kme1 containing peptide:

20 eq of K_2CO_3 (0.0176 mmol) in 50 μ L of 10 mM sodium phosphate buffer (Nap, pH 7.5) and 20 eq of 2d (0.0176 mmol) in 50 μ L ACN were added, the reaction mixture was incubated at 25 $^\circ$ C for 1 h. The reaction mixture was directly used for next step without further purification. 30 eq of CuCl (0.0264 mmol) in 50 μ L ACN was added into the reaction mixture, the reaction mixture was stirred in incubator at 60 $^\circ$ C. After 12 h, O-benzylhydroxylamine (15 eq, 0.014 mmol) was added to the mixture, the reaction mixture was incubated at 25 $^\circ$ C in the incubator for 5 h. The crude was filtered and purified by HPLC to obtain the oxime product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm



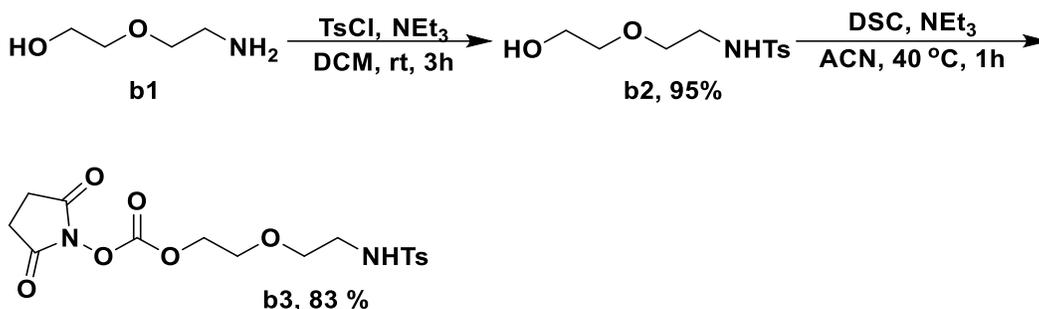


Supplementary Figure 12: Single-molecule sequencing for identification of Kme1 sites by TCC

Supporting Information for Chapter Three:

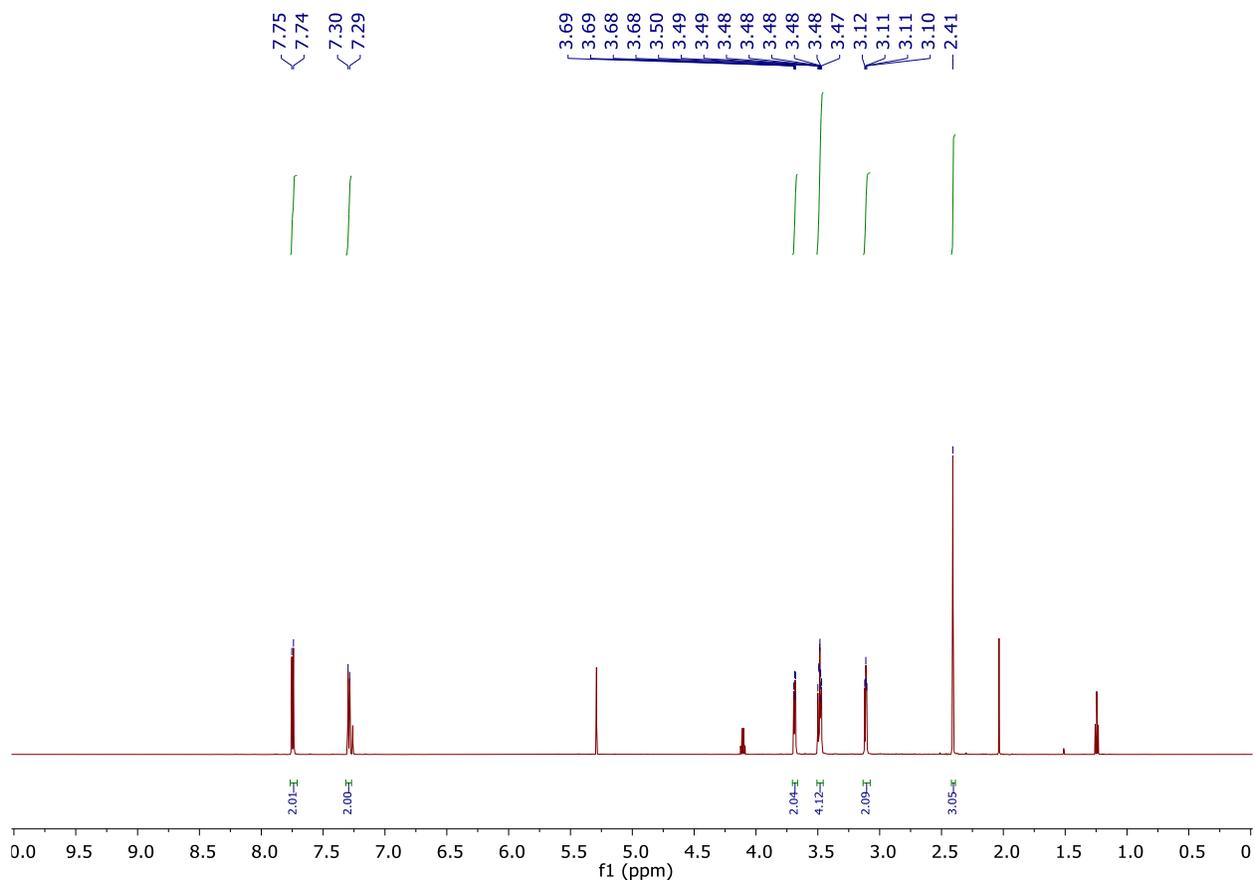
Supplementary Figure 1. Synthesis of different K-me-DM probes

Synthesis of intermediate b3



Synthesis of intermediate b2

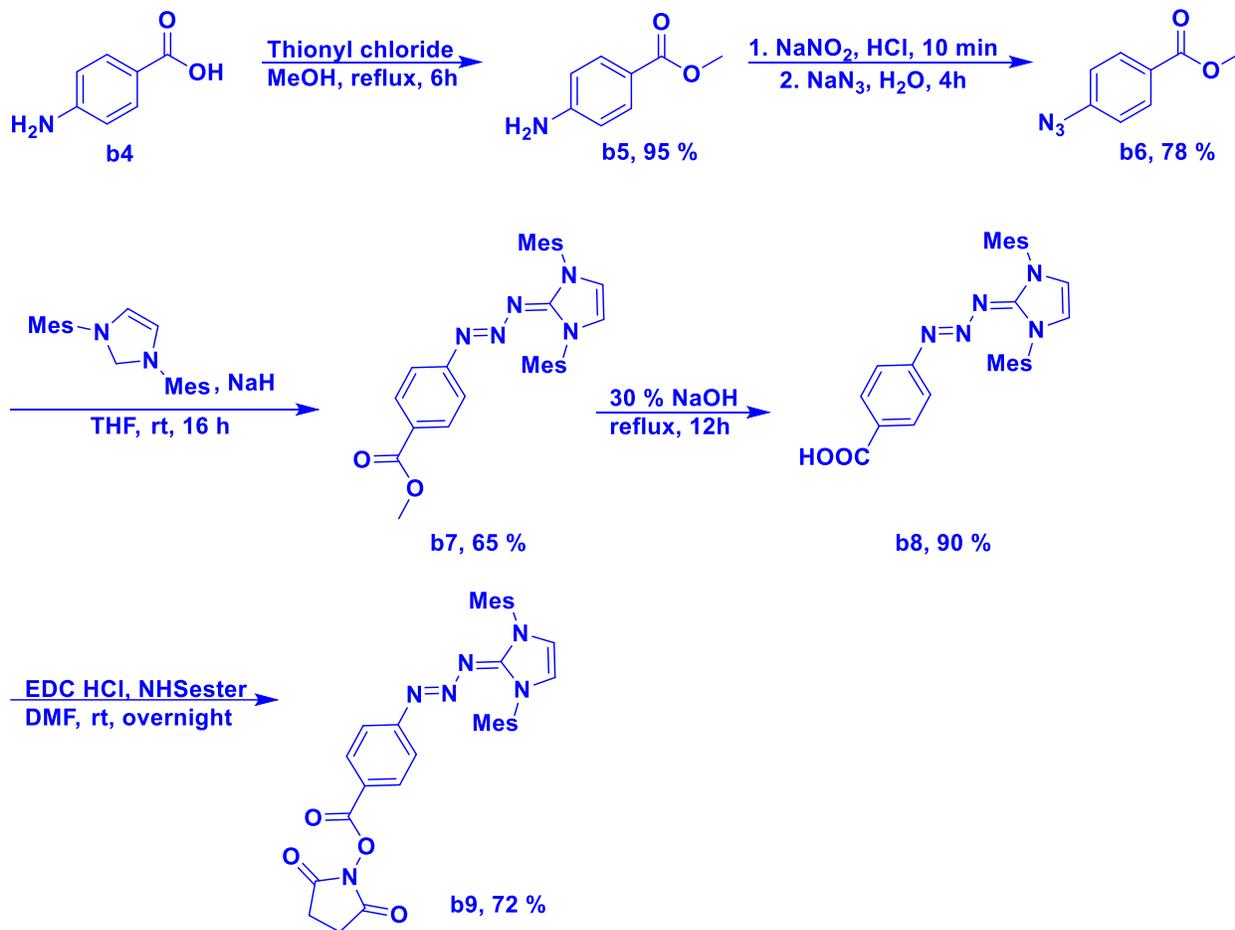
2-(2-aminoethoxy)ethan-1-ol **b1** (3.1 g, 21.0 mmol), Net_3 (4.4 ml, 31.5 mmol), and TsCl (4 g, 21 mmol) were added sequential into dry DCM at 25°C . The mixture was warmed to room temperature and stirred for 3 hours. The reaction mixture was washed with saturated aqueous NaHCO_3 solution for 3 times. The organic layer was collected, dried over anhydrous MgSO_4 , filtered, and concentrated under the reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 2:1) to yield compound **b2** as colorless oil (5.16 g, 95 %). $^1\text{H NMR}$ (600 MHz, Chloroform- d) δ 7.75 (d, $J = 8.3$ Hz, 2H), 7.29 (d, $J = 7.9$ Hz, 2H), 3.71 – 3.66 (m, 2H), 3.48 (ddd, $J = 7.9, 5.0, 4.0$ Hz, 4H), 3.11 (dd, $J = 5.4, 4.5$ Hz, 2H), 2.41 (s, 3H).



Synthesis of intermediate b3

Compound b2 (1 g, 3.86 mmol) was dissolved in 28 mL dry ACN. To the solution, triethylamine (1 mL, 7.72 mmol) and disuccinimidyl carbonate (DSC, 2 g, 7.72 mmol) was added. The reaction was stirred at 40 °C. After 1 hour The reaction mixture was concentrated under reduced pressure and purified by flash column chromatography (hexane: ethyl acetate 1:1) to obtain compound b3 as colorless oil (1.29 g, 83 %).

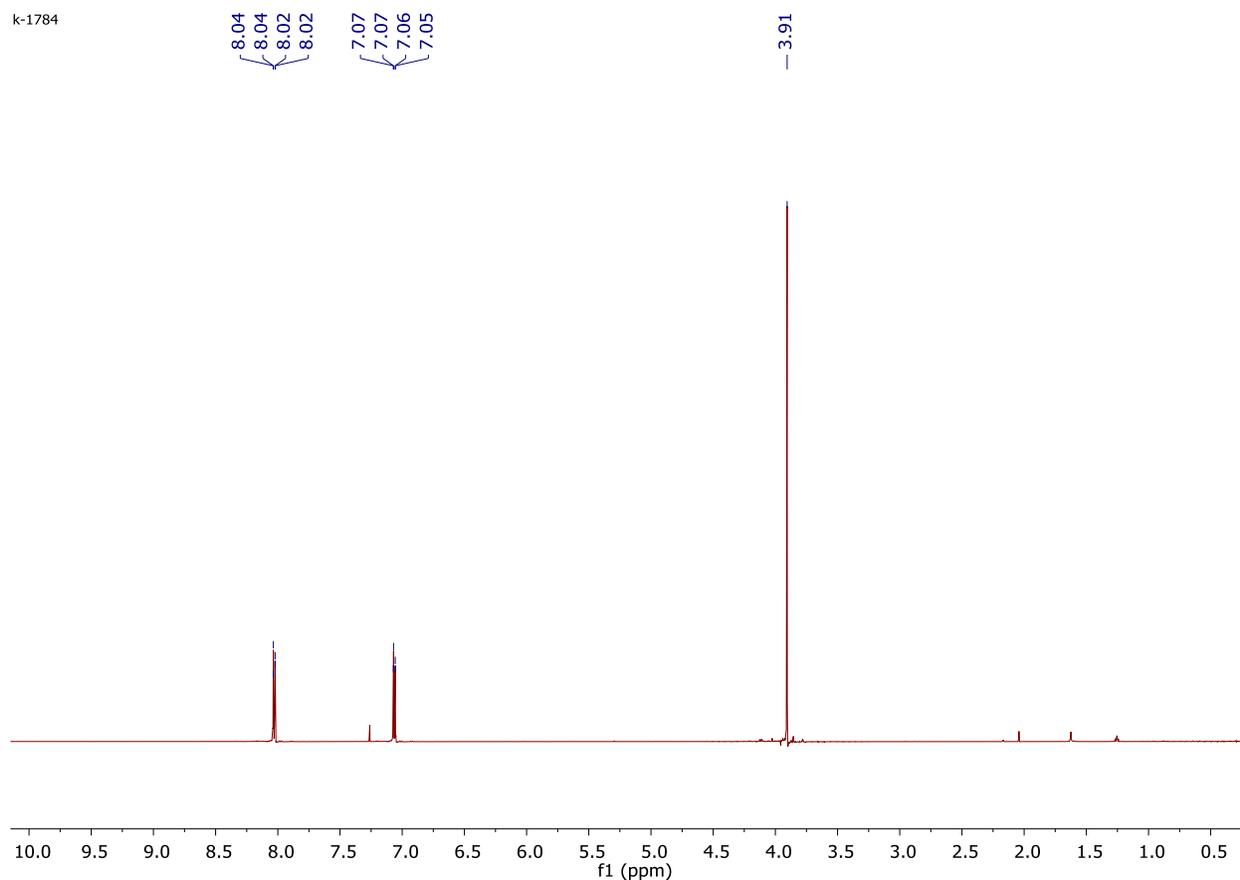
Synthesis of compound b9



Synthesis of intermediate b5

4-Aminobenzoic acid (10 g, 73 mmol) was dissolved in dried methanol 100 ml and the mixture was cooled with an ice bath to 0°C. Subsequently, thionyl chloride (15.8 mL, 219 mmol) was added dropwise and the solution was stirred at room temperature overnight. The solution was neutralized with saturated NaHCO₃-solution and then K₂CO₃ was added till a pH-value of eight was acquired. Precipitating salts were dissolved by addition of water. Afterwards the organic phase was extracted with DCM and the combined organic layers were dried over sodium sulfate, filtered off, and concentrated under the reduced pressure to obtain compound b5 as white solid (10.4 g, 95%). ¹H NMR (600 MHz, Chloroform-*d*) δ 8.06 – 8.00 (m, 2H), 7.09 – 7.02 (m, 2H), 3.91 (s, 3H).

k-1784

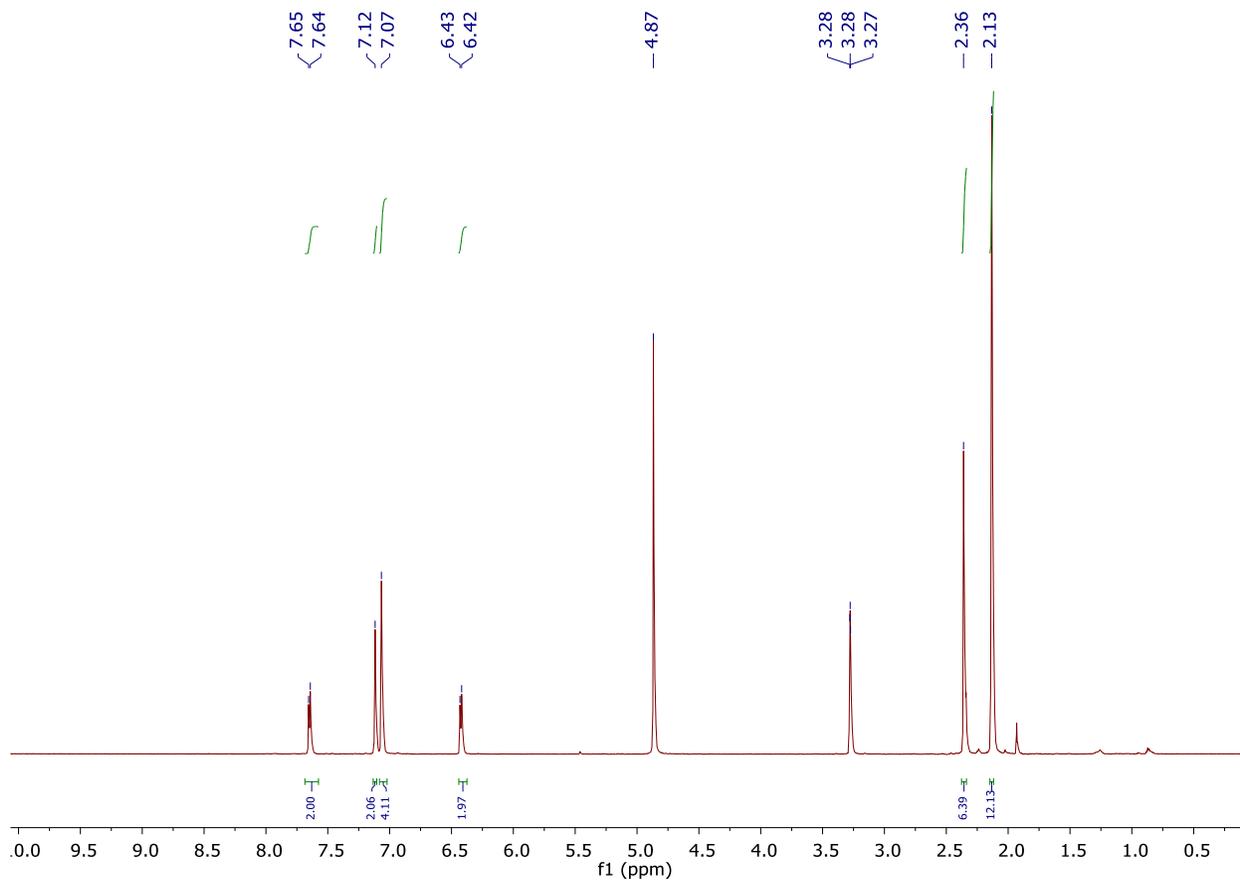


Synthesis of intermediate b6

Compound b5 (3 g, 19.8 mmol) was dissolved in concentrated HCl 13 ml and H₂O 7 ml and the mixture was cooled to 0 °C. A solution of NaNO₂ (1.34 g, 20.06 mmol) in 3 ml of water was added dropwise with a dropping funnel and the reaction was allowed to stir for 10 min at 0 °C. At 0 °C, a solution of NaN₃ (1.3 g, 20.06 mmol) in 7 ml of water was added dropwise with a dropping funnel and the reaction was allowed to heat to r.t for 4 h. The crude reaction mixture was extracted two times with diethyl ether and one time with water. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure yielding the azide as yellow crystal (2.74 g, 78%). The crude products were directly used for the next step without the purification.

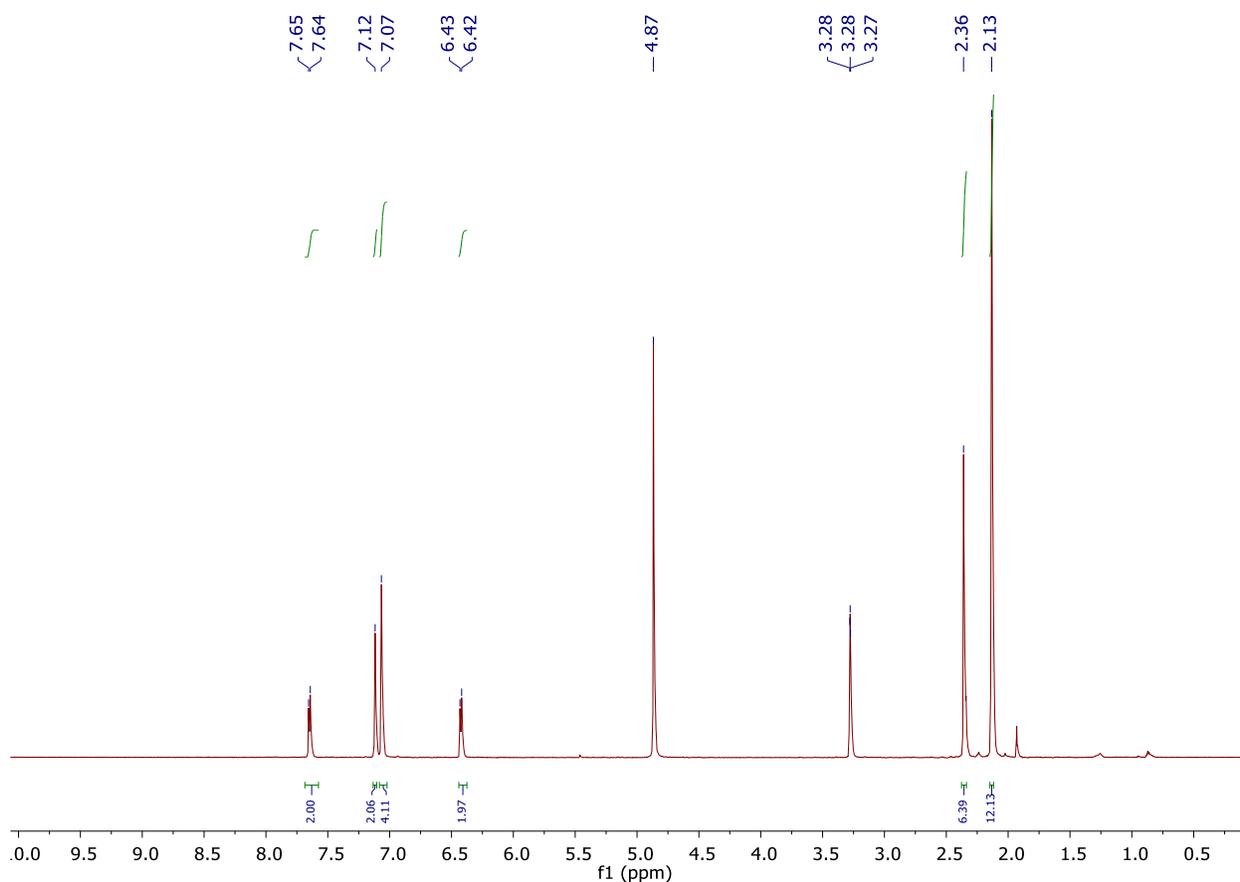
Synthesis of intermediate b7

To a solution of methyl 4-azidobenzoate b6 (0.4 g, 1.18 mmol) in 10 mL dry THF, 1,3-dimesitylimidazolium chloride (0.2 g, 1.18 mmol) was added and the solution was cooled for 20 min. To this solution, 0.11 g NaH (1.24 mmol, 60% in mineral oil) was added at 0 °C, and the reaction was allowed to stir for 18 h at room temperature. Next, the reaction was quenched with 10 mL water and 30 mL ethyl acetate was added to it. The organic layer washed with 10 mL brine and dried on anhydrous sodium sulfate. Finally, the organic layer was evaporated to give compound b7 as bright yellow solid (369.1 mg, 65%). ¹H NMR (600 MHz, Chloroform-d) δ 7.69 (d, J = 8.8 Hz, 2H), 7.00 (s, 4H), 6.63 (d, J = 1.1 Hz, 2H), 6.57 (d, J = 8.8 Hz, 2H), 3.85 (s, 3H), 2.37 (s, 6H), 2.15 (s, 12H).



Synthesis of intermediate b8

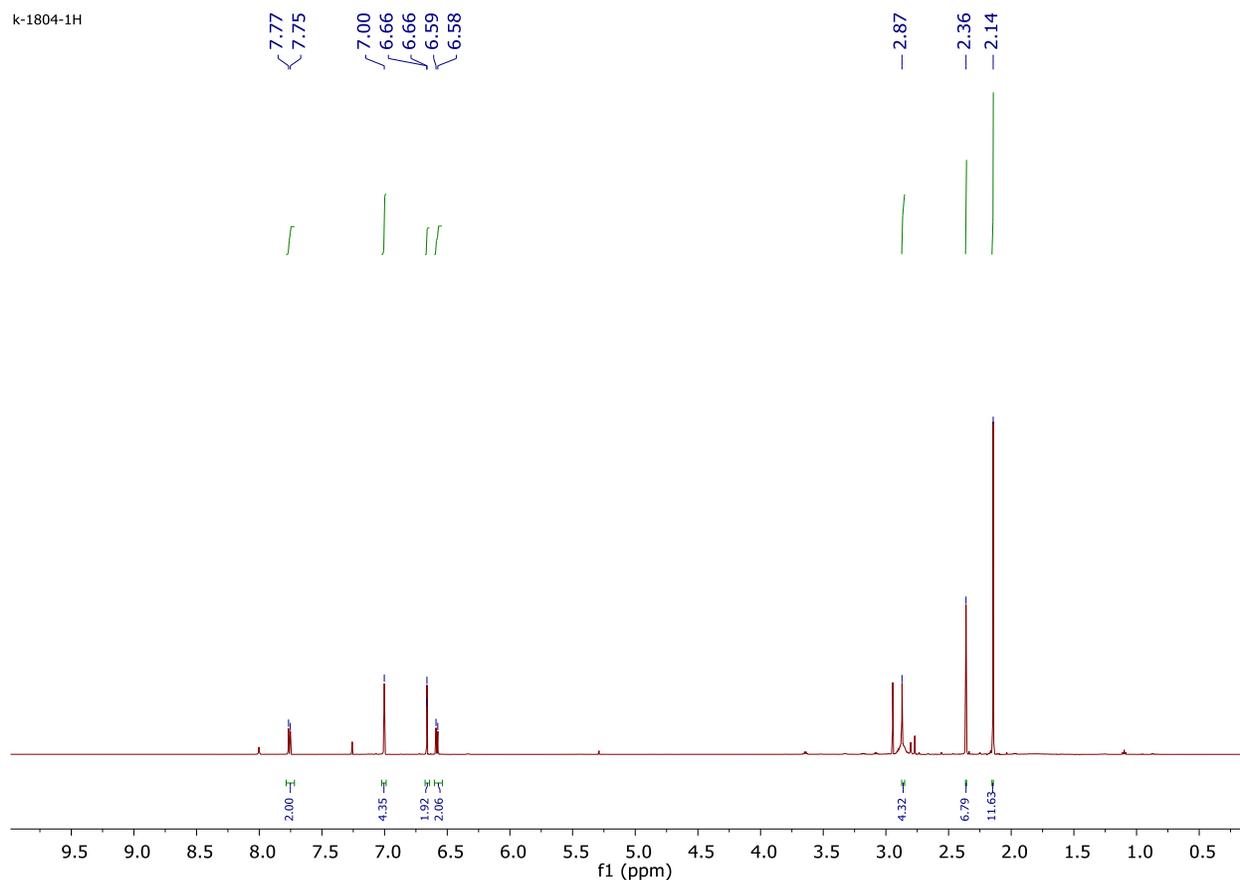
To a methanolic solution of intermediate 7 (57 mg, 0.118 mmol in 3 mL methanol), 380 μ l of NaOH solution was added (from a stock solution of 0.1 g NaOH in 5 mL water) in a drop-wise manner and the reaction was refluxed for 12 h. Next, few drops of glacial acid were added to neutralize the reaction mixture. The precipitate was collected and dried under reduced pressure to obtain compound b6 as bright yellow powder (49.6 mg, 90 %). $^1\text{H NMR}$ (600 MHz, Methanol- d_4) δ 7.65 (d, $J = 8.2$ Hz, 2H), 7.12 (s, 2H), 7.07 (s, 4H), 6.42 (d, $J = 8.2$ Hz, 2H), 2.36 (s, 6H), 2.13 (s, 12H).

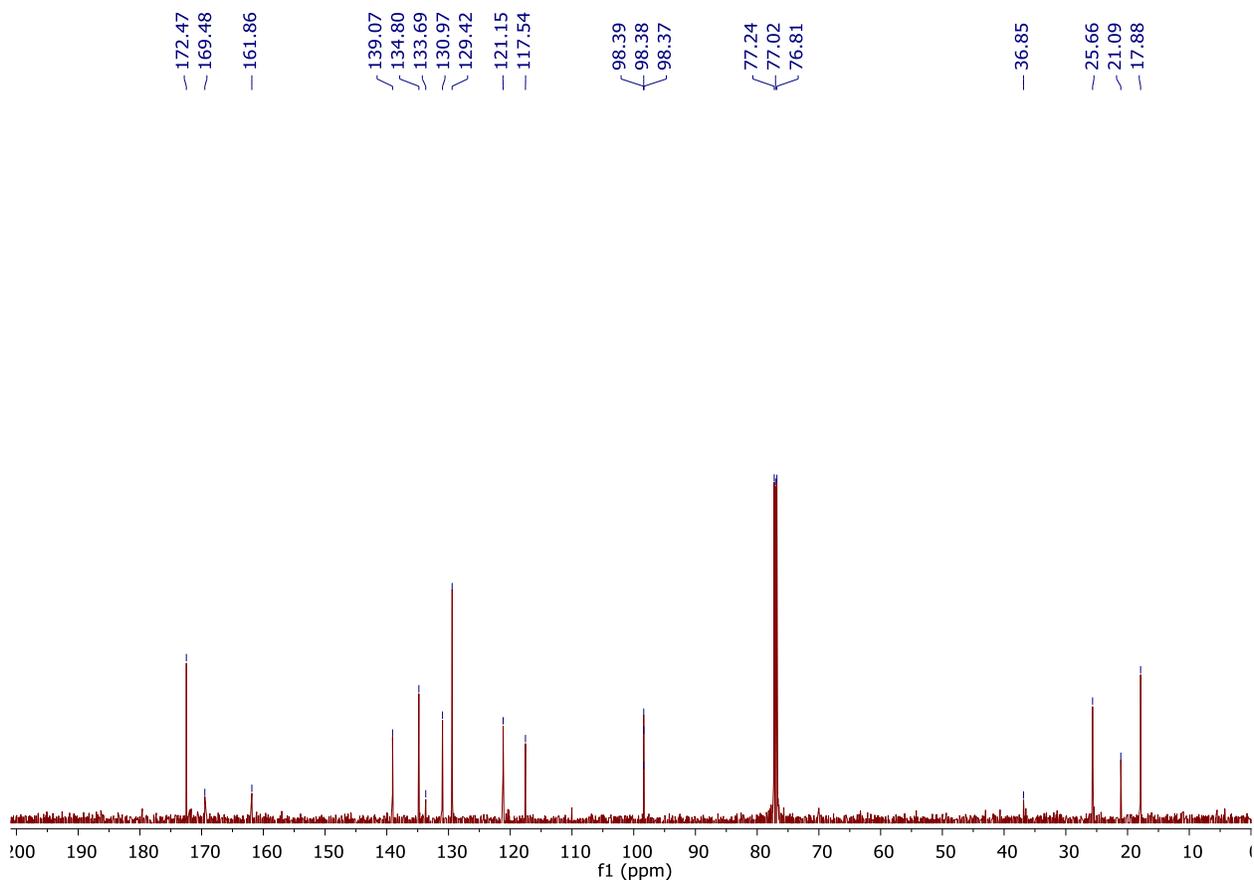


Synthesis of intermediate b9

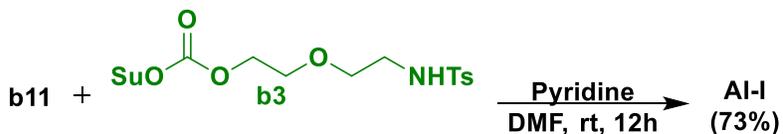
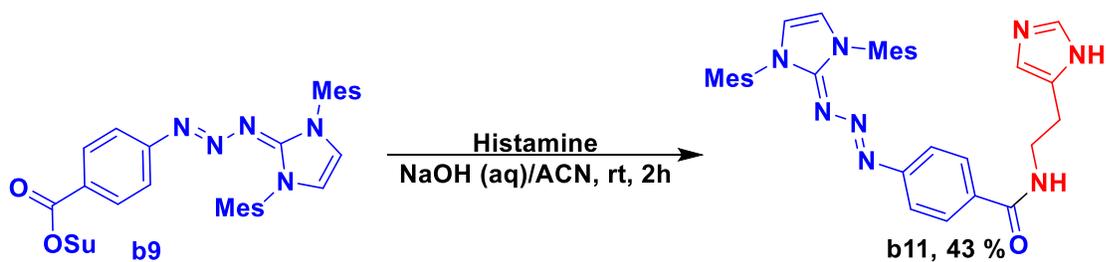
0.17 g of intermediate b8 was dissolved in 7 mL dry DMF. To this solution 0.14 g of EDC (0.73 mmol) and 83.7 mg of NHS (0.73 mmol) was added and stirred for 12 h. Next, the reaction mixture was washed with ethyl acetate and brine for 3 times. The ethyl acetate solution was dried

over anhydrous sodium sulfate and evaporated to give the crude compound. The crude intermediate was purified by resolidification with Hex:DCM (95:5) to obtain the pure compound b9 as bright yellow powder (147.9 mg, 72%). ^1H NMR (600 MHz, Chloroform-d) δ 7.76 (d, $J = 8.6$ Hz, 2H), 7.00 (s, 4H), 6.66 (d, $J = 0.8$ Hz, 2H), 6.58 (d, $J = 8.6$ Hz, 2H), 2.87 (s, 4H), 2.36 (s, 7H), 2.14 (s, 12H). ^{13}C NMR (151 MHz, cdCl_3) δ 172.47, 169.48, 161.86, 139.07, 134.80, 133.69, 130.97, 129.42, 121.15, 117.54, 98.39, 98.38, 98.37, 36.85, 25.66, 21.09, 17.88.





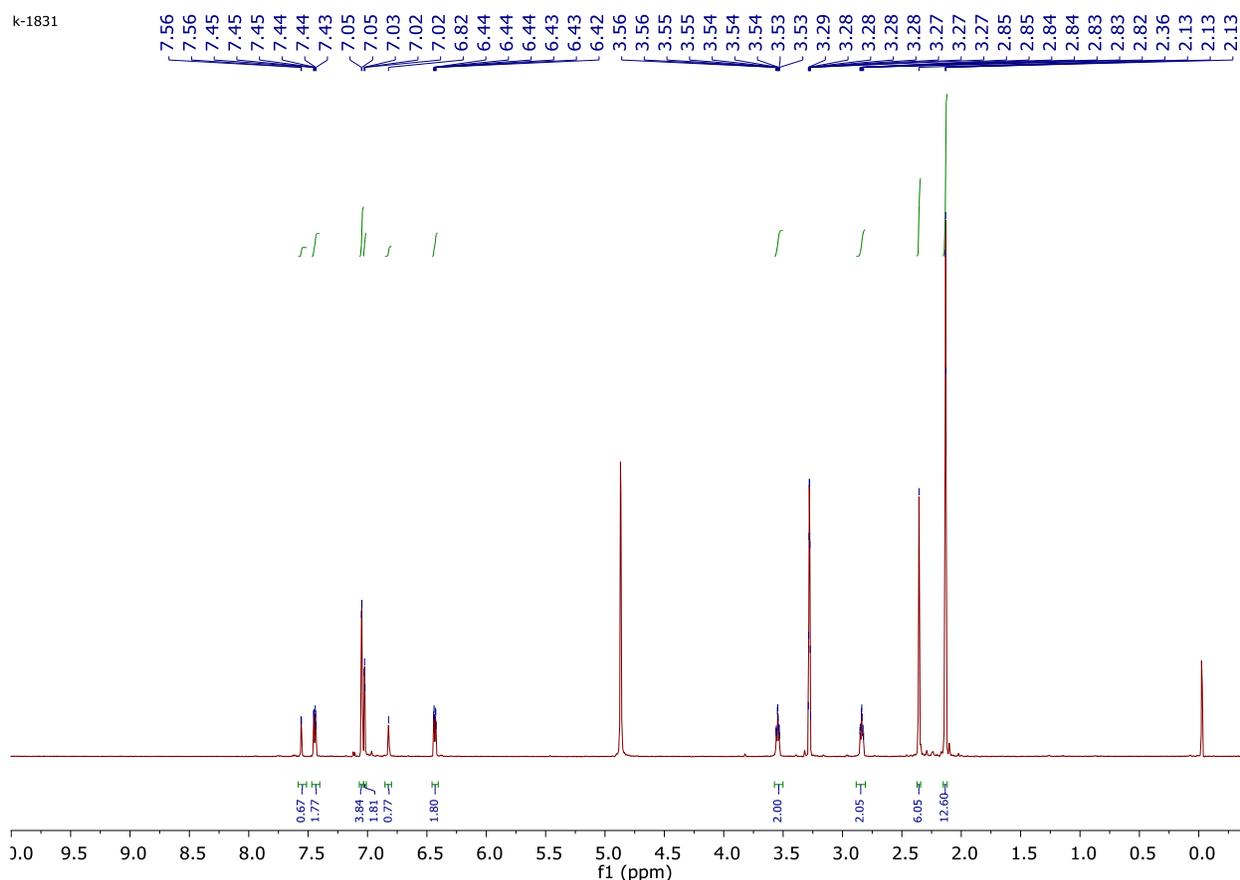
Synthesis of AI-I

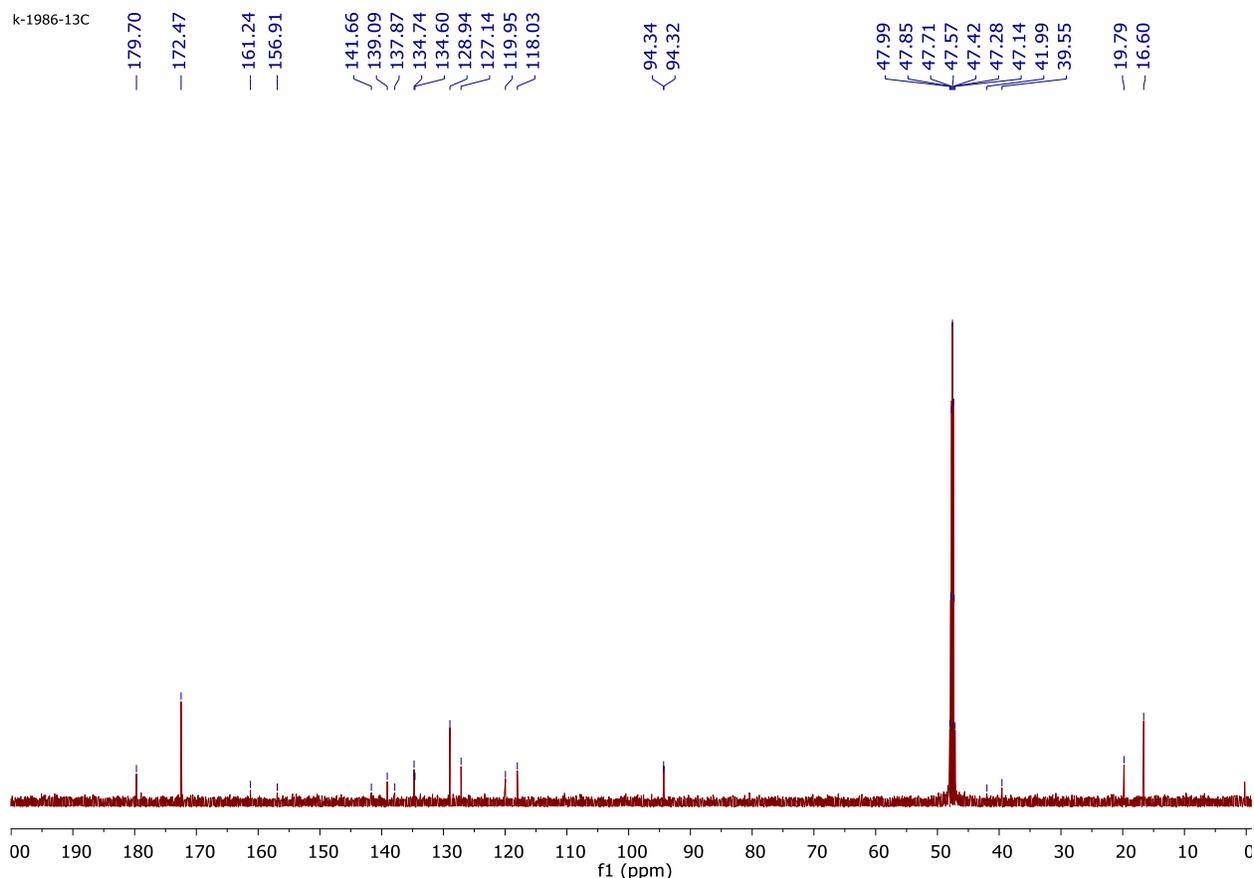


Synthesis of intermediate b11

To a solution of intermediate b9 (0.27 g, 0.48 mmol) in 2.5 mL of ACN, histamine dihydrochloride (83.2 mg, 0.45 mmol) in 1.2 mL of 1M NaOH(aq) was added dropwise at 0 °C. The reaction was allowed to stir for 2 h at room temperature. Next, the ACN was removed under

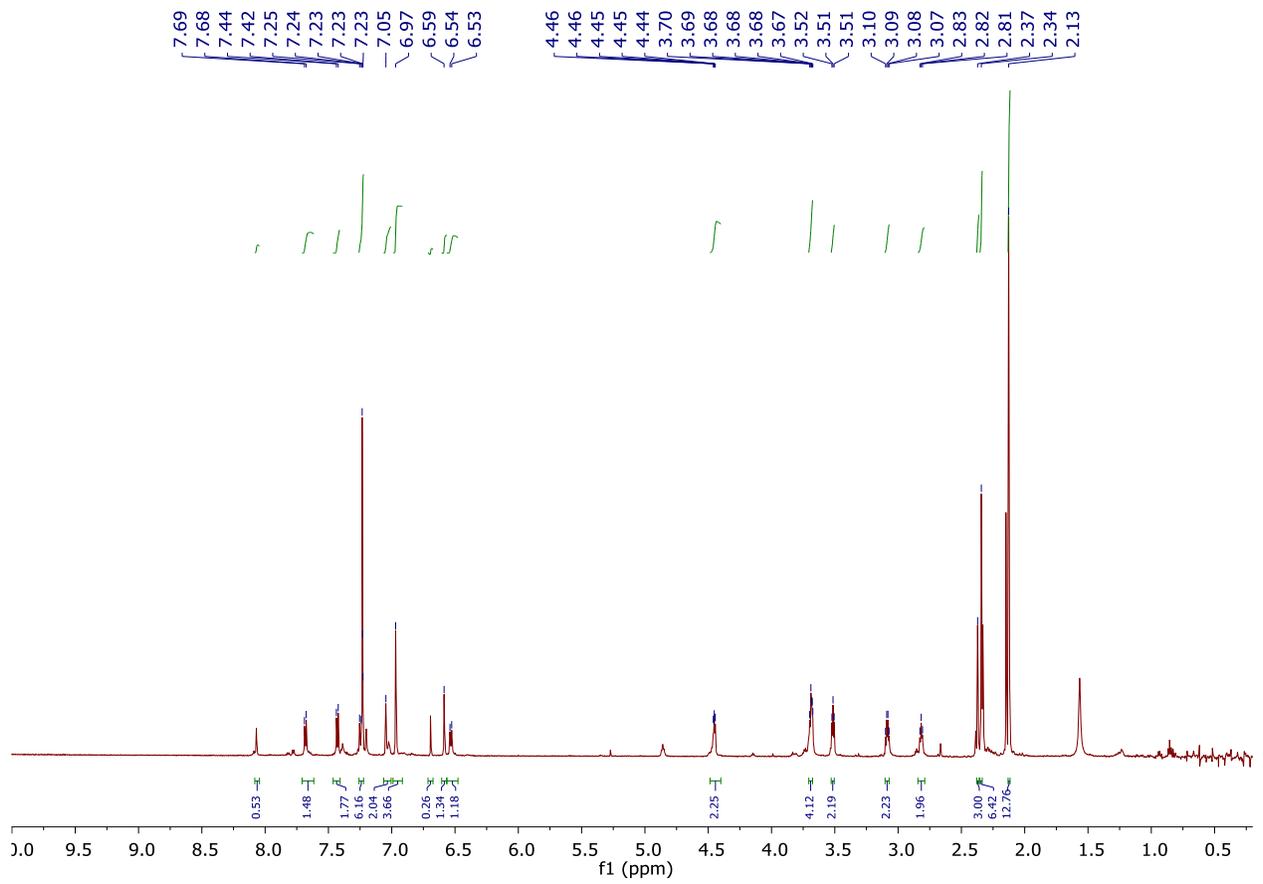
the reduced pressure, the mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product was purified by resolidification with Hex:DCM (95:5) to obtain the pure compound b11 as bright yellow powder (115.2 mg, 43%). ¹H NMR (600 MHz, Methanol-d₄) δ 7.56 (d, J = 2.4 Hz, 1H), 7.44 (dt, J = 8.6, 1.9 Hz, 2H), 7.05 (d, J = 2.6 Hz, 4H), 7.02 (t, J = 2.0 Hz, 2H), 6.82 (s, 1H), 6.43 (dt, J = 8.6, 2.0 Hz, 2H), 3.55 (tt, J = 7.4, 1.9 Hz, 2H), 2.88 – 2.81 (m, 2H), 2.36 (s, 6H), 2.13 (t, J = 1.9 Hz, 13H). ¹³C NMR (151 MHz, cd₃od) δ 179.70, 172.47, 161.24, 156.91, 141.66, 139.09, 137.87, 134.74, 134.60, 128.94, 127.14, 119.95, 118.03, 94.34, 94.32, 41.99, 39.55, 19.79, 16.60.

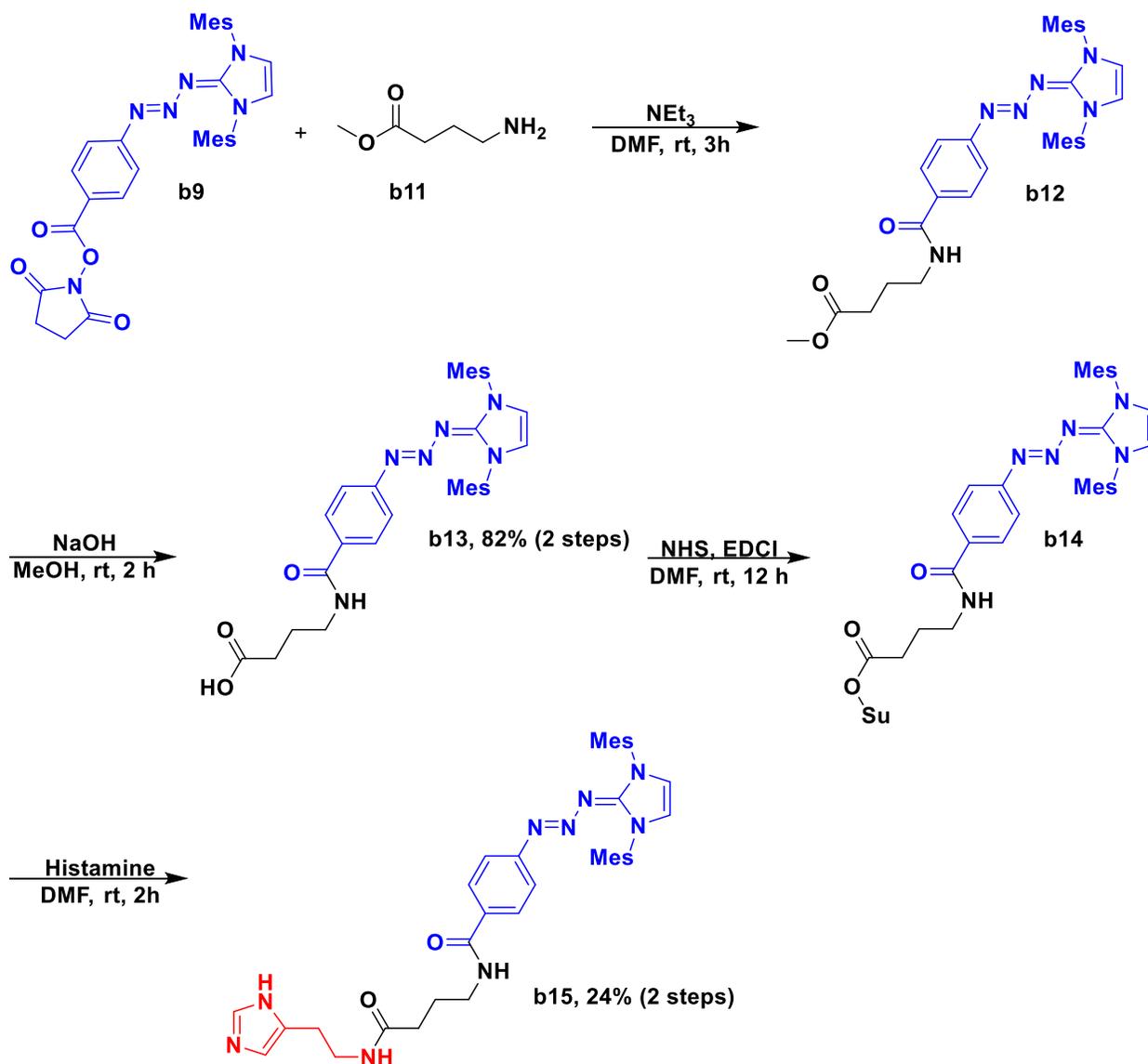




Synthesis of probe AI-I

A solution of b11 (50 mg, 89 μmol), b3 (139 mg, 134 μmol), and pyridine (65 μL , 267 μmol) in dry DMF (5 mL) was stirred for 12 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times. The organic layer was drying over Na_2SO_4 . After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO_2 (DCM : MeOH = 50 : 1 \rightarrow 20 : 1) to produce the probe AI-I as a bright yellow oil (55 mg, 73%). ^1H NMR (600 MHz, Chloroform- d) δ 8.07 (d, J = 1.3 Hz, 1H), 7.68 (d, J = 8.3 Hz, 1H), 7.43 (d, J = 8.7 Hz, 2H), 7.26 – 7.22 (m, 6H), 7.05 (s, 2H), 6.97 (s, 4H), 6.59 (s, 1H), 6.53 (d, J = 8.1 Hz, 1H), 4.49 – 4.40 (m, 2H), 3.71 – 3.68 (m, 4H), 3.51 (t, J = 5.2 Hz, 2H), 3.09 (q, J = 5.4 Hz, 2H), 2.82 (t, J = 6.2 Hz, 2H), 2.37 (s, 3H), 2.34 (s, 6H), 2.13 (s, 13H).





Synthesis of compound b12

To a solution of intermediate b9 (0.29 g, 0.53 mmol) in 9 mL of dry DMF, compound b11 (121 mg, 0.79 mmol), and triethyl amine (110 μ L, 0.79 mmol) were added at room temperature. The reaction was allowed to stir for 3 h at room temperature. Next, the mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product b12 was applied to the next step directly without further purification.

Synthesis of compound b13

To a methanolic solution of intermediate b14 (0.53 mmol in 5 mL methanol), 5 mL of 2.5 M NaOH solution was added in a drop-wise manner and the reaction was stirred at 25 °C for 12 h. Next, few drops of glacial acid were added to neutralize the reaction mixture. The mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product was purified by resolidification with Hex:DCM (95:5) to obtain the pure compound b13 as bright yellow powder (239.9 mg, 82% for 2 steps).

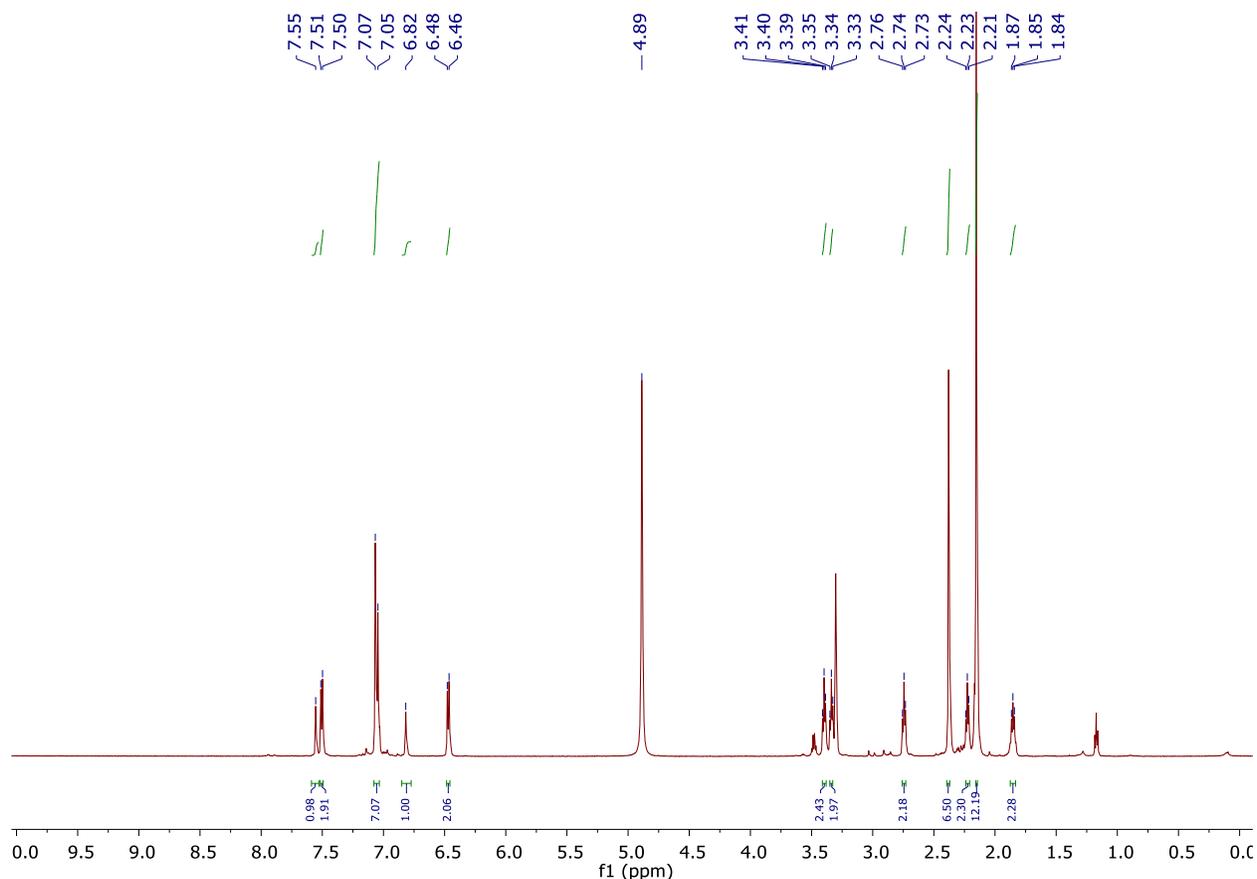
Synthesis of compound b14

0.24 g of intermediate b13 (0.43 mmol) was dissolved in 12 mL dry DMF. To this solution 0.16 g of EDC (0.87 mmol) and 100 mg of NHS (0.87 mmol) was added and stirred for 12 h. Next, the reaction mixture was washed with ethyl acetate and brine for 3 times. The ethyl acetate solution was dried over anhydrous sodium sulfate and evaporated to give the crude compound. The crude product b14 was applied to the next step directly without further purification.

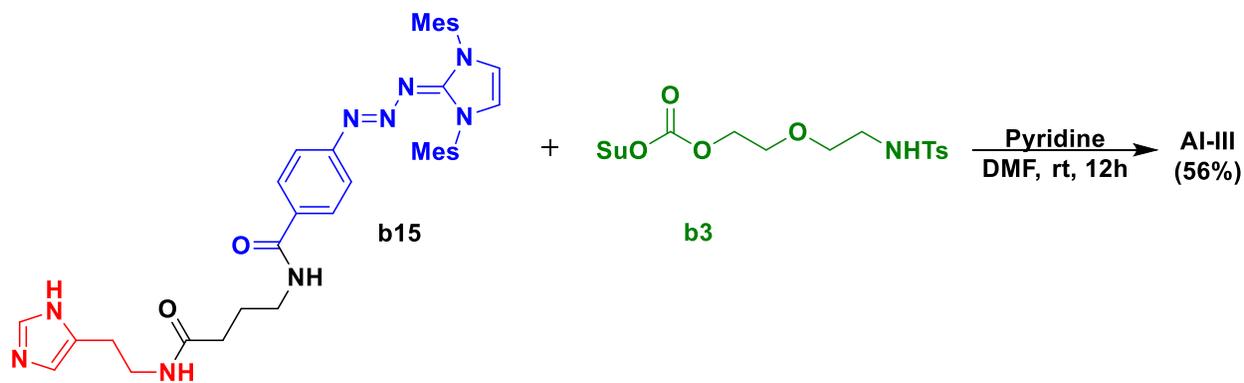
Synthesis of intermediate b15

To a solution of intermediate b14 (0.43 mmol) in 10 mL of dry DMF, histamine (241 mg, 2.17 mmol) was added at room temperature. The reaction was allowed to stir for 2 h at room temperature. Next, the mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product was purified by resolidification in water to obtain the pure compound b15 as bright yellow powder (66.5 mg, 24% for 2 steps). ¹H NMR (600 MHz, Methanol-d₄) δ 7.55 (s, 1H),

7.51 (d, J = 8.5 Hz, 2H), 7.06 (d, J = 12.9 Hz, 7H), 6.82 (s, 1H), 6.47 (d, J = 8.7 Hz, 2H), 3.40 (t, J = 7.2 Hz, 2H), 3.34 (d, J = 6.9 Hz, 2H), 2.74 (t, J = 7.2 Hz, 2H), 2.38 (s, 6H), 2.22 (d, J = 7.4 Hz, 2H), 2.15 (s, 12H), 1.88 – 1.83 (m, 2H).



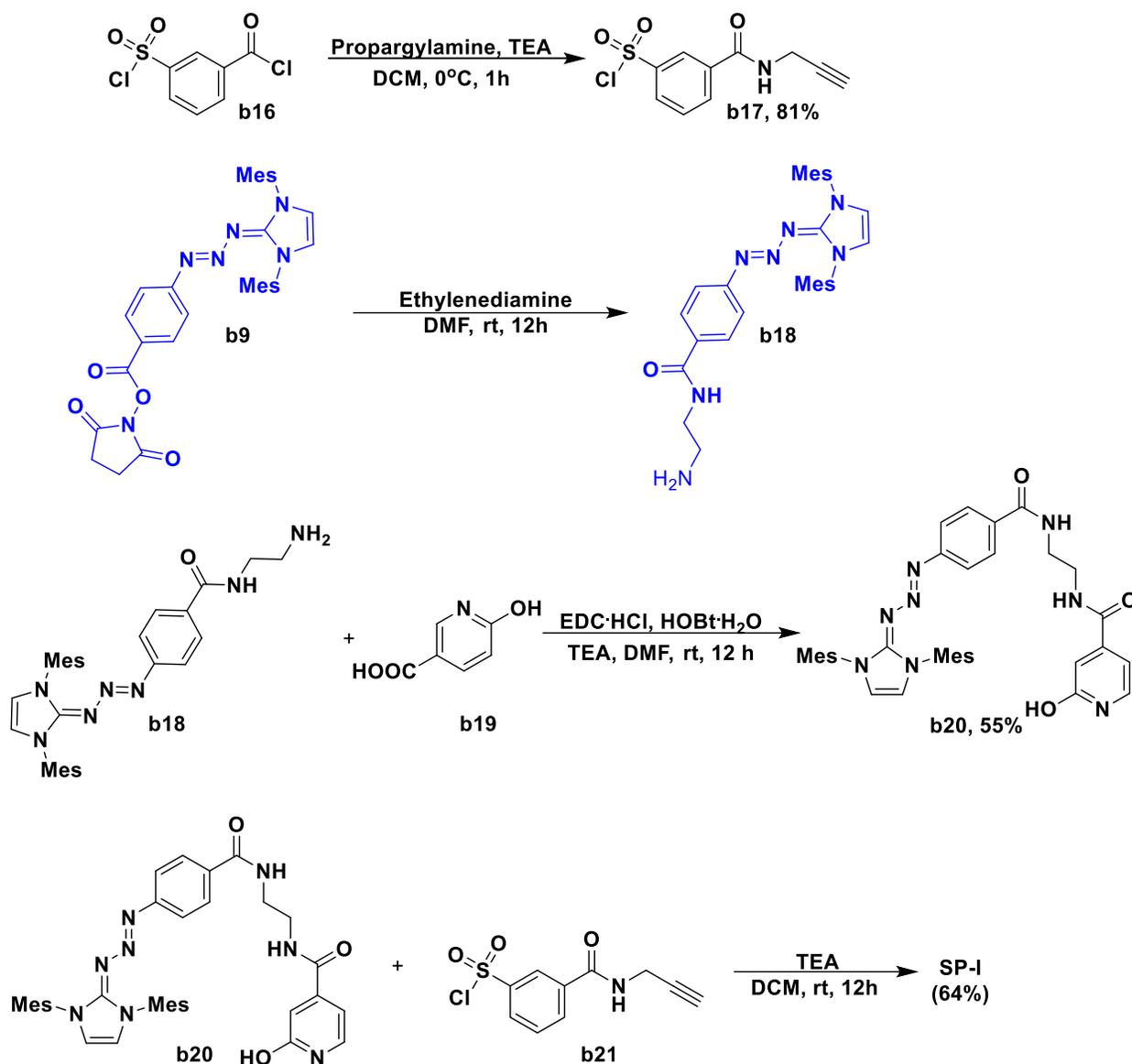
Synthesis of AI-III



Synthesis of probe AI-III

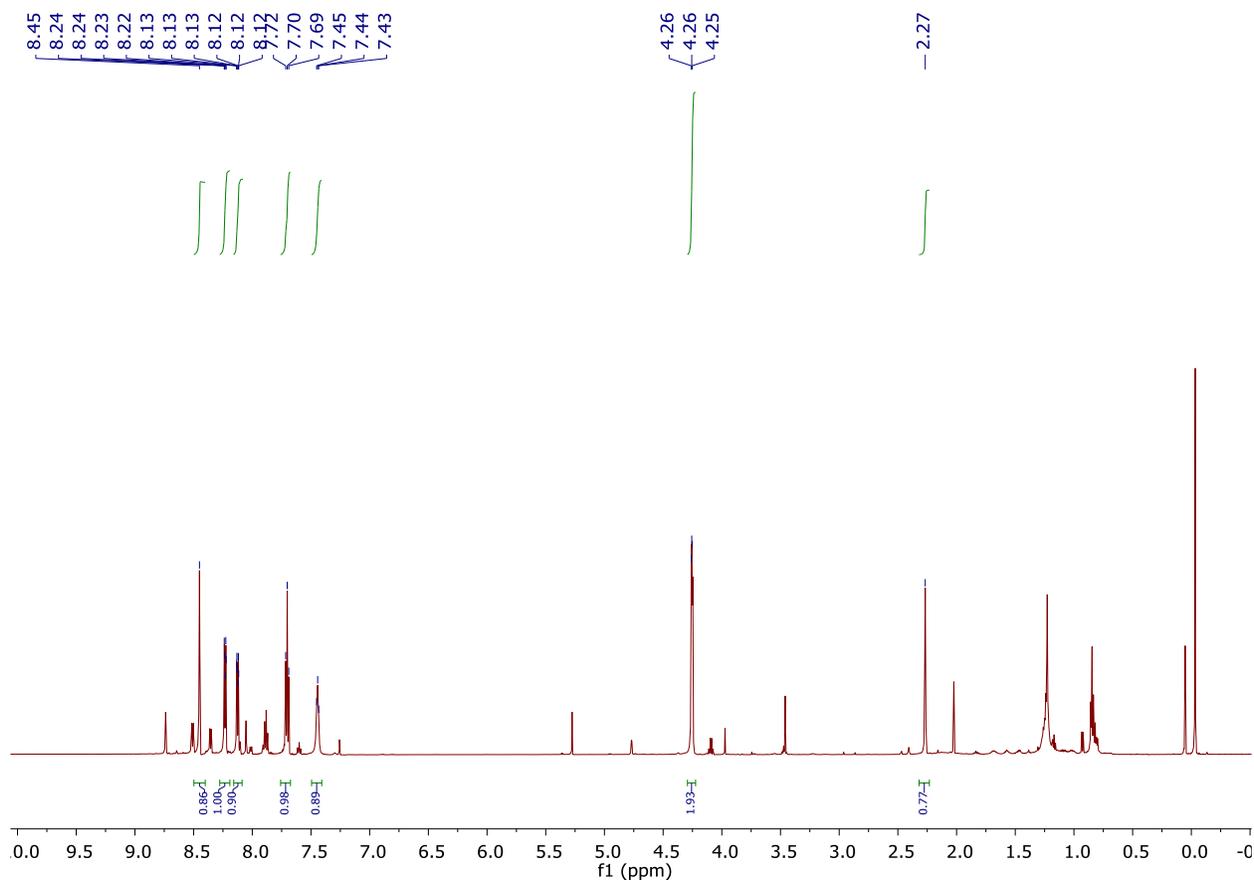
A solution of b15 (50 mg, 104 μmol), b3 (139 mg, 156 μmol), and pyridine (60 μL , 312 μmol) in dry DMF (6 mL) was stirred for 12 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times . The organic layer was drying over Na_2SO_4 . After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO_2 (DCM : MeOH = 50 : 1 \rightarrow 20 : 1) to produce the probe AI-I as a bright yellow oil (54.1 mg, 56%).

Synthesis of SP-I



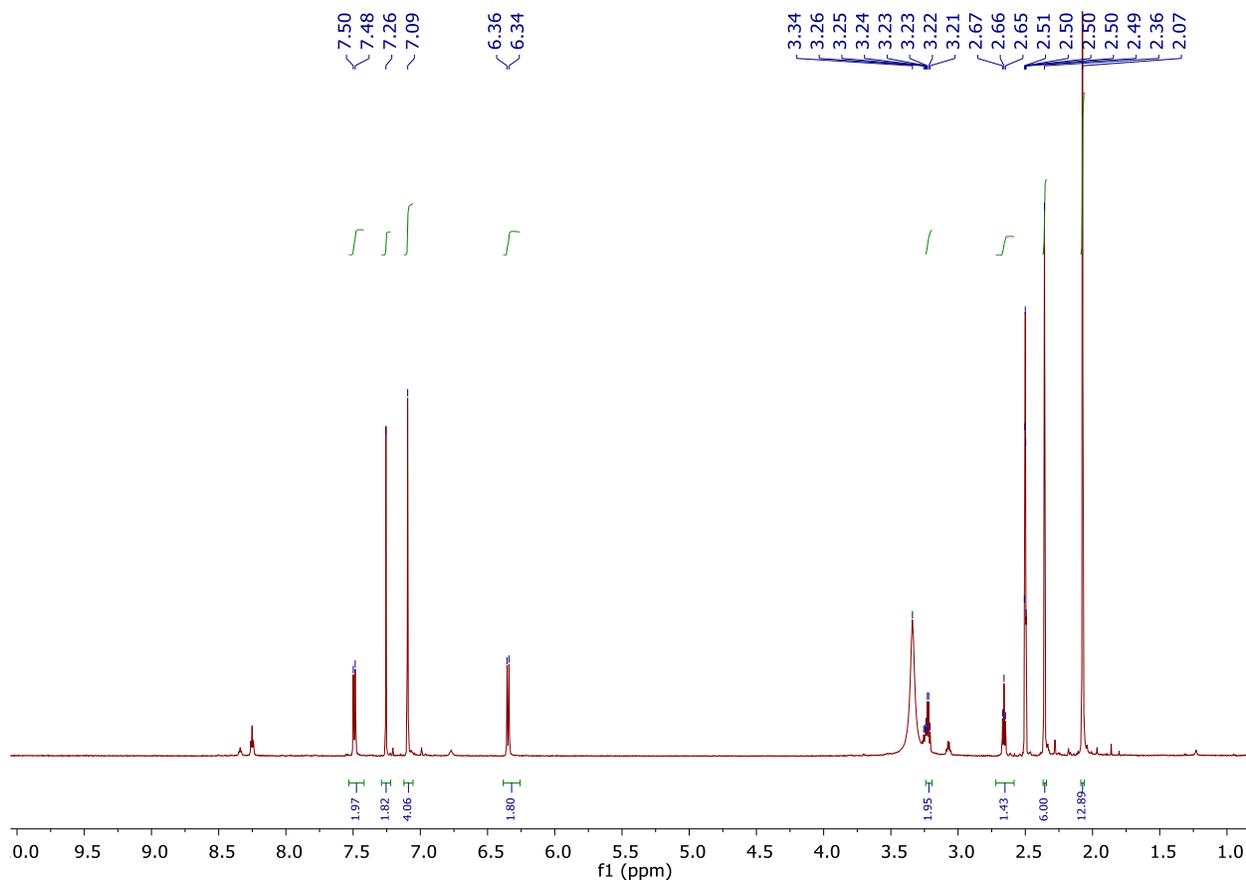
Synthesis of intermediate b17

To the solution of **b16** (chlorosulfonyl)benzoyl chloride (287 mg, 1.2mmol) and TEA (280 μ L, 2.0 mmol) in DCM (5mL) was added dropwise the solution of propargylamine (64 μ L, 1.0mmol) in DCM (10 mL). The reaction mixture was stirred for 1 h at 0°C. After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO₂ (DCM : AcOEt = 9 : 1) to give **b17** as a colorless oil (231 mg, 96 %). ¹H NMR (600 MHz, Chloroform-d) δ 8.45 (s, 1H), 8.23 (dd, J = 7.8, 1.4 Hz, 1H), 8.16 – 8.09 (m, 1H), 7.70 (t, J = 7.9 Hz, 1H), 7.44 (t, J = 5.3 Hz, 1H), 4.29 – 4.22 (m, 2H), 2.27 (s, 1H).



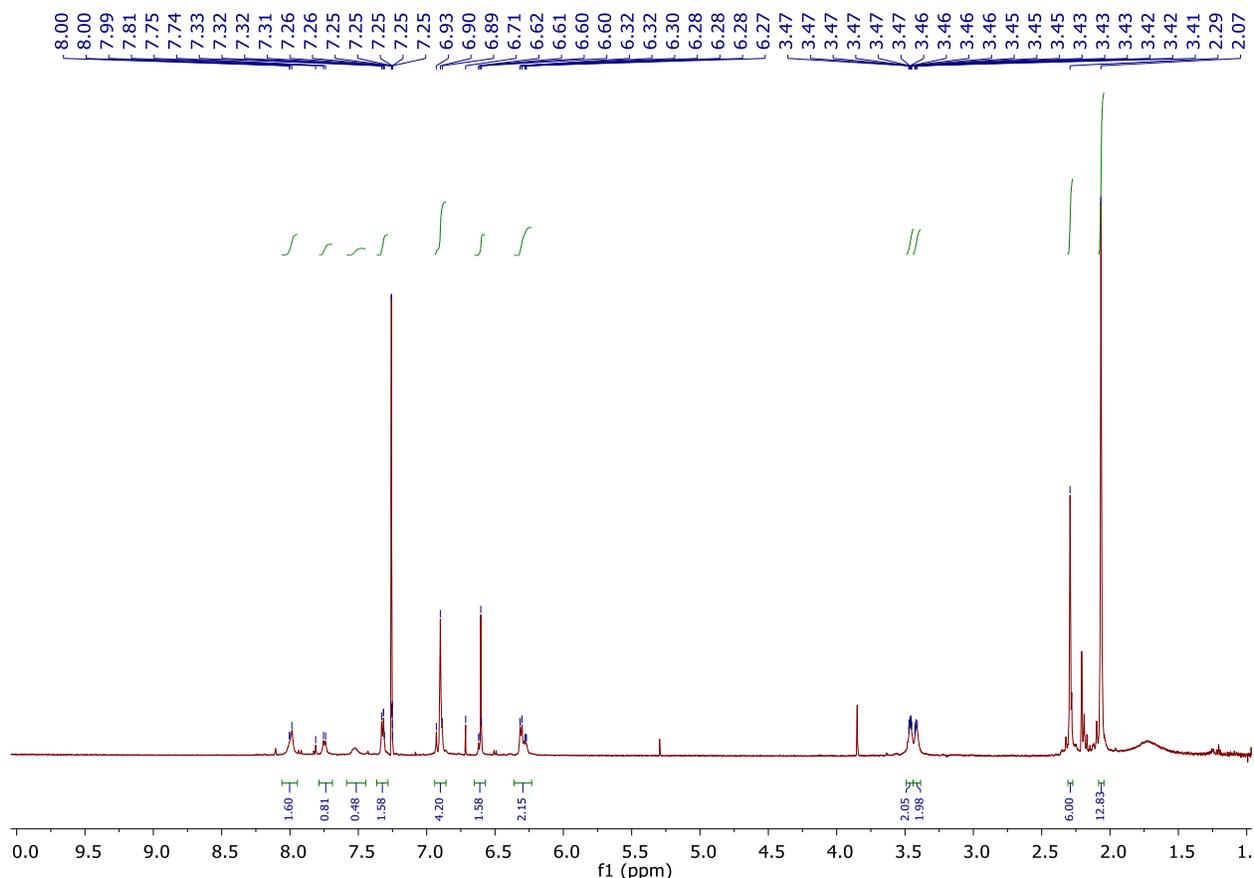
Synthesis of intermediate b18

To a solution of intermediate b9 (0.088 mmol) in 5 mL of dry DMF, ethylenediamine (245 μ L, 3.67 mmol) was added at room temperature. The reaction was allowed to stir for 2 h at room temperature. Next, the mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product was purified by resolidification in DCM : Hex = (95:5) to obtain the pure compound b18 as bright yellow powder (38.0 mg, 85%). ^1H NMR (600 MHz, DMSO- d_6) δ 7.49 (d, J = 8.7 Hz, 2H), 7.26 (s, 2H), 7.09 (s, 4H), 6.35 (d, J = 8.6 Hz, 2H), 3.22 (q, J = 6.3 Hz, 2H), 2.66 (t, J = 6.5 Hz, 1H), 2.36 (s, 6H), 2.07 (s, 13H).



Synthesis of intermediate b20

The solution of 6-Hydroxynicotinic acid (14.6 mg, 0.106 mmol), b18 (79 mg, 0.155 mmol), HOBt•H₂O (20 mg, 0.13 mmol), EDCI•HCl (25 mg, 0.13 mmol) and TEA (36.5 μL, 0.26 mmol) in DMF (0.5 mL) was stirred for 12 h at rt. After removal of the solvent by evaporation, the residue was purified by resolidification in Hex : DCM (95:5) to obtain the pure compound b20 as bright yellow powder (36.7 mg, 55%). ¹H NMR (600 MHz, Chloroform-d) δ 7.99 (s, 2H), 7.75 (d, J = 9.2 Hz, 1H), 7.32 (d, J = 8.1 Hz, 2H), 6.91 (d, J = 17.5 Hz, 4H), 6.60 (d, J = 2.9 Hz, 2H), 6.36 – 6.23 (m, 2H), 3.49 – 3.44 (m, 2H), 3.44 – 3.39 (m, 2H), 2.29 (s, 6H), 2.07 (s, 13H).

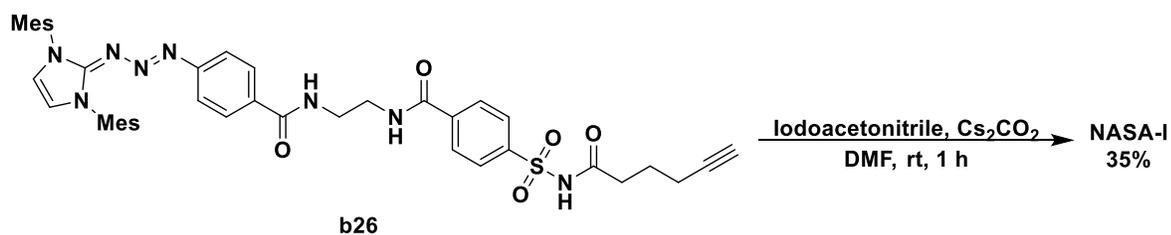
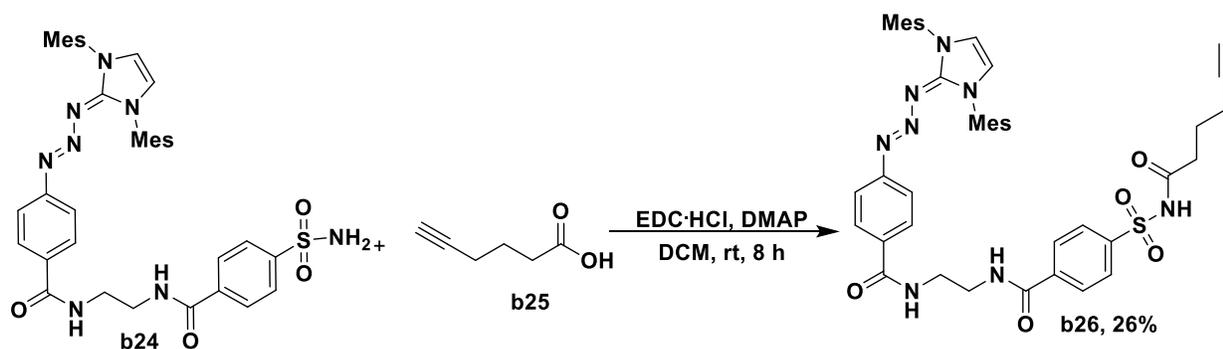
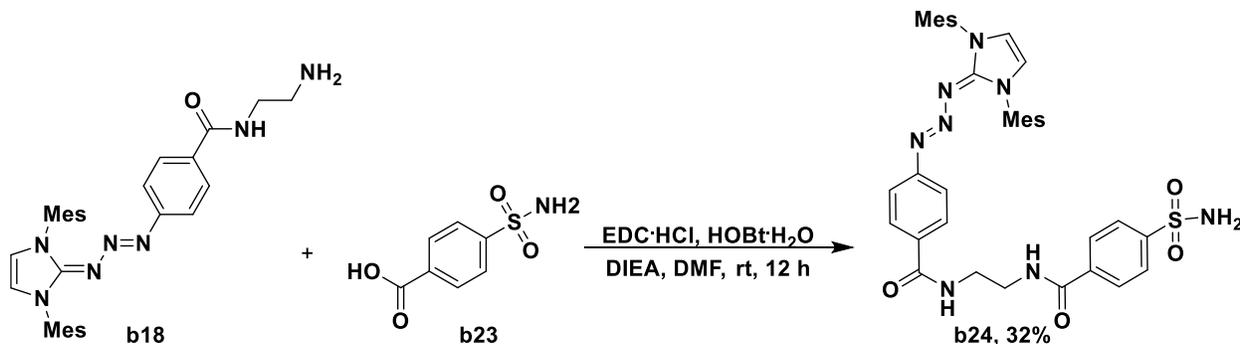


Synthesis of probe SP-I

To the solution of b20 (25 mg, 39 μ mol) and TEA (5.5 mL, 39 mmol) in DCM (10.4 mL), b17 (101 mg, 396 μ mol) was added. The reaction mixture was stirred for 12 h at rt. After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO₂ (DCM : MeOH = 50 : 1 \rightarrow 20 : 1) to produce the probe SP-I as a bright yellow solid (31.2 mg, 64%). ¹H NMR (600 MHz, Chloroform-d) δ 8.38 (d, J = 38.9 Hz, 1H), 8.23 (d, J = 8.6 Hz, 1H), 8.02 (dd, J = 24.4, 8.1 Hz, 1H), 7.40 – 7.29 (m, 2H), 7.27 (s, 1H), 7.01 – 6.90 (m, 5H), 6.78 (s, 1H), 6.29 (d, J = 8.4 Hz, 1H), 4.14 (d, J = 5.6 Hz, 1H), 3.54 (s, 2H), 3.47 (q, J = 6.5, 5.5 Hz, 4H), 2.31 (s, 6H), 2.07 (s, 12H).

(DCM : MeOH = 20 : 1 \rightarrow 15 : 1 \rightarrow 7 : 1) to produce the probe SP-II as a bright yellow solid (17.4 mg, 70%).

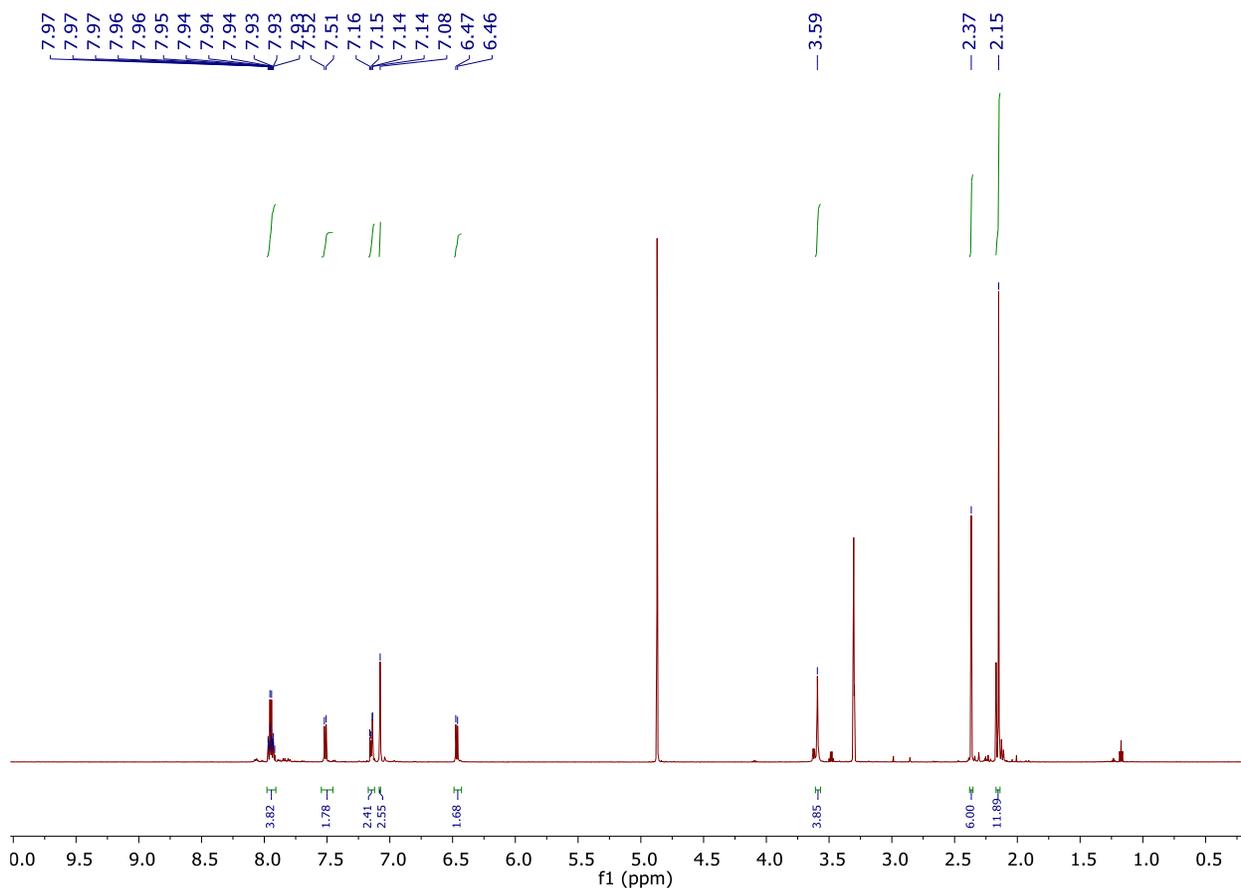
Synthesis of NASA-I



Synthesis of intermediate b24

To a stirred solution of 4-sulfamoylbenzoic acid **b23** (40 mg, 0.19 mmol) in dry DMF 1 mL was added **b18** (100 mg, 0.19 mmol), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) (56.2 mg, 0.29 mmol), HOBT·H₂O (20 mg, 0.13 mmol) (45 mg, 0.29 mmol) and N, N-diisopropylethylamine (DIEA) (0.1 mL, 0.59 mmol). The mixture was stirred overnight at room

temperature. The solution was dissolved in DCM and washed with Sat. NaHCO₃ for three times. The organic layer was dried over MgSO₄, filtered and evaporated to yield crude compound. The residue was purified by resolidification in ether : methanol (95:5) to obtain the pure compound b24 as bright yellow powder (43.4 mg, 32%). ¹H NMR (600 MHz, Methanol-d₄) δ 7.98 – 7.91 (m, 4H), 7.51 (d, J = 8.7 Hz, 2H), 7.17 – 7.12 (m, 2H), 7.08 (s, 3H), 6.47 (d, J = 8.7 Hz, 2H), 3.59 (s, 4H), 2.37 (s, 6H), 2.15 (s, 12H).



Synthesis of intermediate b26

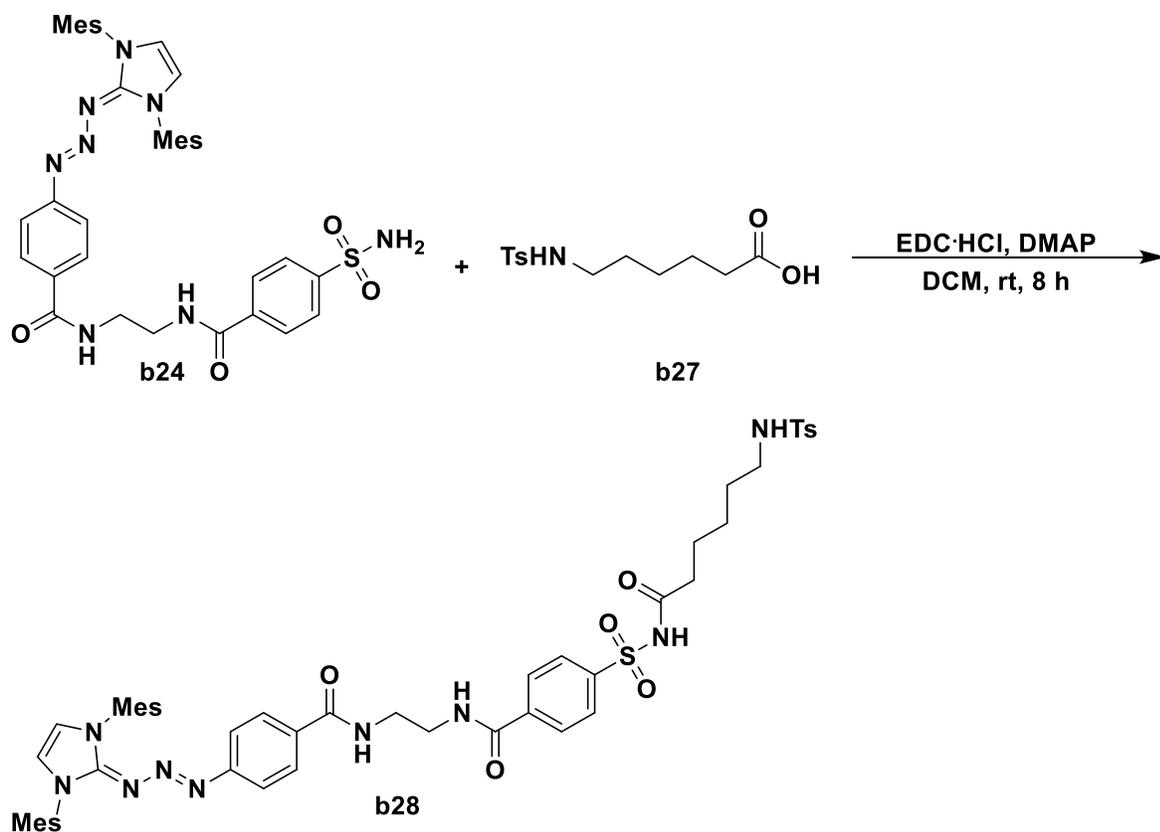
To a stirred solution of compound b24 (27 mg, 0.04 mmol) in dry DMF (0.5 mL) was added hexynoic acid b25 (7 μL, 0.06 mmol), EDC (11.2 mg, 0.058 mmol), 4-dimethylaminopyridine (DMAP) (1.5 mg, 0.011 mmol) and DIEA (34 μL, 0.2 mmol). The mixture was stirred 8 h at room temperature. The solvent was removed under reduced pressure, and the residue was

purified by resolidification in Hex : DCM (95:5) to obtain the pure compound b256 as bright yellow powder (8.1 mg, 32%).

Synthesis of NASA-I

To a stirred solution of compound b26 (8.3 mg, 10 μ mol) in dry DMF (0.1 mL) was added iodoacetonitrile (7.6 μ L, 0.105 mmol) and Cs₂CO₃ (5.1 mg, 15 μ mol). The mixture was stirred 1 h at room temperature. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography on silica gel (DCM : MeOH = 50:1→30:1→10:1) to yield NASA-I as brown oil (2.9 mg, 35%).

Synthesis of NASA-II



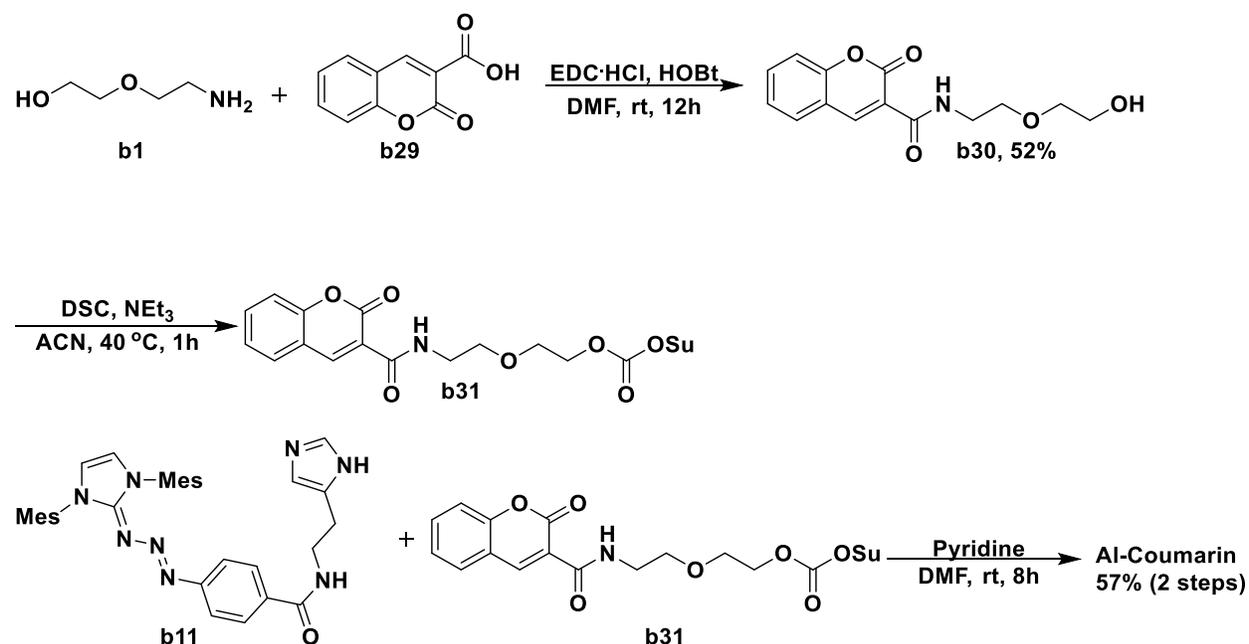
Synthesis of intermediate **b28**

To a stirred solution of compound **b24** (40 mg, 0.057 mmol) in dry DMF (0.7 mL) was added **b27** (24.3 mg, 0.086 mmol), EDC (17 mg, 0.086 mmol), 4-dimethylaminopyridine (DMAP) (2.5 mg, 0.017 mmol) and DIEA (50 μL , 0.285 mmol). The mixture was stirred 12 h at room temperature. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography on silica gel (DCM : MeOH = 10:1) to yield **b28** as yellow oil (13.6 mg, 25%).

Synthesis of **NASA-II**

To a stirred solution of compound b28 (10 mg, 10 μ mol) in dry DMF (0.1 mL) was added iodoacetonitrile (2.2 μ L, 0.03 mmol) and Cs₂CO₃ (5.1 mg, 15 μ mol). The mixture was stirred 1 h at room temperature. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography on silica gel (DCM : MeOH = 50:1→30:1→10:1) to yield NASA-II as brown oil (4.6 mg, 46%).

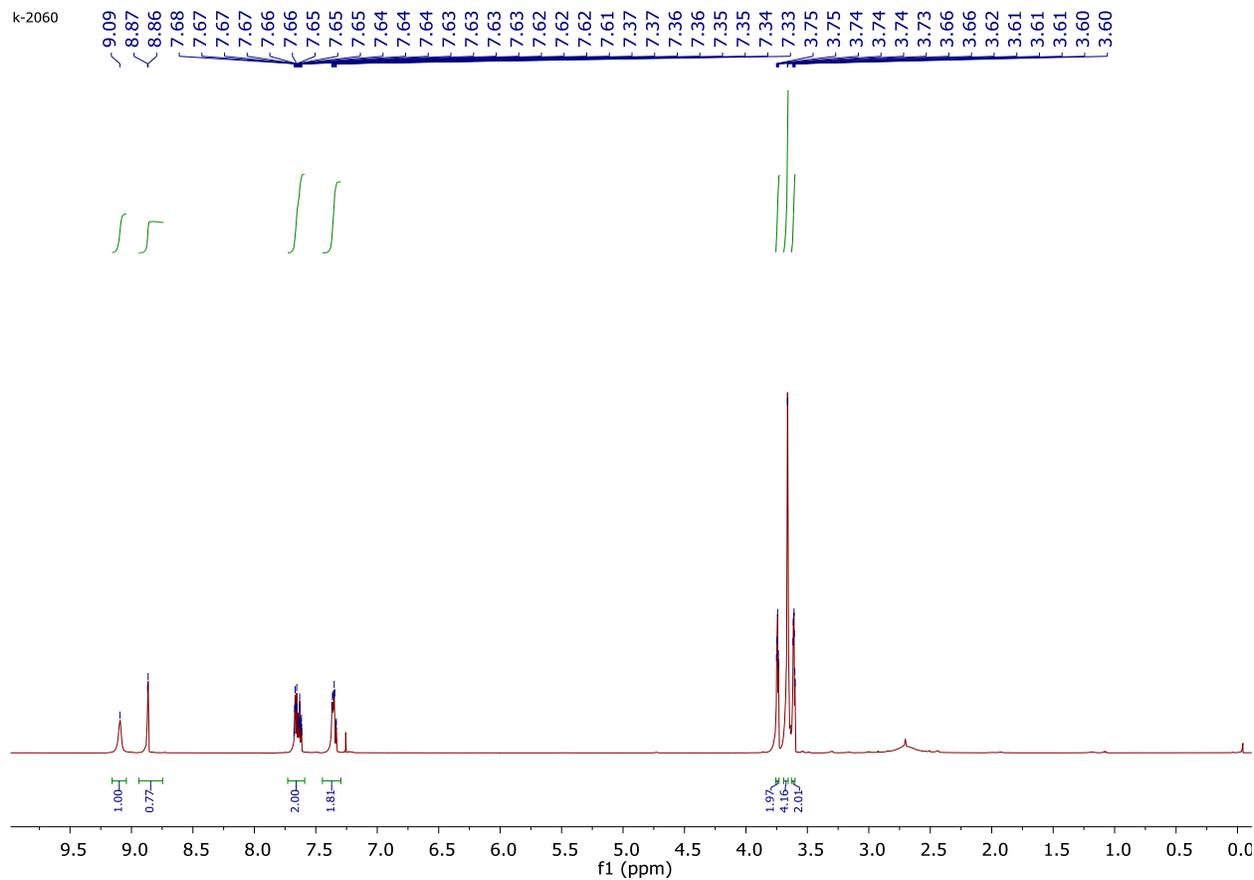
Synthesis of Al-Coumarin



Synthesis of intermediate b30

To a solution of 2-Oxo-2H-chromene-6-carboxylic acid (1.4 g, 7.65 mmol) in anhydrous DMF (15 mL) was added 1-hydroxybenzotriazole monohydrate (HOBT) (1.5 g, 9.79 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) (1.94 g, 10.12 mmol). After stirring at room temperature for 10 min, 2-(2-aminoethoxy)ethanol (990 μ L, 9.98 mmol) was added to the above solution. The mixture was allowed to stir at 25 °C for 12h. After evaporation, the crude was dissolved in DCM and washed three times with saturated aqueous NaHCO₃. The

organic layer was dried over anhydrous Na_2SO_4 , filtered, and concentrated. The resulting residue was recrystallized with Et_2O to yield compound b30 as yellow solid (1.52 g, 53%). ^1H NMR (600 MHz, Chloroform- d) δ 9.09 (s, 1H), 8.87 (d, $J = 2.9$ Hz, 1H), 7.73 – 7.59 (m, 2H), 7.35 (ddd, $J = 10.4, 7.3, 2.4$ Hz, 2H), 3.76 – 3.73 (m, 2H), 3.66 (d, $J = 1.8$ Hz, 4H), 3.63 – 3.60 (m, 2H).



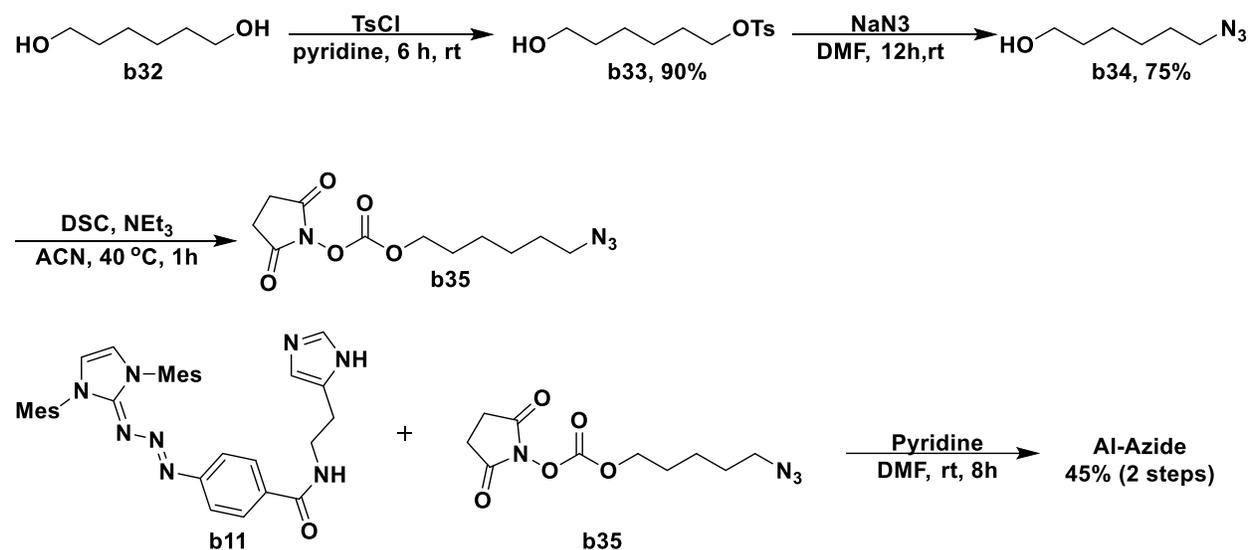
Synthesis of intermediate b31

Compound b30 (500 mg, 1.8 mmol) was dissolved in 13 mL dry ACN. To the solution, triethylamine (0.5 mL, 3.6 mmol) and disuccinimidyl carbonate (DSC, 924.2 mg, 3.6 mmol) was added. The reaction was stirred at 40 °C. After 1 hour the reaction mixture was concentrated under reduced pressure and purified by flash column chromatography (DCM : MeOH = 50 : 1) to obtain compound b31 as yellow oil. The intermediate b31 was applied to the next step directly without purification.

Synthesis of probe Al-Coumarin

A solution of b11 (50 mg, 89 μ mol), b31 (56 mg, 134 μ mol), and pyridine (50 μ L, 267 μ mol) in dry DMF (5 mL) was stirred for 12 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times. The organic layer was drying over Na₂SO₄. After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO₂ (DCM : MeOH = 50 : 1 \rightarrow 20 : 1) to produce the probe Al-I as a bright yellow oil (43.9 mg, 52% for 2 steps).

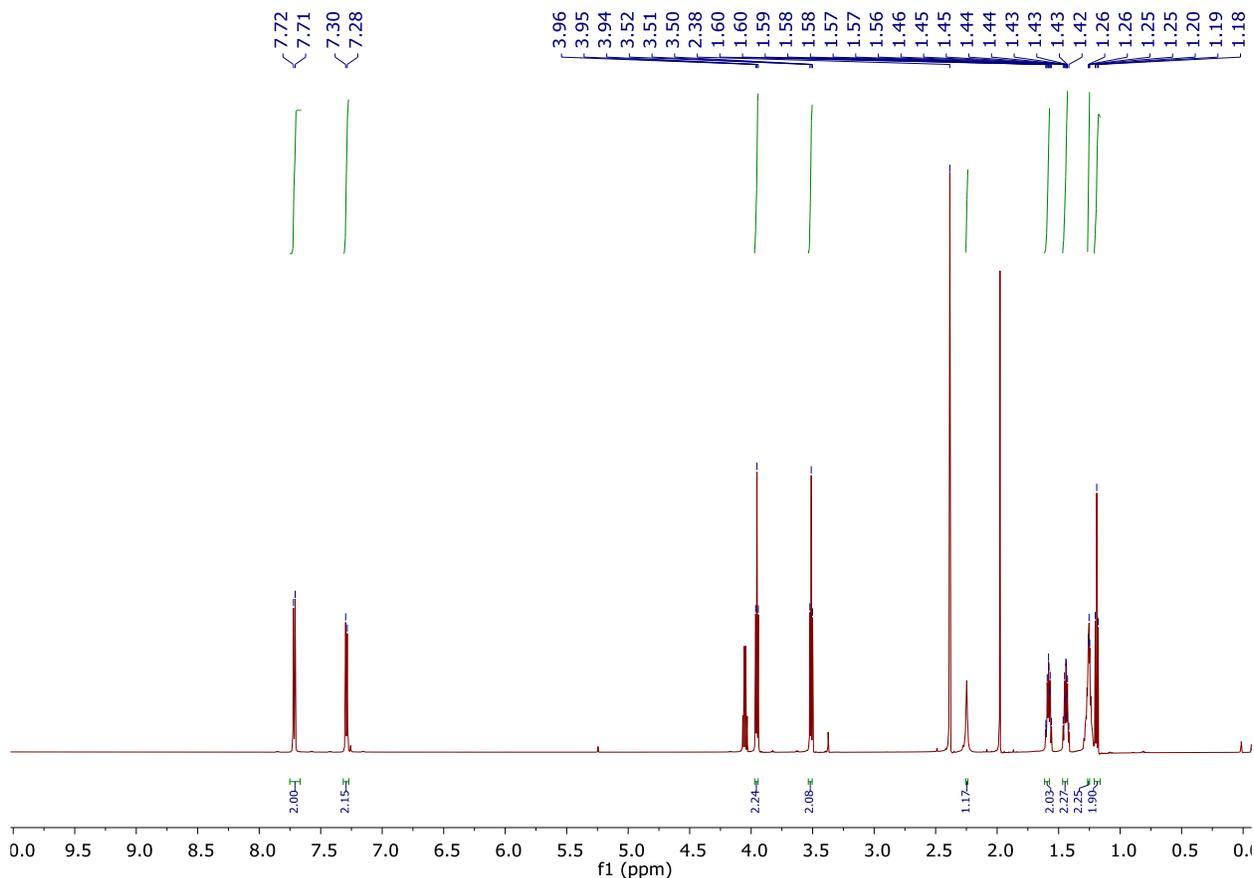
Synthesis of Al-Azide



Synthesis of intermediate b33

To a stirring solution of 1,6-hexanediol (4.6 g, 39.1 mmol) in pyridine (5 mL) was added Ts-Cl (4.03 g, 19.5 mmol) and the reaction mixture was stirred 6 h at room temperature. Next, the reaction was quenched by the addition of an excess of 1 M HCl. The organic layer was isolated and washed sequentially with saturated aqueous NaHCO₃, then brine. The organic layer was then dried over Na₂SO₄ and concentrated by rotary evaporation. The crude compound was purified by flash silica gel chromatography (EtOAc/hexanes = 1 : 2) to afford b33 as a

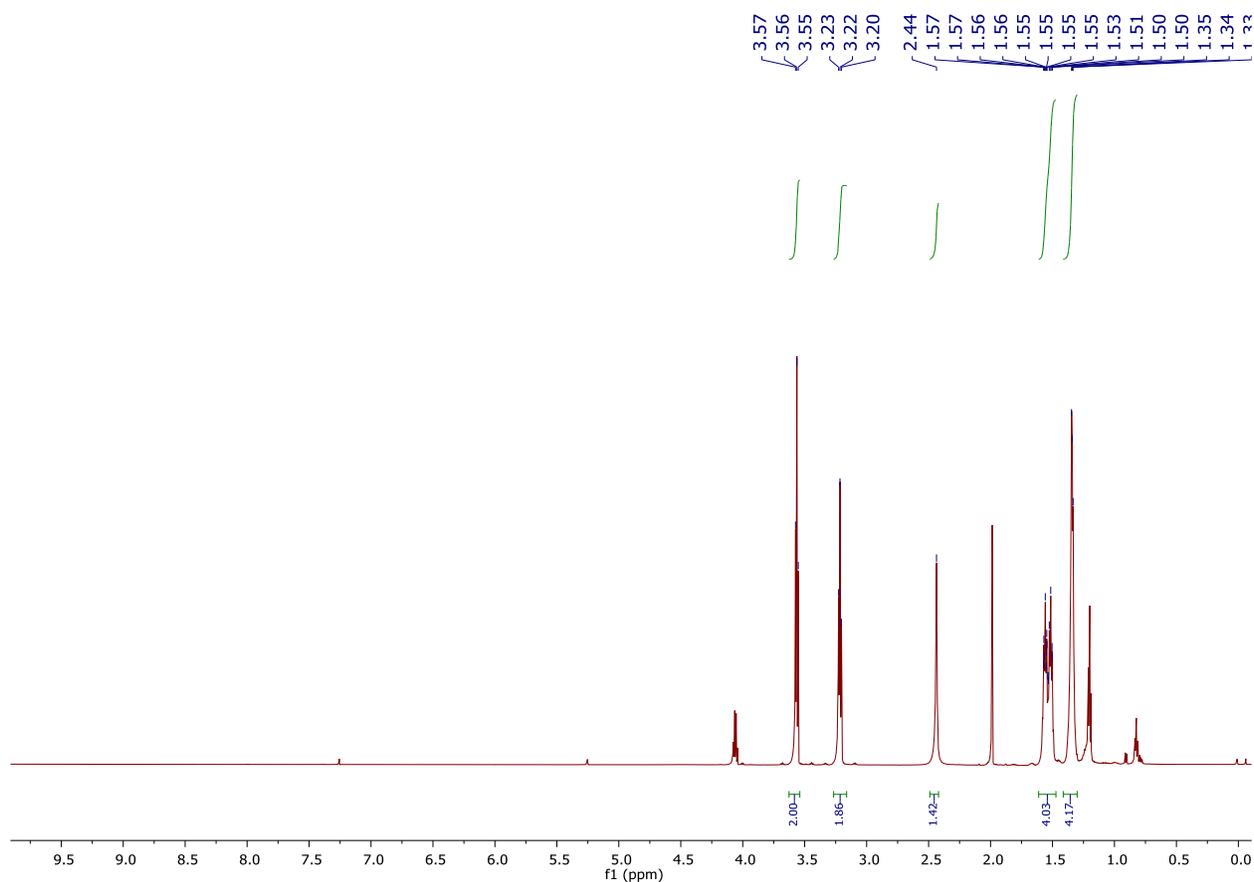
colorless oil in 90% yield (9.5 g, 90%). ^1H NMR (600 MHz, Chloroform- d) δ 7.71 (d, J = 8.3 Hz, 2H), 7.29 (d, J = 7.9 Hz, 2H), 3.96 (d, J = 6.5 Hz, 2H), 3.52 (d, J = 6.6 Hz, 2H), 2.38 (s, 3H), 2.25 (s, 1H), 1.61 – 1.57 (m, 2H), 1.47 – 1.43 (m, 2H), 1.26 – 1.25 (m, 2H), 1.19 (t, J = 7.2 Hz, 2H).



Synthesis of intermediate b34

To a stirring solution of the b33 (326 mg, 1.2 mmol) in DMF (5 mL) was added sodium azide (94 mg, 1.45 mmol) and the mixture was stirred at 25 °C for 12h. The reaction was diluted with water and extracted with ethyl acetate for three times. The combined organic extracts were washed with brine for three times. The organic layer was then dried over Na_2SO_4 and concentrated by rotary evaporation to afford the crude. The residue was purified by flash silica gel chromatography (EtOAc/hexanes = 1 : 10) to afford b34 as a yellow oil (128 mg, 75%). ^1H

NMR (600 MHz, Chloroform-d) δ 3.56 (t, $J = 6.6$ Hz, 2H), 3.22 (t, $J = 6.8$ Hz, 2H), 2.44 (s, 1H), 1.61 – 1.47 (m, 4H), 1.41 – 1.30 (m, 4H).



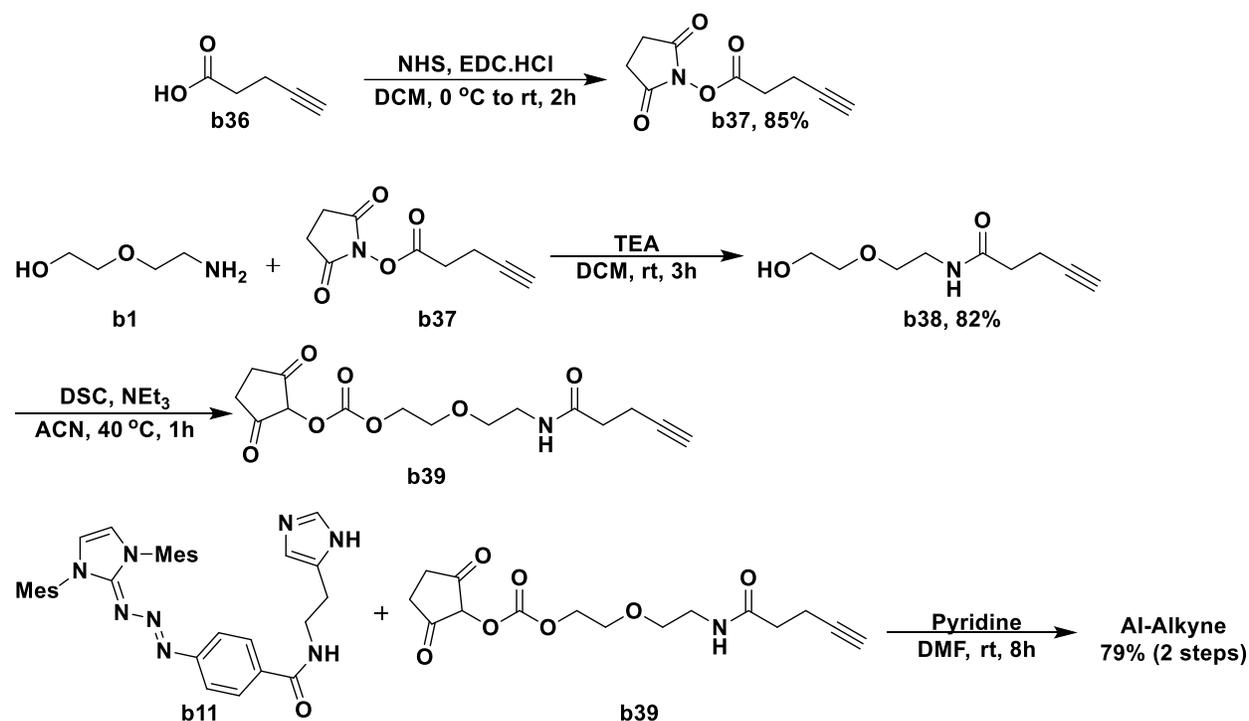
Synthesis of intermediate b35

Compound b34 (227 mg, 1.58 mmol) was dissolved in 12 mL dry ACN. To the solution, triethylamine (0.44 mL, 3.17 mmol) and disuccinimidyl carbonate (DSC, 813.3 mg, 3.17 mmol) was added. The reaction was stirred at 40 °C. After 1 hour the reaction mixture was concentrated under reduced pressure and purified by flash column chromatography (EtOAc/hexanes = 1 : 1) to obtain compound b35 as yellow oil. The intermediate b35 was applied to the next step directly without purification.

Synthesis of Al-Azide

A solution of b11 (44 mg, 78.5 μmol), b35 (33.4 mg, 117 μmol), and pyridine (44 μL , 235.5 μmol) in dry DMF (4.5 mL) was stirred for 8 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times. The organic layer was drying over Na_2SO_4 . After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO_2 ($\text{DCM} : \text{MeOH} = 50 : 1 \rightarrow 20 : 1$) to produce the probe AI-Azide as a bright yellow oil (25.7 mg, 45% for 2 steps).

Synthesis of AI-Alkyne



Synthesis of intermediate b37

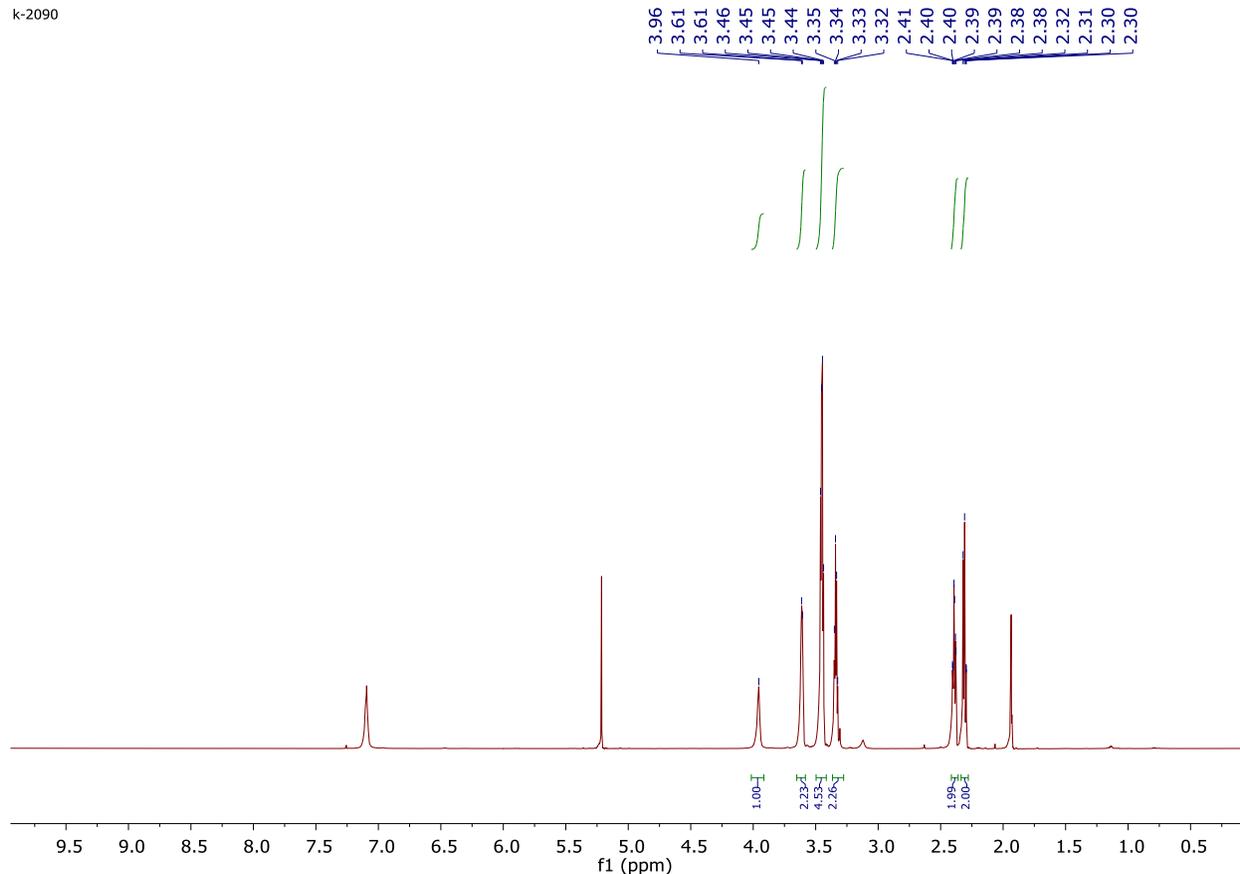
0.3 g of 4-pentynoic acid b36 was dissolved in 15 mL dry DCM. To this solution 1.17 g of EDC (6.1 mmol) and 402 mg of NHS (3.49 mmol) was added and stirred for 2 h. Next, the reaction mixture was washed with ethyl acetate and brine for 3 times. The ethyl acetate solution

was dried over anhydrous sodium sulfate and evaporated to yield compound b38 as colorless crystal (507.4 mg, 85%).

Synthesis of intermediate b38

To the solution of b37 (1.2 g, 6.15 mmol) and TEA (1.7 mL, 12.3 mmol) in DCM (20 mL), 2-(2-aminoethoxy)ethan-1-ol b1 (2.6 ml, 25.84 mmol) was added. The reaction mixture was stirred for 3 h at rt. Next, the reaction mixture was washed three times with saturated aqueous NaHCO₃. The organic layer was then dried over Na₂SO₄ and concentrated by rotary evaporation to yield compound b38 as colorless oil (933.9 mg, 82%). ¹H NMR (600 MHz, Chloroform-d) δ 3.96 (s, 1H), 3.61 (d, J = 4.1 Hz, 2H), 3.45 (q, J = 5.3, 4.9 Hz, 5H), 3.34 (q, J = 5.3 Hz, 2H), 2.39 (dt, J = 7.0, 3.5 Hz, 2H), 2.31 (t, J = 7.6 Hz, 2H).

k-2090



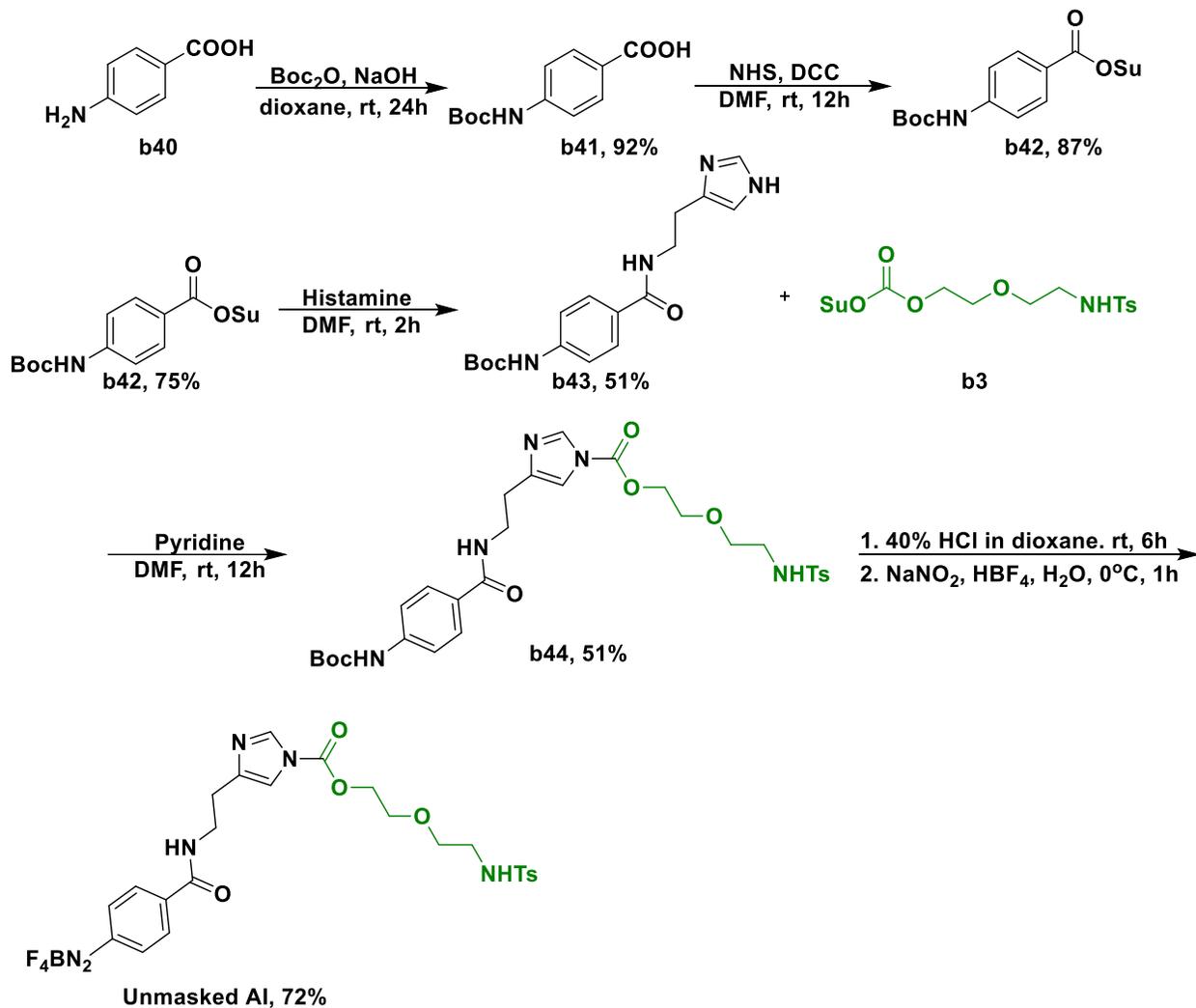
Synthesis of intermediate b39

Compound b38 (240 mg, 1.29 mmol) was dissolved in 9.4 mL dry ACN. To the solution, triethylamine (0.36 mL, 2.59 mmol) and disuccinimidyl carbonate (DSC, 664.6 mg, 2.59 mmol) was added. The reaction was stirred at 40 °C. After 1 hour the reaction mixture was concentrated under reduced pressure and purified by flash column chromatography (EtOAc/hexanes = 1 : 1) to obtain compound b39 as yellow oil. The intermediate b39 was applied to the next step directly without purification.

Synthesis of AI-Alkyne

A solution of b11 (62 mg, 110 μmol), b39 (54 mg, 166 μmol), and pyridine (54 μL, 332 μmol) in dry DMF (6 mL) was stirred for 12 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times. The organic layer was drying over Na₂SO₄. After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO₂ (DCM : MeOH = 50 : 1 → 20 : 1) to produce the probe **AI-Alkyne** as a bright yellow oil (67.0 mg, 79% for 2 steps).

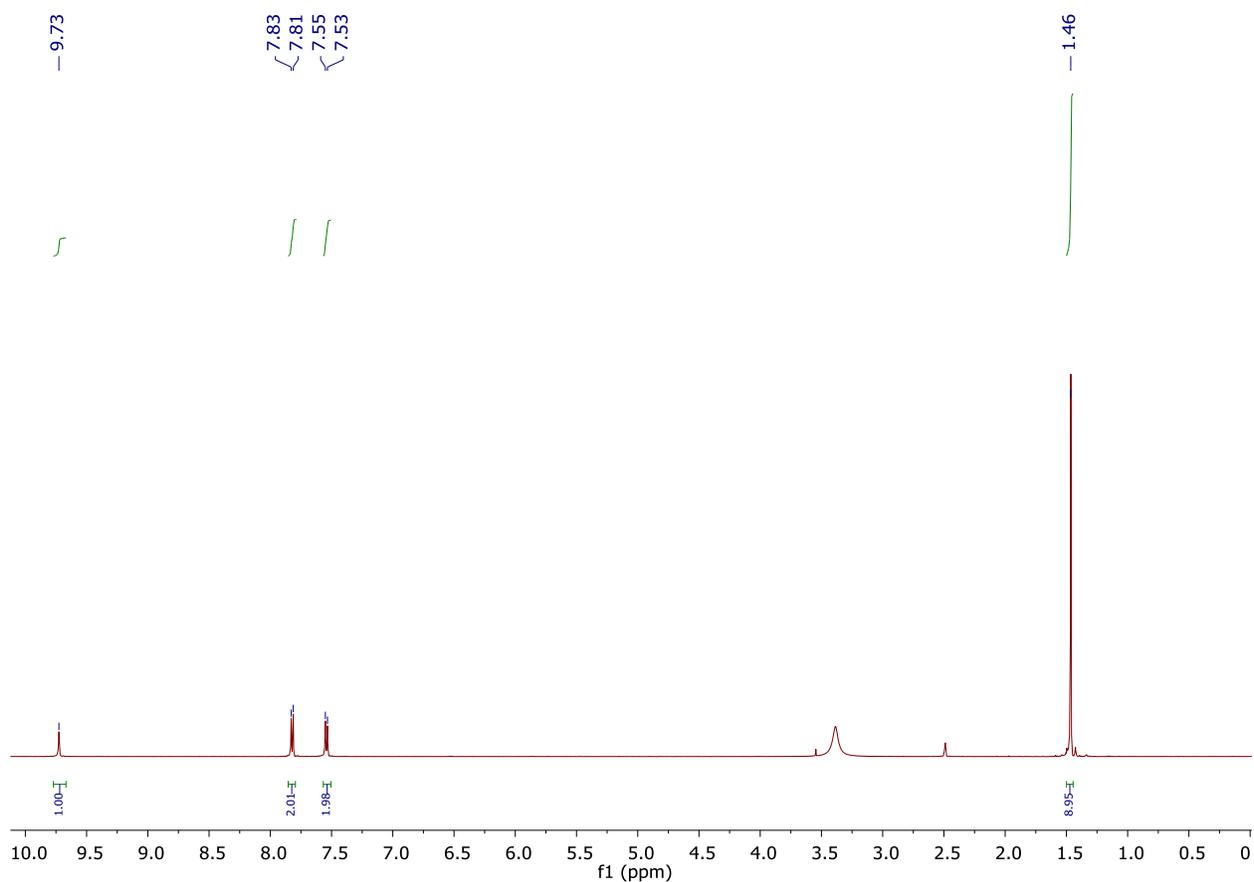
Synthesis of unmasked-AI-I



Synthesis of intermediate b41

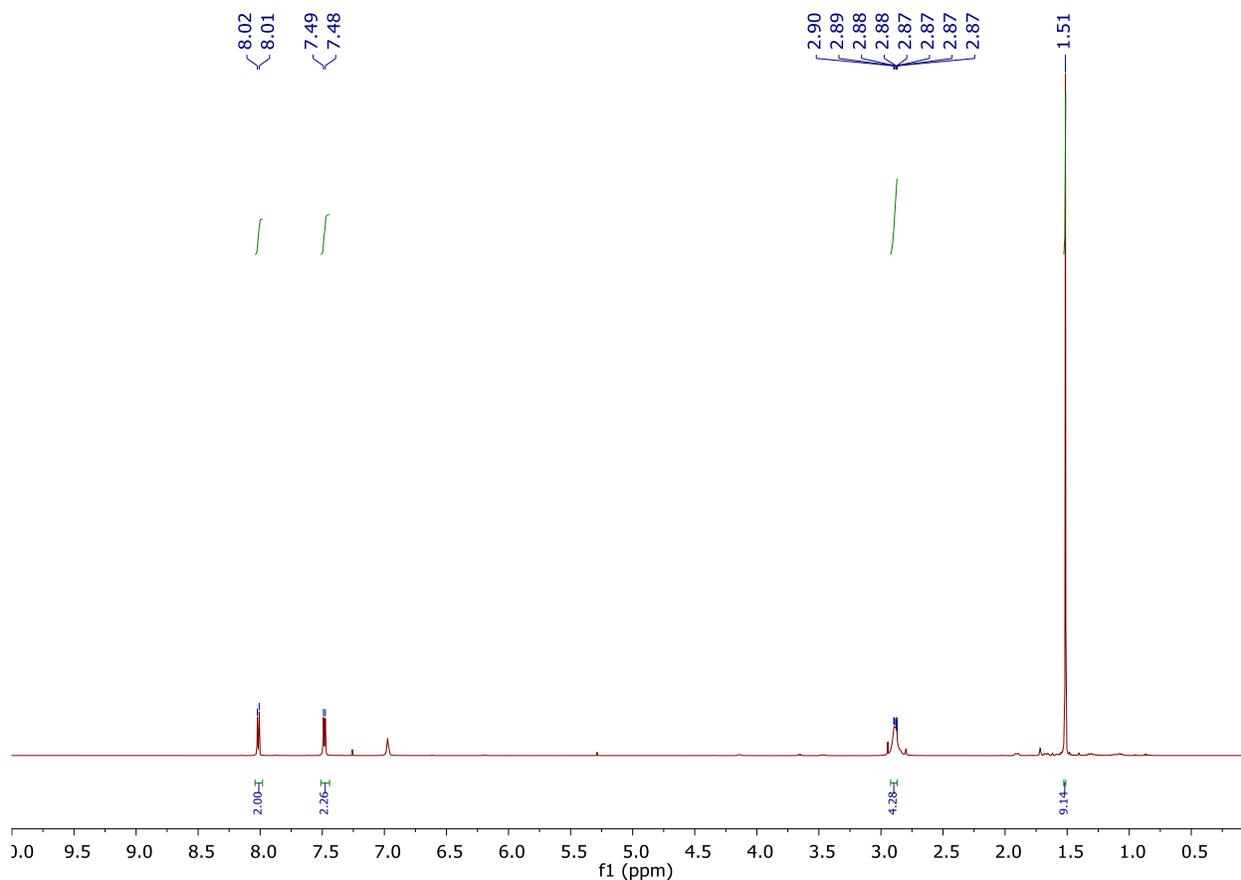
4-aminobenzoic acid **b40** (73 mmol) was dissolved in a mixture of 1,4 dioxane (30 mL) and NaOH 0.5 N (30 mL) at 0°C and Boc_2O (80.3 mmol) was added. Then the mixture was allowed to react at room temperature, under magnetic stirring for 24 hours. Next, the dioxane was removed under reduced pressure and the pH of the reaction was adjusted to 2 with 2N HCl. The obtained mixture was then extracted with ethyl acetate for three times, and the combined organic phases were dried over Na_2SO_4 , filtered off and concentrated under reduced pressure to give the

desired intermediates b41 (15.9 g, 92%). ^1H NMR (500 MHz, DMSO- d_6) δ 9.73 (s, 1H), 7.82 (d, $J = 8.3$ Hz, 2H), 7.54 (d, $J = 8.7$ Hz, 2H), 1.46 (s, 9H).



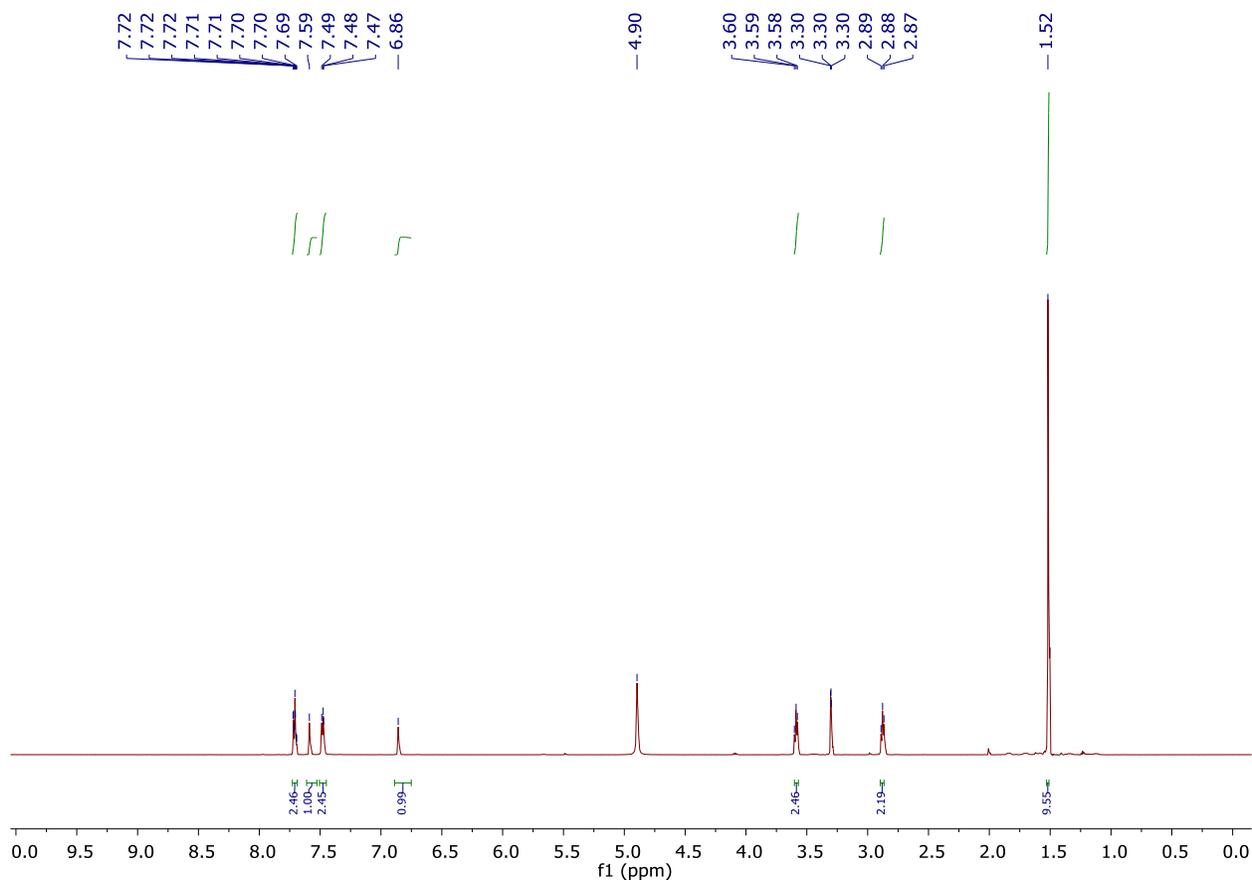
Synthesis of intermediate b42

2 g of b41 was dissolved in 30 mL dry DMF. To this solution 2.1 g of DCC (10.12 mmol) and 1.16 g of NHS (10.12 mmol) was added and stirred for 12 h at room temperature. Next, the reaction mixture was washed with ethyl acetate and brine for 3 times. The ethyl acetate solution was dried over anhydrous sodium sulfate and evaporated to yield the crude compound. The residue was purified by resolidification in Hex : DCM (95:5) to obtain the pure compound b42 as white solid (2.45 g, 87%). ^1H NMR (600 MHz, Chloroform- d) δ 8.02 (d, $J = 8.9$ Hz, 2H), 7.48 (d, $J = 8.9$ Hz, 2H), 2.92 – 2.87 (m, 4H), 1.51 (s, 9H).



Synthesis of intermediate b43

To a solution of intermediate b42 (1.49 mmol) in 10 mL of dry DMF, histamine (332.6 mg, 2.99 mmol) was added at room temperature. The reaction was allowed to stir for 2 h at room temperature. Next, the mixture was washed with EA and brine for three times. The organic layer was dried using anhydrous sodium sulphate and concentrated under reduced pressure. The crude product was purified by resolidification in Hex : DCM = 95 : 5 to obtain the pure compound b43 as white solid (740 mg, 75%). ^1H NMR (600 MHz, Methanol- d_4) δ 7.71 (ddd, $J = 9.1, 5.5, 2.2$ Hz, 2H), 7.59 (s, 1H), 7.51 – 7.45 (m, 2H), 6.86 (s, 1H), 3.58 (d, $J = 6.2$ Hz, 2H), 2.88 (t, $J = 7.2$ Hz, 2H), 1.52 (s, 10H).



Synthesis of intermediate b44

A solution of b43 (27.5 mg, 83 μmol), b3 (50 mg, 125 μmol), and pyridine (50 μL , 250 μmol) in dry DMF (5 mL) was stirred for 12 h at rt. After removal of the solvent in vacuo, the residue was washed with EA and brine for 3 times. The organic layer was drying over Na_2SO_4 . After removal of the solvent by evaporation, the residue was purified by flash column chromatography on SiO_2 (DCM : MeOH = 50 : 1 \rightarrow 20 : 1) to produce the intermediate b44 as a colorless oil (26.0 mg, 51%).

Synthesis of unmasked-AI-I

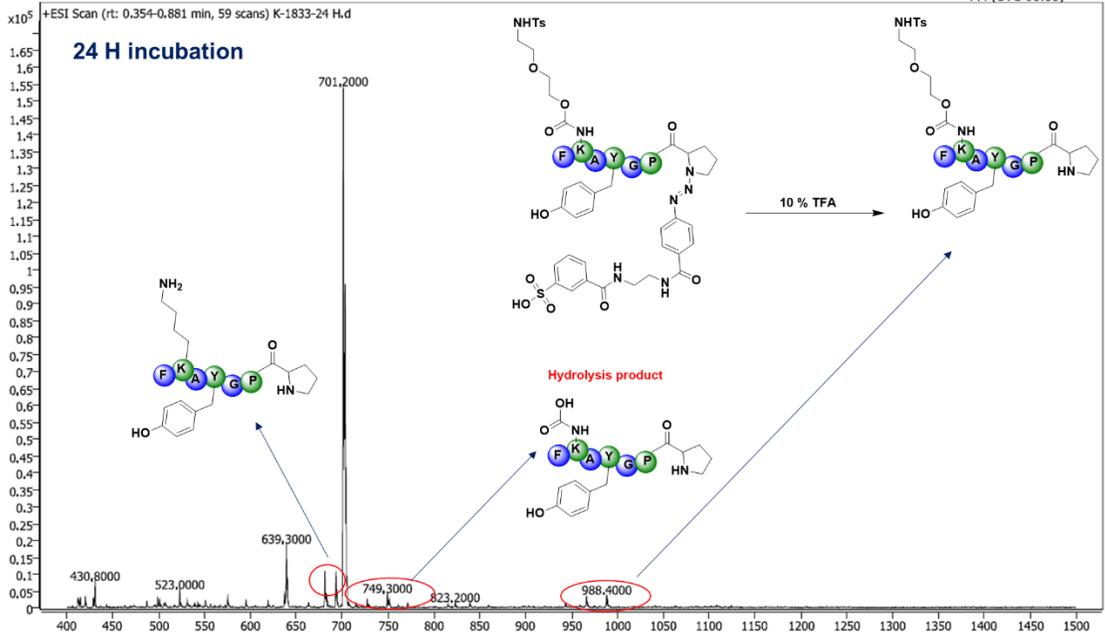
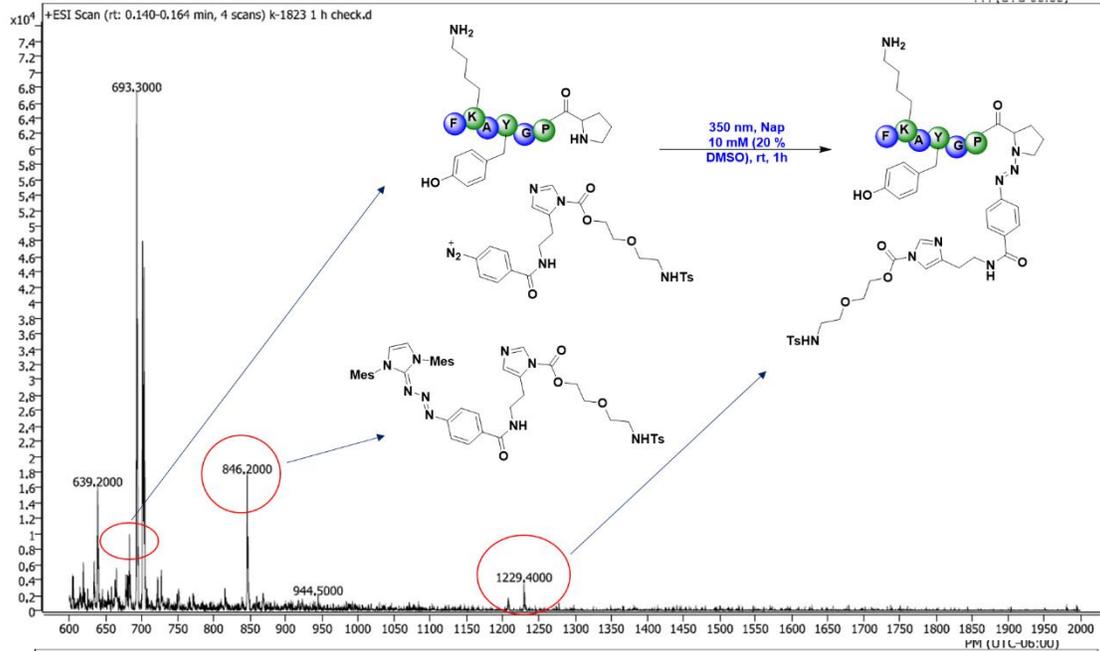
The appropriate compound b44 (0.083 mmol) was dissolved in a mixture of dioxane (5 mL) and HCl 4N (2 mL) and stirred for 24 h. Then the solvent was evaporated under vacuum and the crude was purified resolidification in ether : methanol = 95 : 5 to obtain the pure deprotect

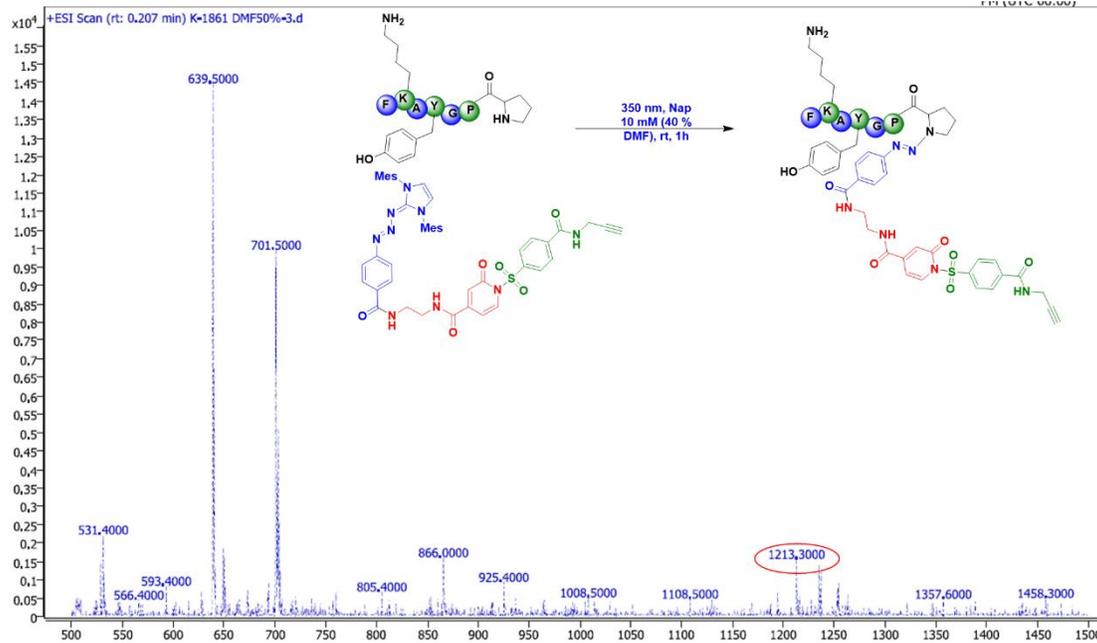
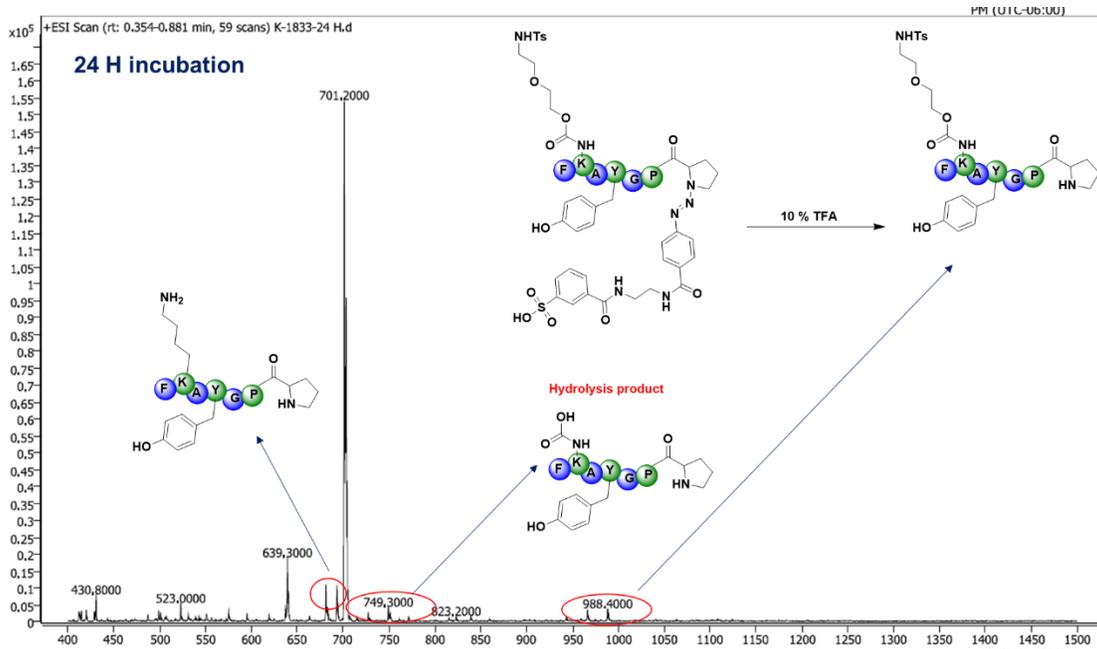
intermediate. To this 48% HBF_4 (100 μL), EtOH (50 μL) and H_2O (50 μL), the reaction mixture was cooled down to 0 $^\circ\text{C}$, the NaNO_2 (1.2 equiv., 0.1 mmol) in distilled H_2O (50 μL) was added dropwise to the mixture over a period of 5 min. The reaction was allowed to stir at 0 $^\circ\text{C}$ for 45 min. The residue was washed with Et_2O , and dried under vacuum to afford **unmasked-AI-I** as yellow powder (3.6 mg, 72%)

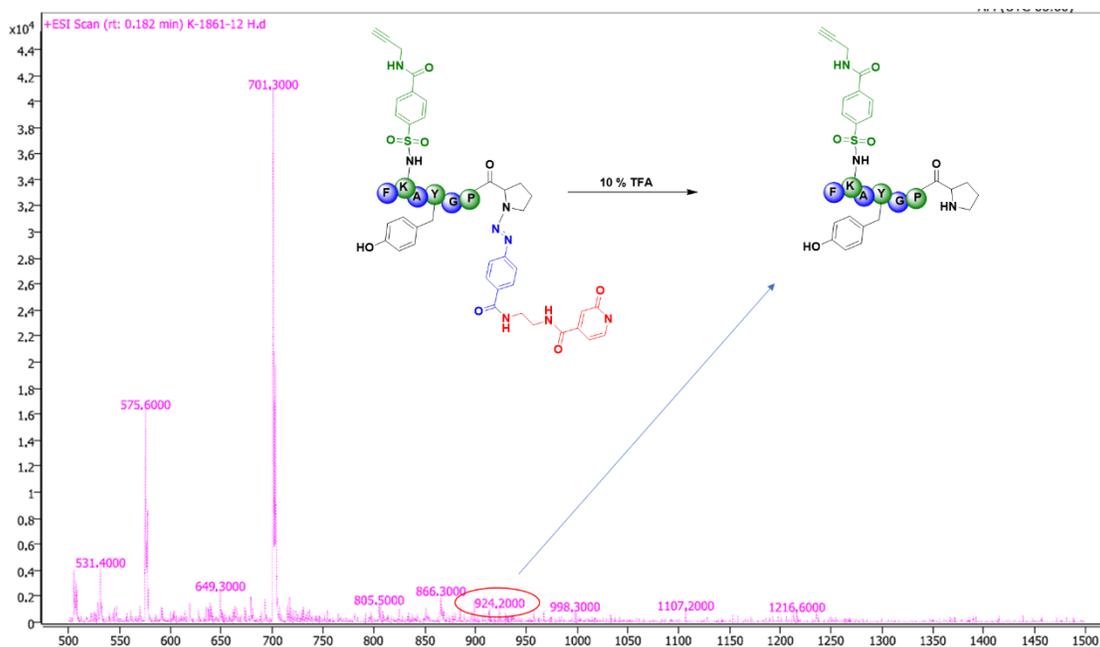
Supplementary Figure 2. Optimization of intramolecular reaction under UV lamp with different proline containing peptides and chemo selectivity study

General procedure for triazene cyclization:

1 mg (1eq, 0.0014 mmol) of k-2, k-3, k-4 peptide with 3 eq of K_2CO_3 (0.0042 mmol) in 240 μl of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of AI, SP, NASA (0.0042 mmol) in 160 μl DMF were mixed together, the reaction mixture was incubated at 25 $^\circ\text{C}$ for 1 h under 365 nm light. Next, the reaction mixture was incubated at 37 $^\circ\text{C}$. After 12 h, 20 μl of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



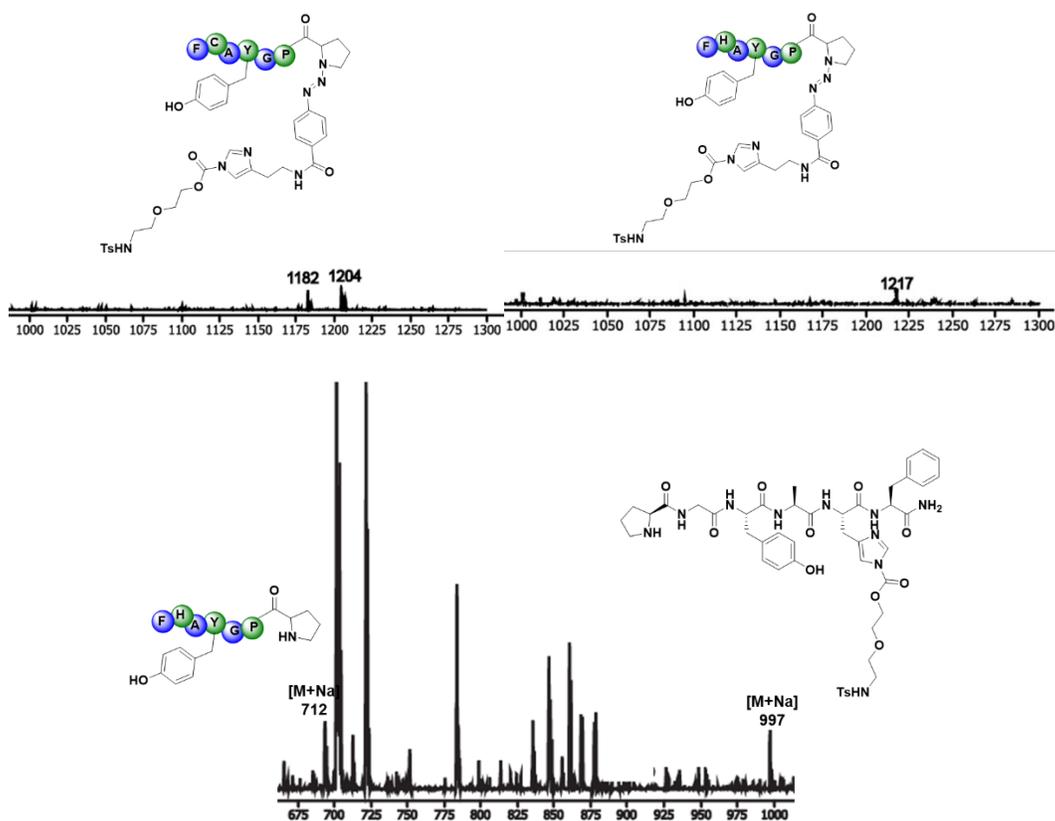




Supplementary Figure 3. Chemoselectivity and control study of AI-I probe

General procedure for triazene cyclization:

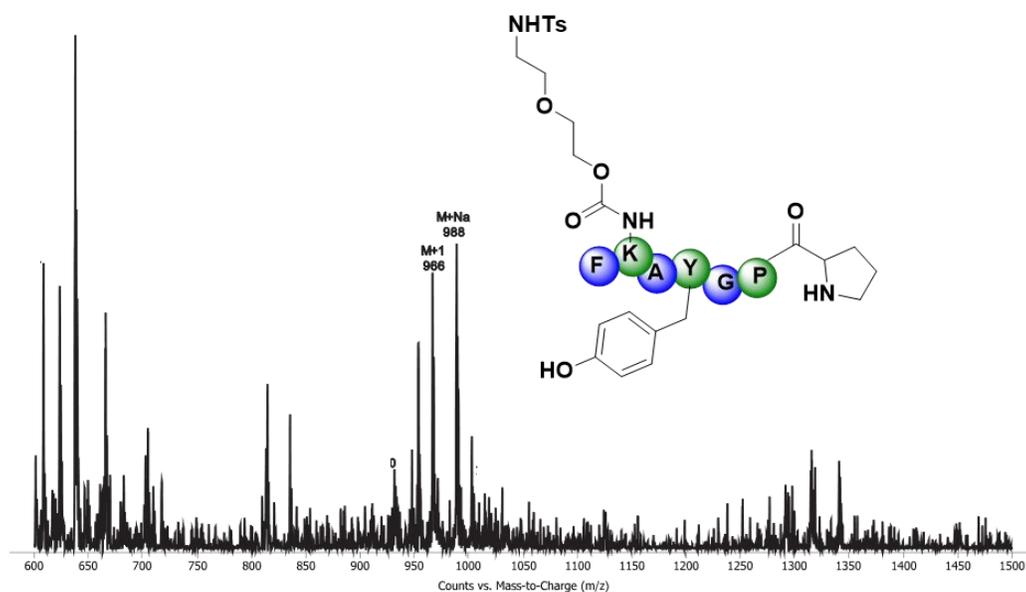
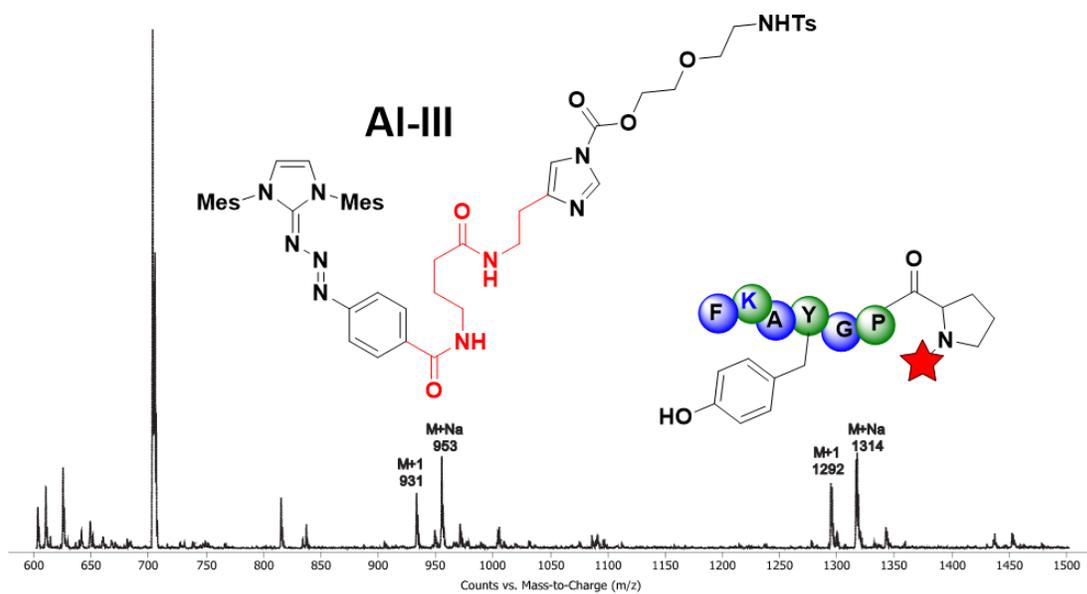
(1eq, 0.0014 mmol) of C2, H2, FRQDW-NH-Ac with 3 eq of K_2CO_3 (0.0042 mmol) in 240 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of AI-I (0.0042 mmol) in 160 μ l DMF were mixed together, the reaction mixture was incubated at 25 $^{\circ}C$ for 1 h under 365 nm light. Next, the reaction mixture was incubated at 37 $^{\circ}C$. After 12 h, 20 μ l of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



Supplementary Figure 4. Screening of AI-III probe with K-2, K-3 and K-4 peptides under UV lamp

General procedure for triazene cyclization:

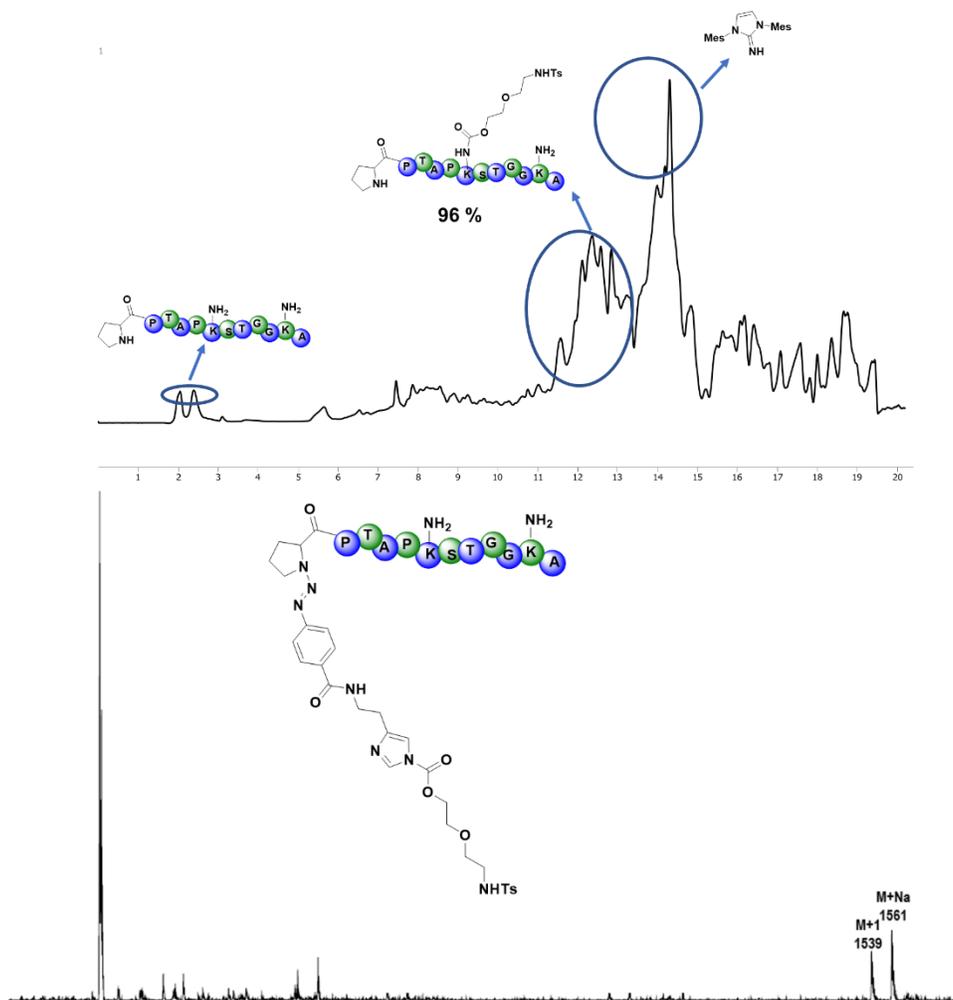
1 mg (1eq, 0.0014 mmol) of k-2, k-3, k-4 peptide with 3 eq of K_2CO_3 (0.0042 mmol) in 240 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of AI-III (0.0042 mmol) in 160 μ l DMF were mixed together, the reaction mixture was incubated at 25 $^\circ$ C for 1 h under 365 nm light. Next, the reaction mixture was incubated at 37 $^\circ$ C or 25 $^\circ$ C. After 12 h, 20 μ l of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

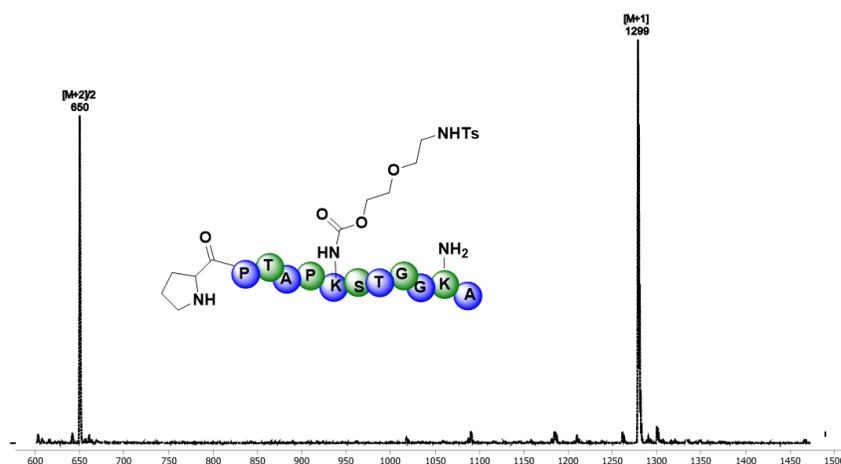


Supplementary Figure 5. Site-specific study with multiple lysins containing peptide
General procedure for triazene cyclization:

(1eq, 0.98 μmol) of PTAPKSTGGKA peptide with 3 eq of K_2CO_3 (2.94 μmol) in 120 μl of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of AI-III (2.94 μmol) in 80 μl DMF were mixed together, the reaction mixture was incubated at 25 $^\circ\text{C}$ for 1 h under 365 nm light. Next, the reaction mixture was incubated at 37 $^\circ\text{C}$. After 12 h, 20 μl of TFA was added to yield the

intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.





Supplementary Figure 6. Digestion of modified peptide.

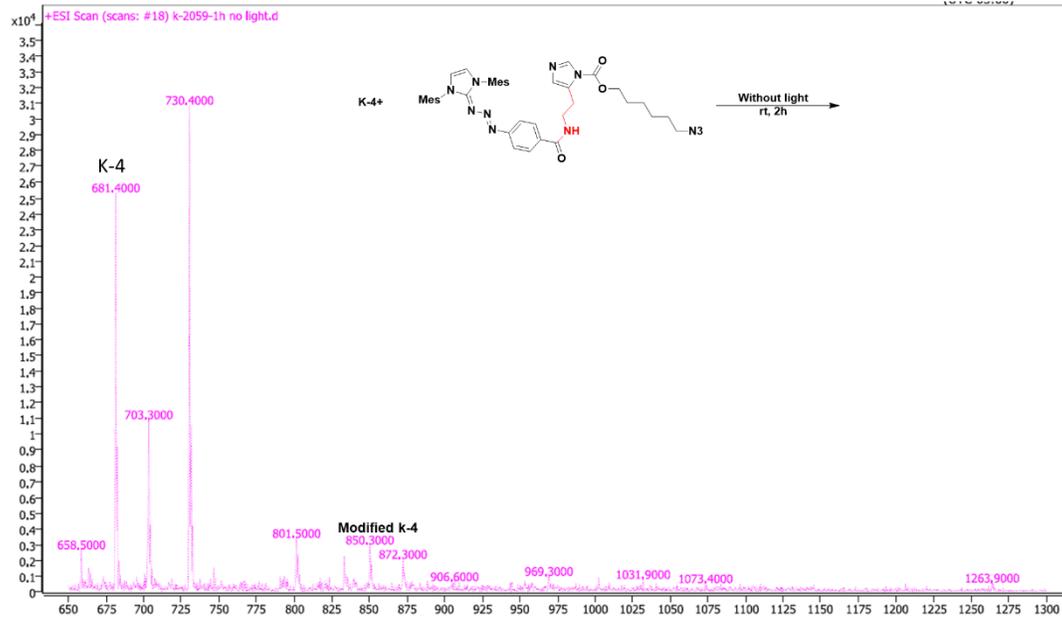
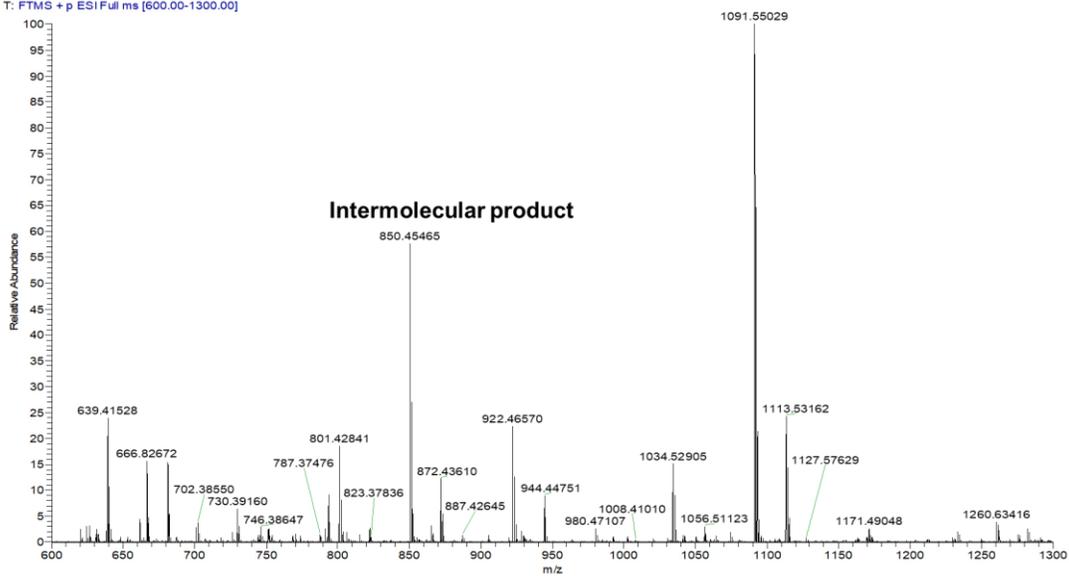
Modified peptide (0.98 μmol) in 120 μl of 10 mM sodium phosphate buffer (Nap, pH 7.5) for 30 min with 20 % of trypsin at 37 $^{\circ}\text{C}$. The reaction mixture was filtered and purified by HPLC to obtain the fragments. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

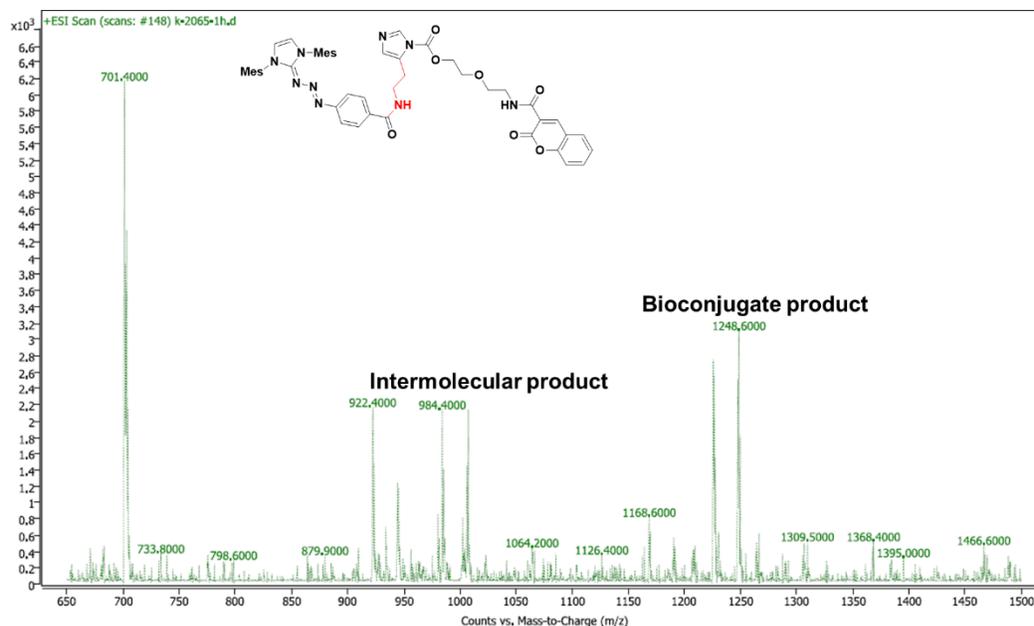
Supplementary Figure 7. Intramolecular reaction study of various AI analogs with different affinity tags.

1 mg (1eq, 0.0014 mmol) of k-2 peptide with 3 eq of K_2CO_3 (0.0042 mmol) in 240 μl of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of AI analogs (0.0042 mmol) in 160 μl DMF were mixed together, the reaction mixture was incubated at 25 $^{\circ}\text{C}$ for 1 h under 365 nm light. Next, the reaction mixture was incubated at 37 $^{\circ}\text{C}$ or 25 $^{\circ}\text{C}$. After 12 h, 20 μl of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

k-2059-1h #18 RT: 0.31 AV: 1 NL: 7.08E6
 T: FTMS + p ESI Full ms [600.00-1300.00]

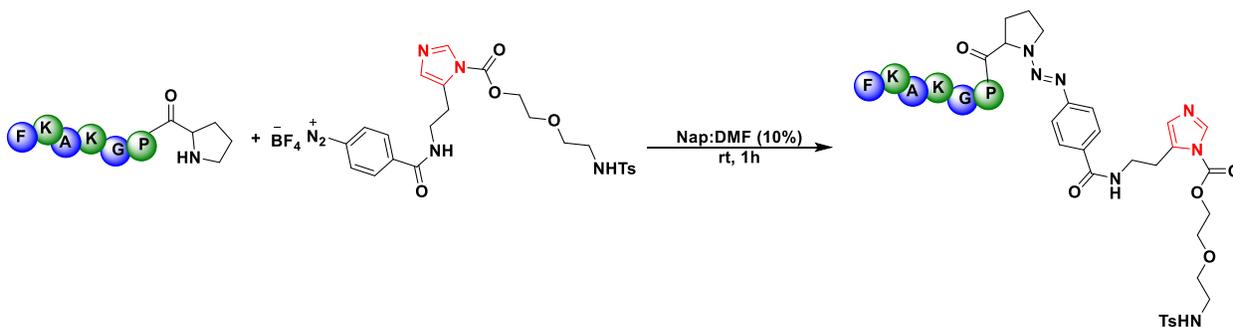
Bioconjugate product

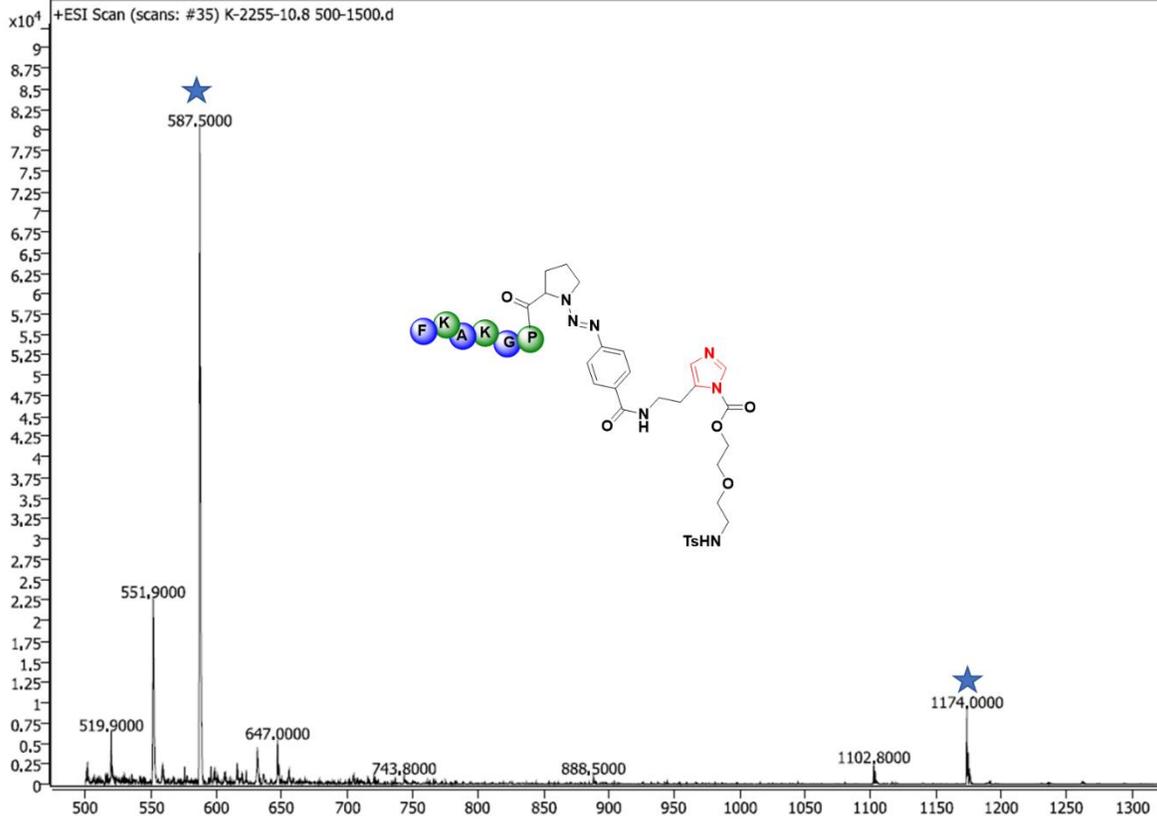
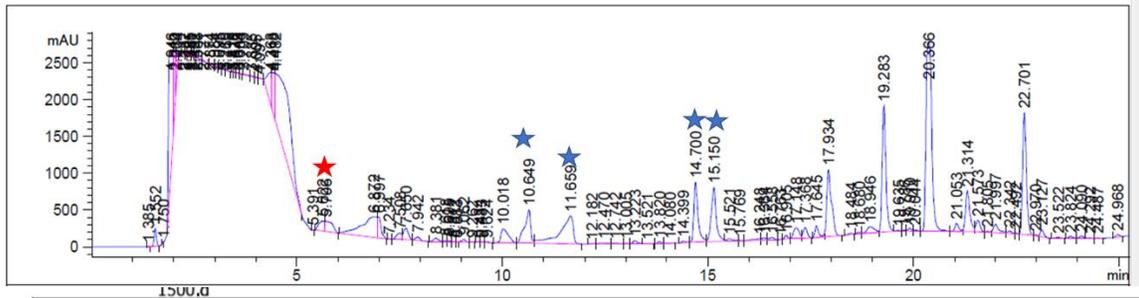




Supplementary Figure 8. Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe

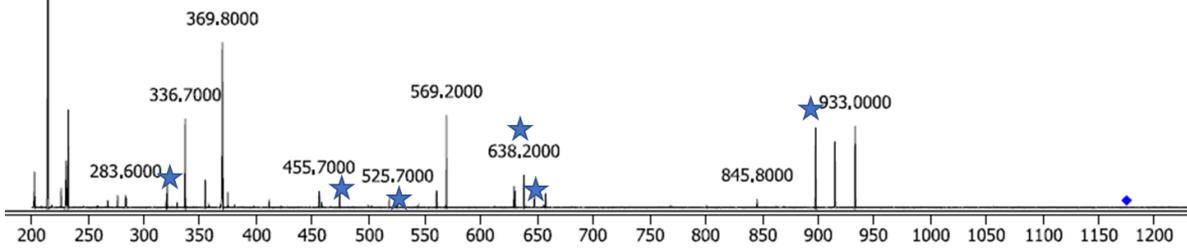
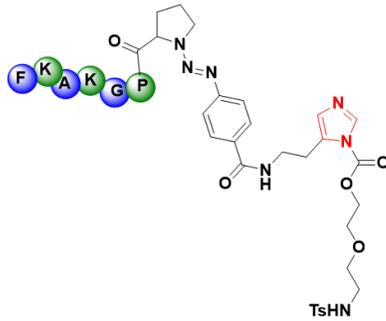
1 mg (1eq, 0.00155 mmol) of PGKAKF with 3 eq of K_2CO_3 (0.0042 mmol) in 320 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of unmasked-AI (0.0042 mmol) in 80 μ l DMF were mixed together, the reaction mixture was incubated at 25 $^{\circ}C$ for 1 h. Next, the reaction mixture was incubated at 37 $^{\circ}C$. After 6 h, 20 μ l of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.





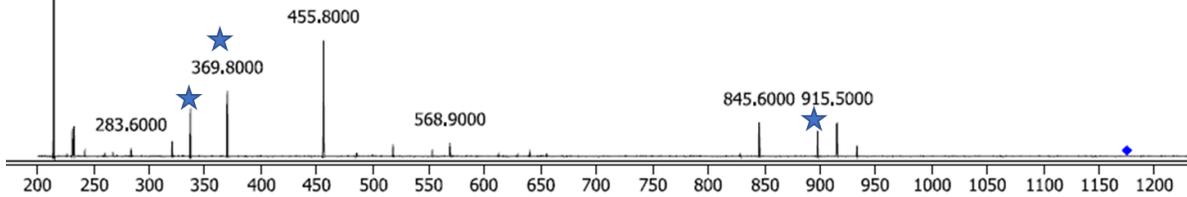
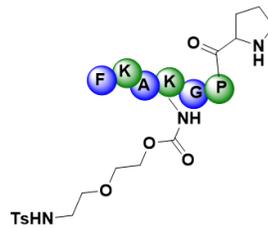
214,5000

MS/MS of peak 10.8

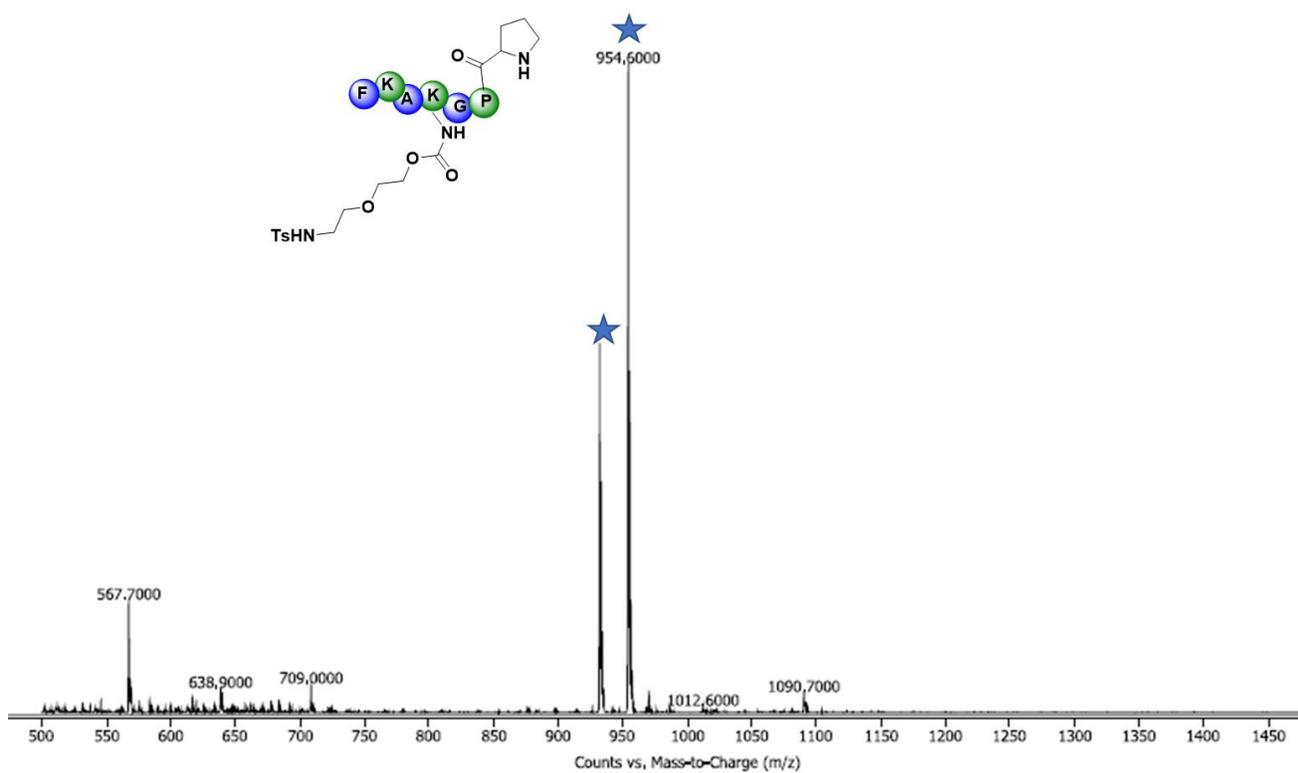


214,5000

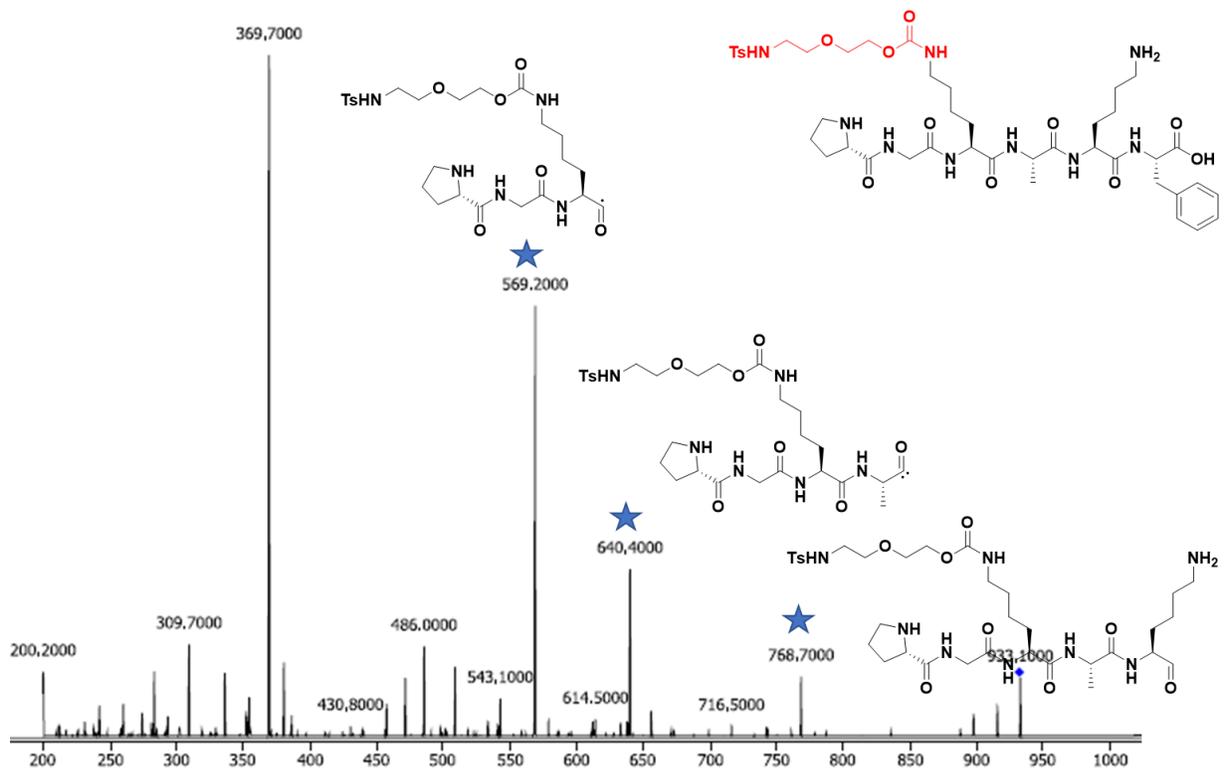
MS/MS of peak 14.6



+ESI Scan (scans: #26) k-2255-TFA-10,4,d

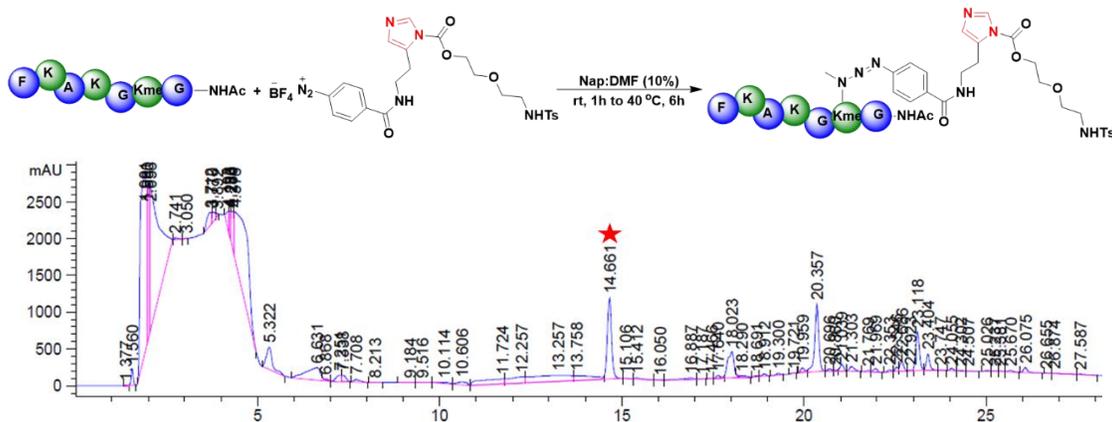


+ESI Product Ion (scans: #32) Frag=135,0V CID@50,0 (932,0 -> **) k-2255-final product,d

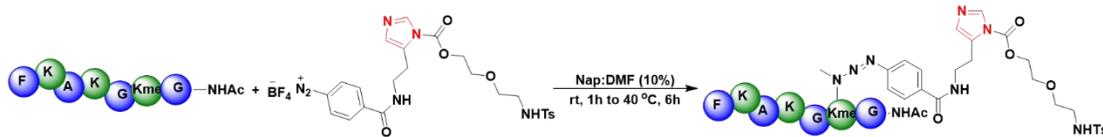
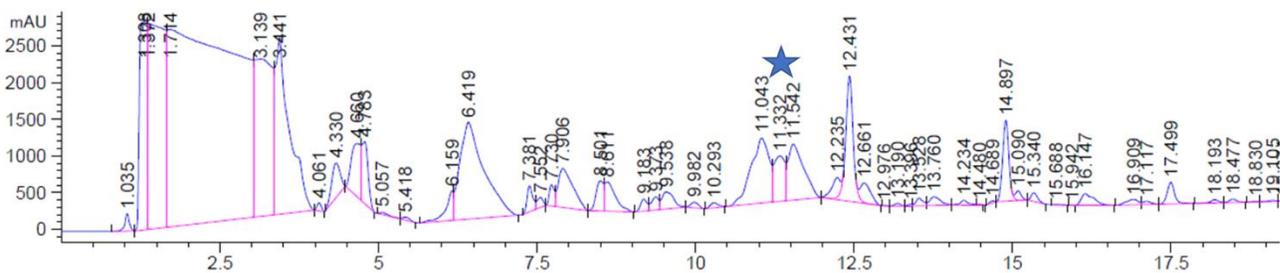
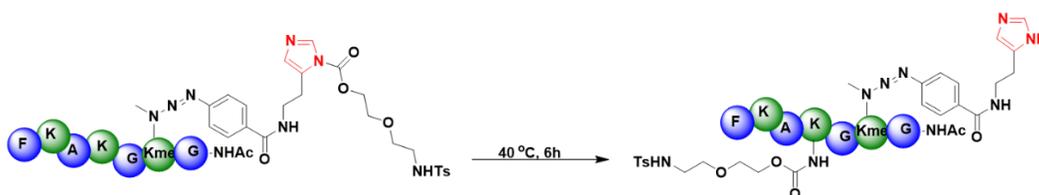
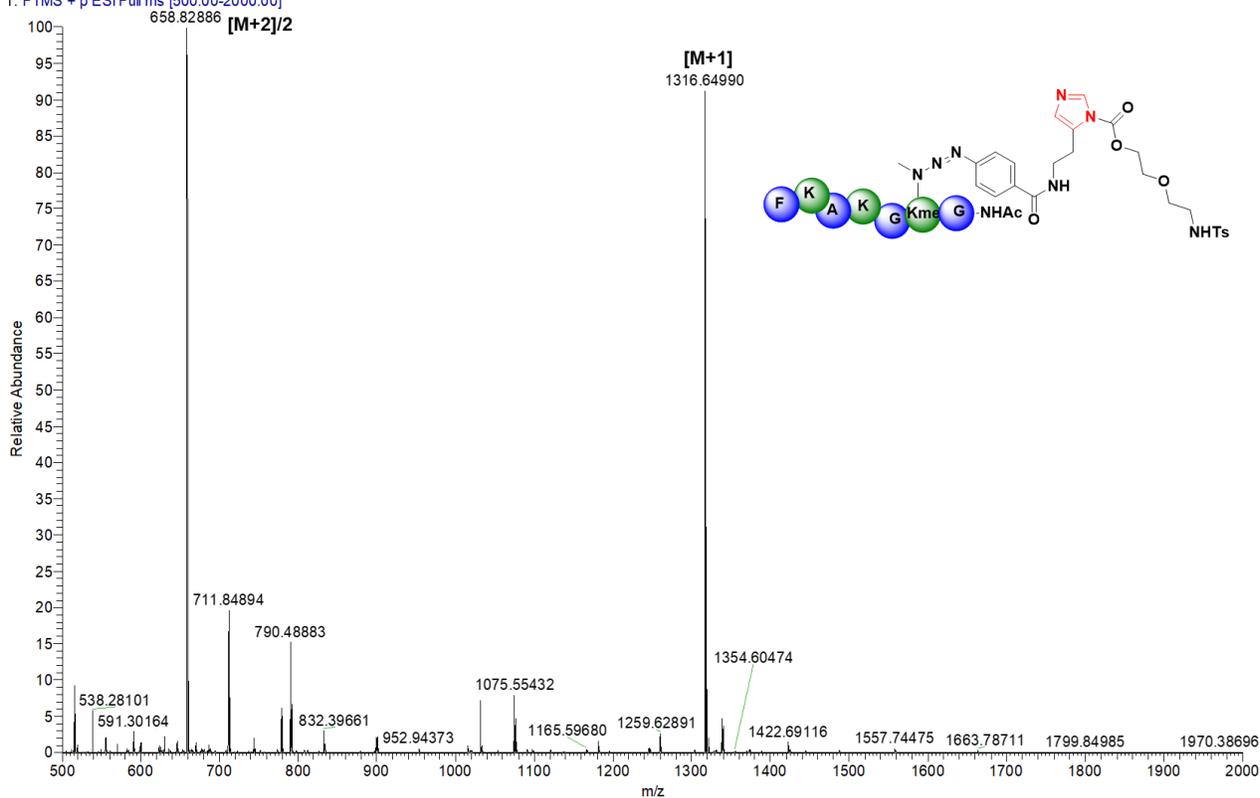


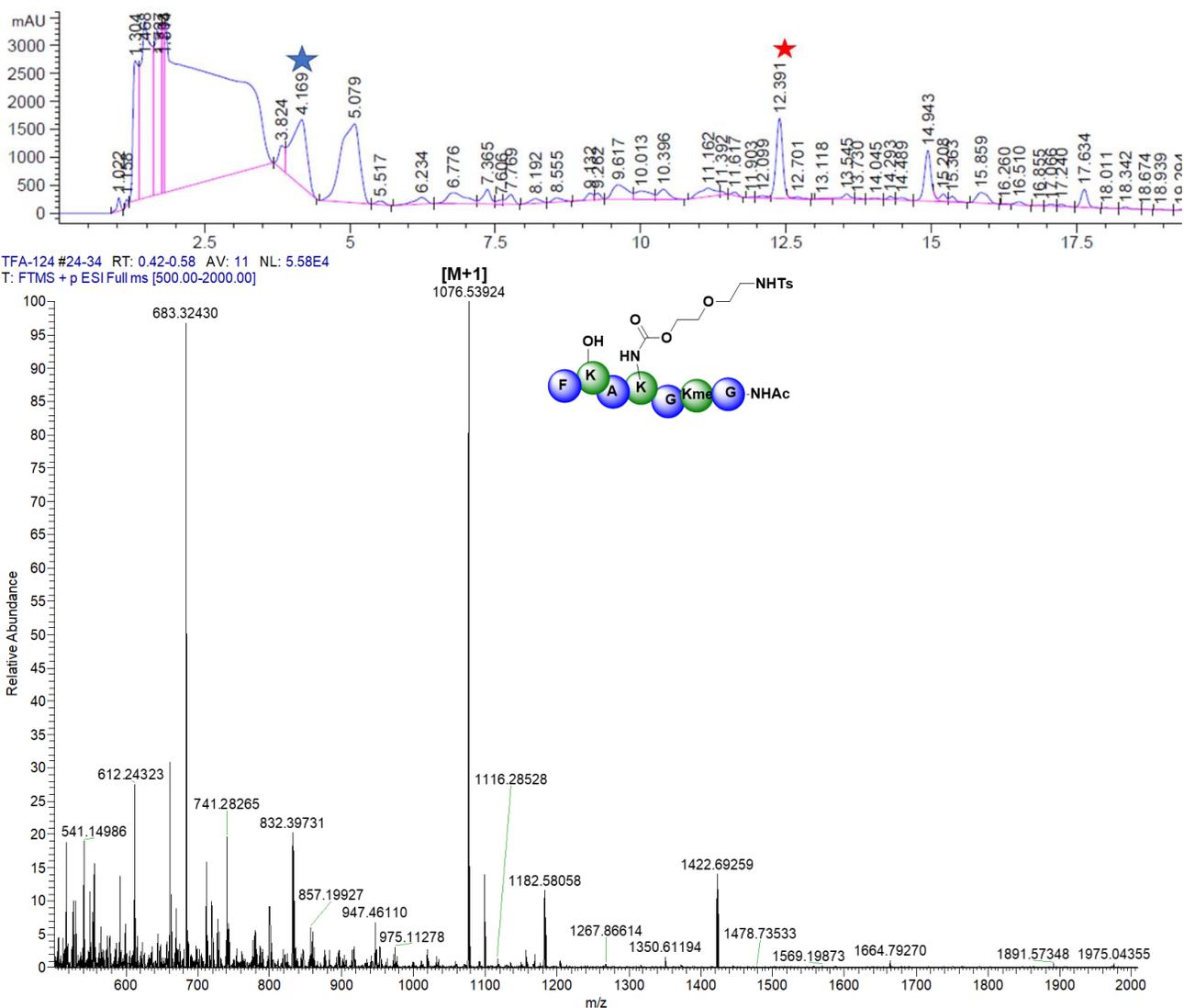
Supplementary Figure 9. Intramolecular reactivity and chemo selectivity study of unmasked AI-I probe with K-me peptide

1 mg (1eq, 0.00155 mmol) of Ac-GKmeGKAKF with 3 eq of K_2CO_3 (0.0042 mmol) in 320 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of unmasked-AI (0.0042 mmol) in 80 μ l DMF were mixed together, the reaction mixture was incubated at 25 $^{\circ}C$ for 1 h. Next, the reaction mixture was incubated at 37 $^{\circ}C$. After 6 h, 20 μ l of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



112-116 #23 RT: 0.36 AV: 1 NL: 2.90E7
 T: FTMS + p ESI Full ms [500.00-2000.00]





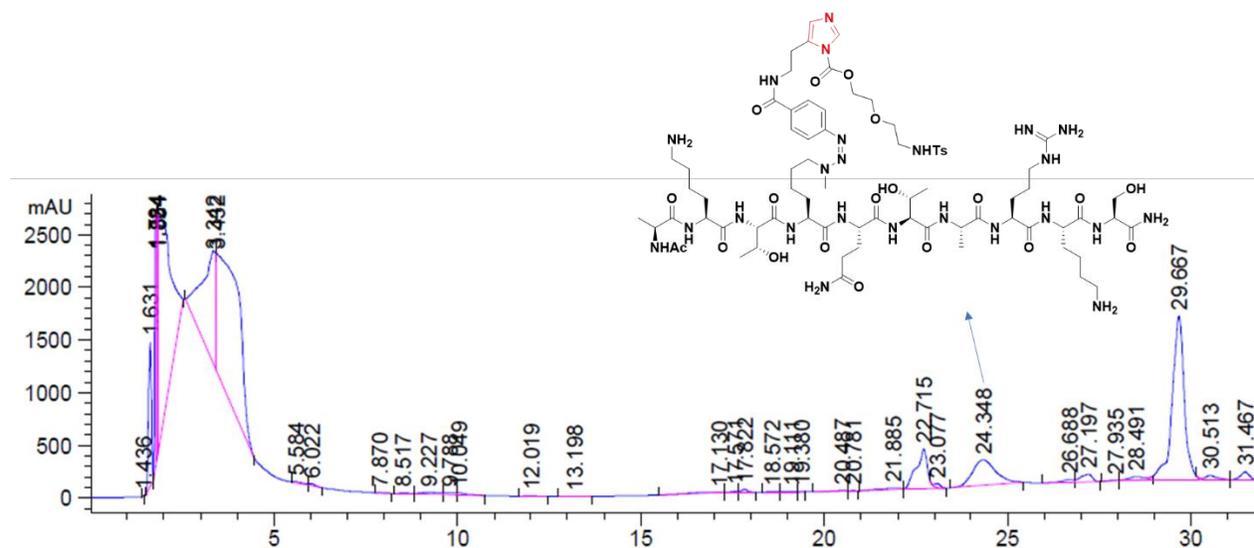
Supplementary Figure 10. Chemo-selective study of unmasked AI probe

(1eq, 0.00155 mmol) of NH₂-GHGKAKF or NH₂-GCGKAKF with 3 eq of K₂CO₃ (0.0042 mmol) in 320 μl of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of unmasked-AI (0.0042 mmol) in 80 μl DMF were mixed together, the reaction mixture was incubated at 25 °C for 1 h. Next, the reaction mixture was incubated at 37 °C. After 6 h, 20 μl of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water

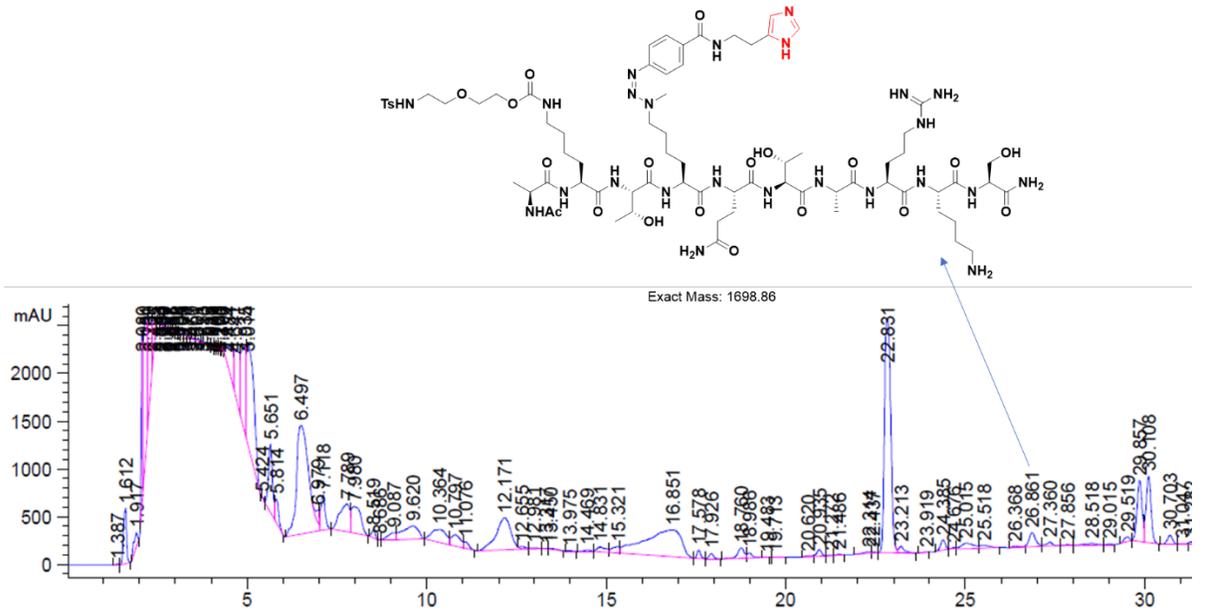
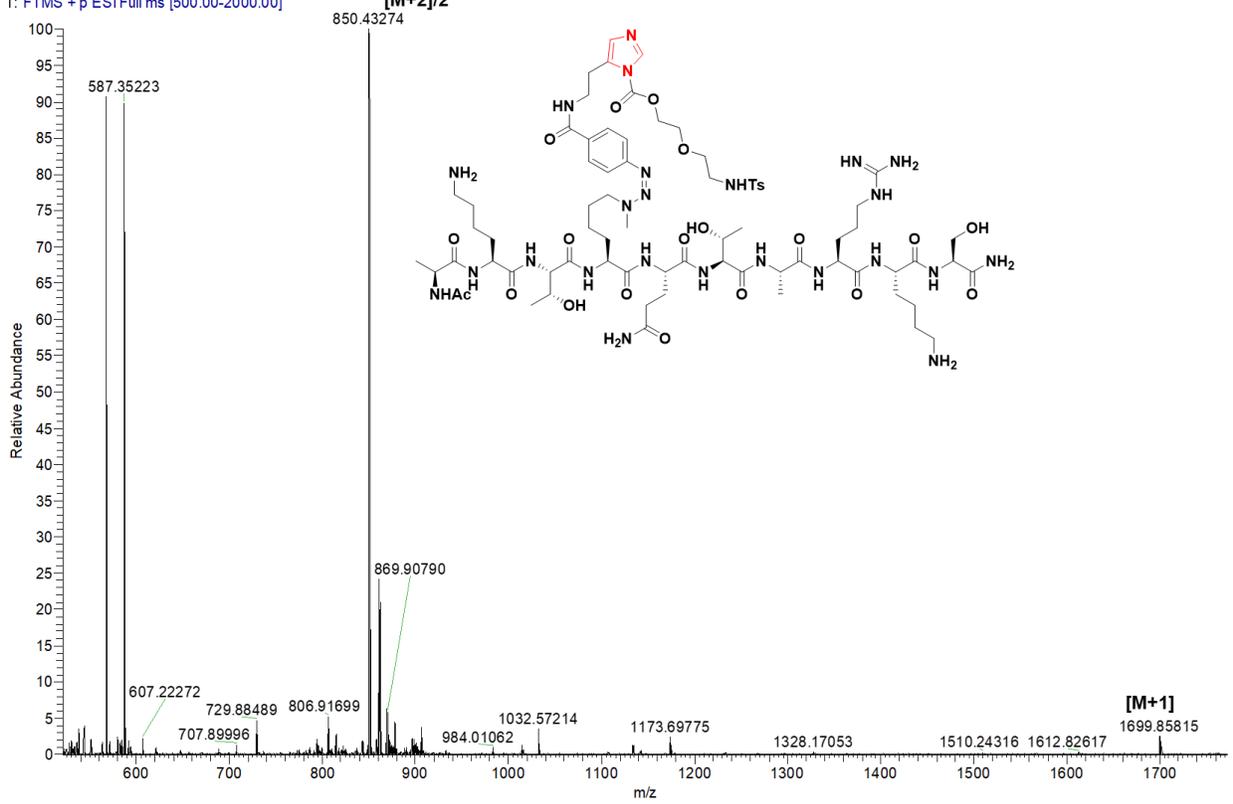
(solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

Supplementary Figure 11. Kme-Directed modification of truncated histone peptide

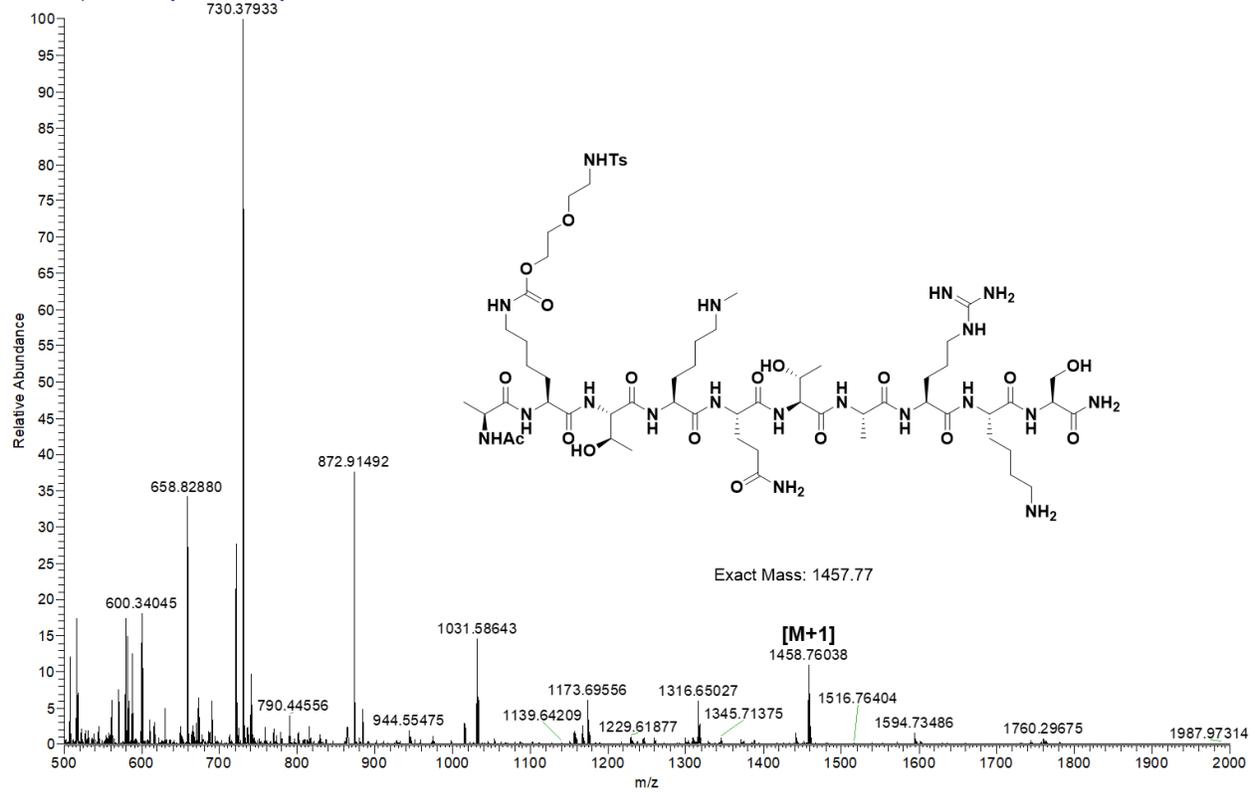
1.8 mg (1eq, 0.00155 mmol) of Ac-AKTK(Me)QTARKS with 3 eq of K_2CO_3 (0.0042 mmol) in 320 μ l of 10 mM sodium phosphate buffer (Nap, pH 7.5), and 3 eq of unmasked-AI (0.0042 mmol) in 80 μ l DMF were mixed together, the reaction mixture was incubated at 25 °C for 1 h. Next, the reaction mixture was incubated at 37 °C. After 6 h, 20 μ l of TFA was added to yield the intermolecular modification product. The reaction mixture was filtered and purified by HPLC to obtain the aldehyde product. HPLC was carried out with 1 % formic acid: water (solvent A): acetonitrile (solvent B); 0-80 % in 30 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.



248_220216195208 #34 RT: 0.51 AV: 1 NL: 7.75E5
 T: FTMS + p ESI Full ms [500.00-2000.00]

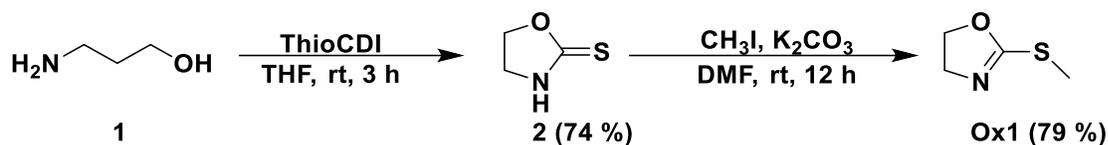


k-2301-h5-tfa #194 RT: 2.83 AV: 1 NL: 3.57E6
T: FTMS + p ESI Full ms [500.00-2000.00]
730.37933

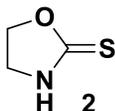


NMR for Chapter One

Synthesis of Ox1



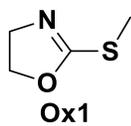
To a stirred solution of ethanolamine **1** (1.0 ml, 16.37 mmol) in THF (20 mL) was added thiocarbonyldiimidazole (3.2 g, 16.37 mmol) at 0 °C. The reaction was allowed to warm to room temperature and stirred at room temperature for 3 h. The reaction mixture was concentrated, washed with aqueous sodium bicarbonate solution three times, and dried over anhydrous MgSO₄. The crude was purified by flash chromatography (hexane: ethyl acetate 1:1) to afford compound **2** as white solid (74 %, 1.4g).



¹H NMR (600 MHz, d-DMSO): δ 9.89 (s, 1H), 4.56 (t, *J* = 9.1 Hz 2H), 3.65 (t, *J* = 9.1 Hz, 2H).

¹³C NMR (150 MHz, d-DMSO): δ 188.63, 69.82, 43.81.

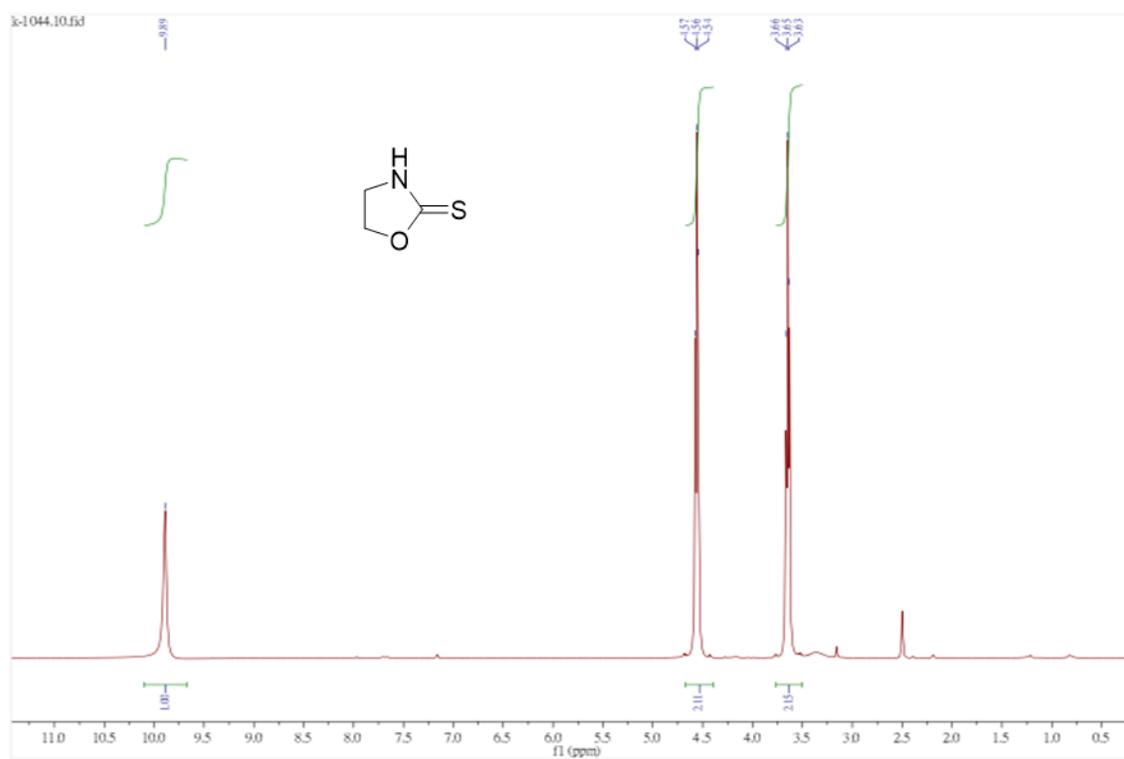
To a solution of oxazolidine-2-thione **2** (967 mg, 9.35 mmol), in 10 mL DMF at 0 °C, K₂CO₃ (1.3g, 9.35 mmol) was added. After 10 mins methyl iodide (641 μL, 10.3 mmol) was added drop wise and the reaction mixture was stirred at room temperature for 12 h. The reaction mixture was washed with ethyl acetate and brine. The organic part was dried over anhydrous MgSO₄, filtered, and concentrated to afford compound **Ox1** as brown oil (79%, 864 mg).



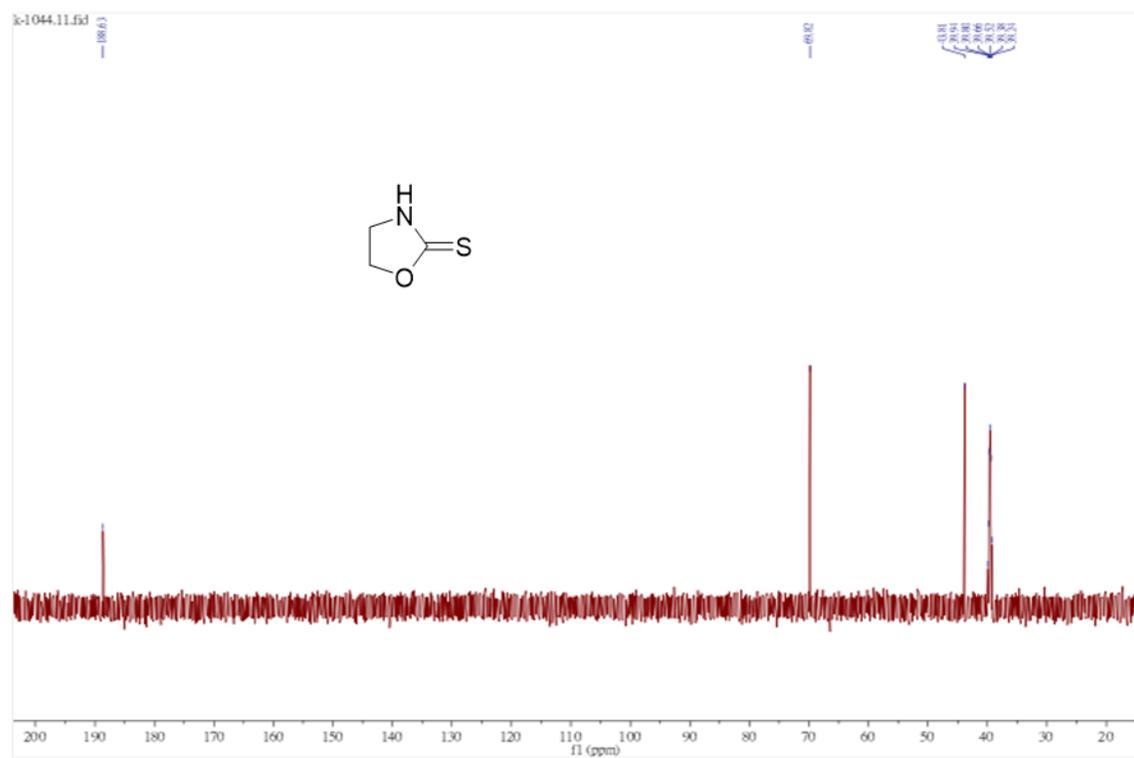
^1H NMR (600 MHz, CDCl_3): δ 4.34 (t, $J = 9.4$ Hz 2H), 3.85 (t, $J = 9.4$ Hz, 2H), 2.41 (s, 3H). ^{13}C

NMR (150 MHz, CDCl_3): δ 166.81, 69.35, 54.49, 14.34.

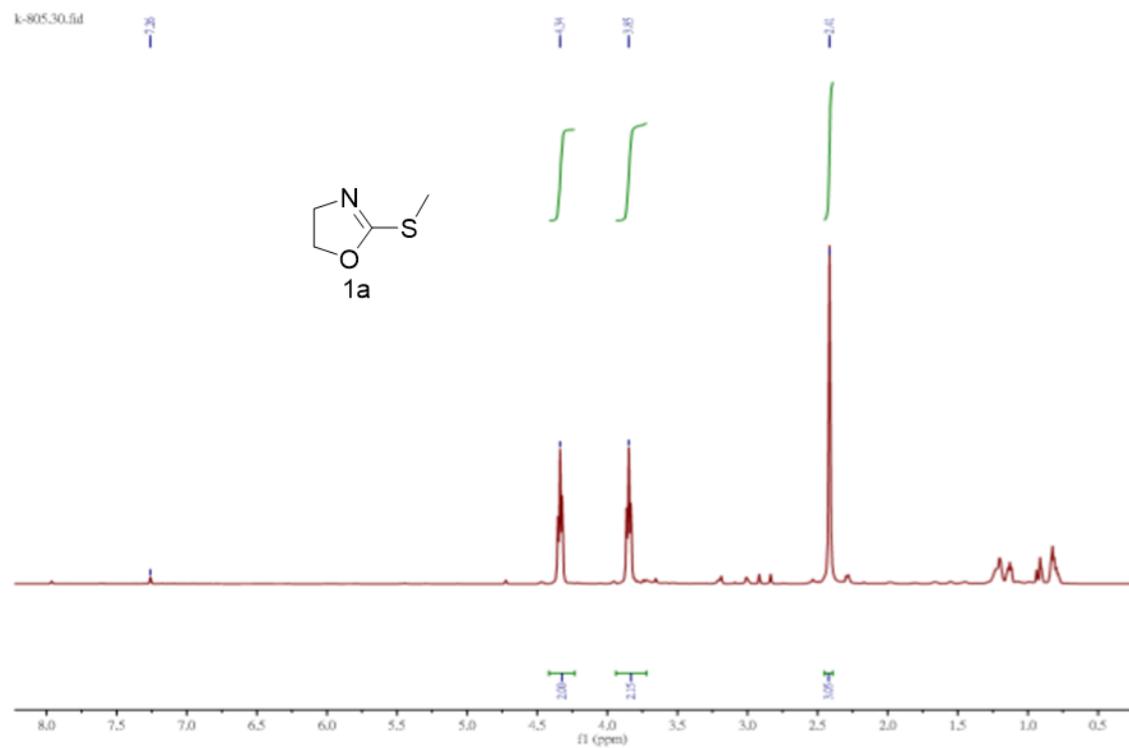
^1H NMR spectra of **2**



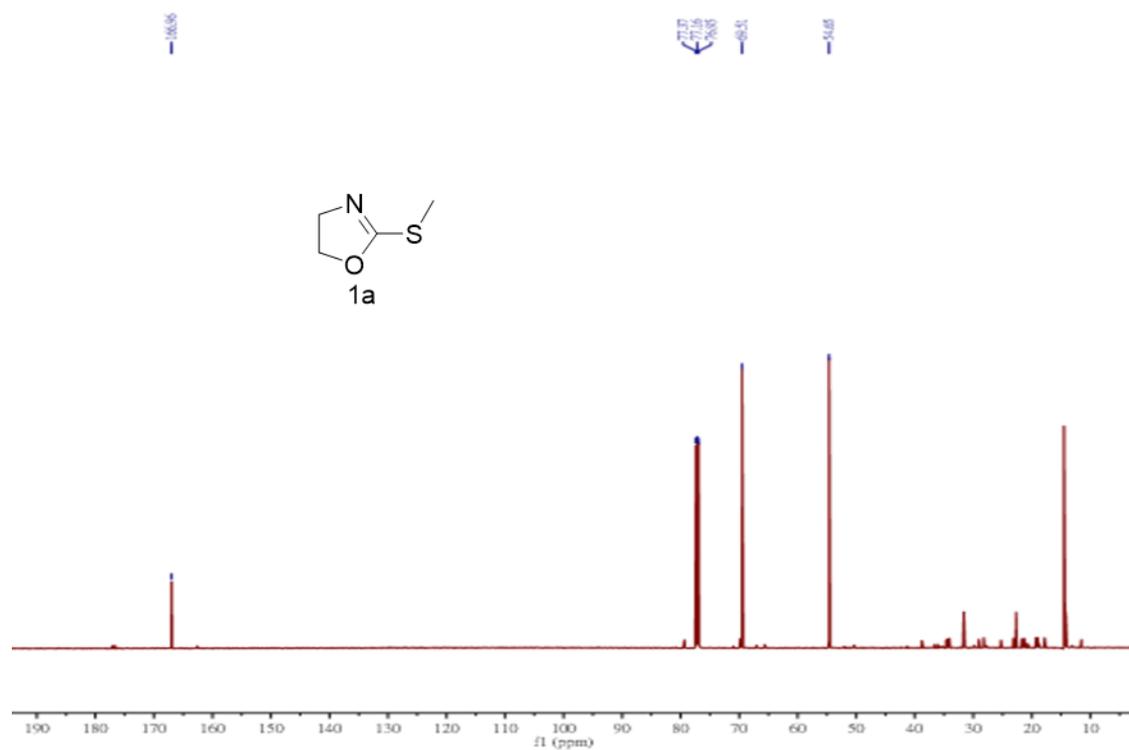
^{13}C NMR spectrum of **2**



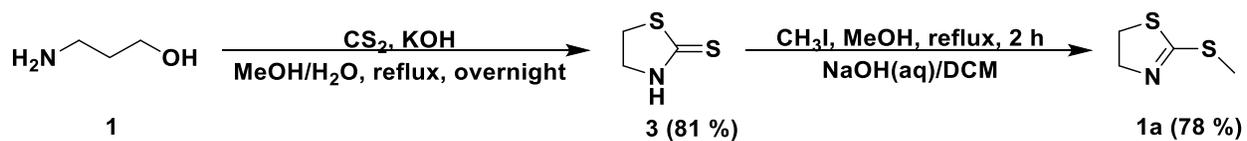
¹H NMR spectra of Ox1



¹³C NMR spectrum of Ox1

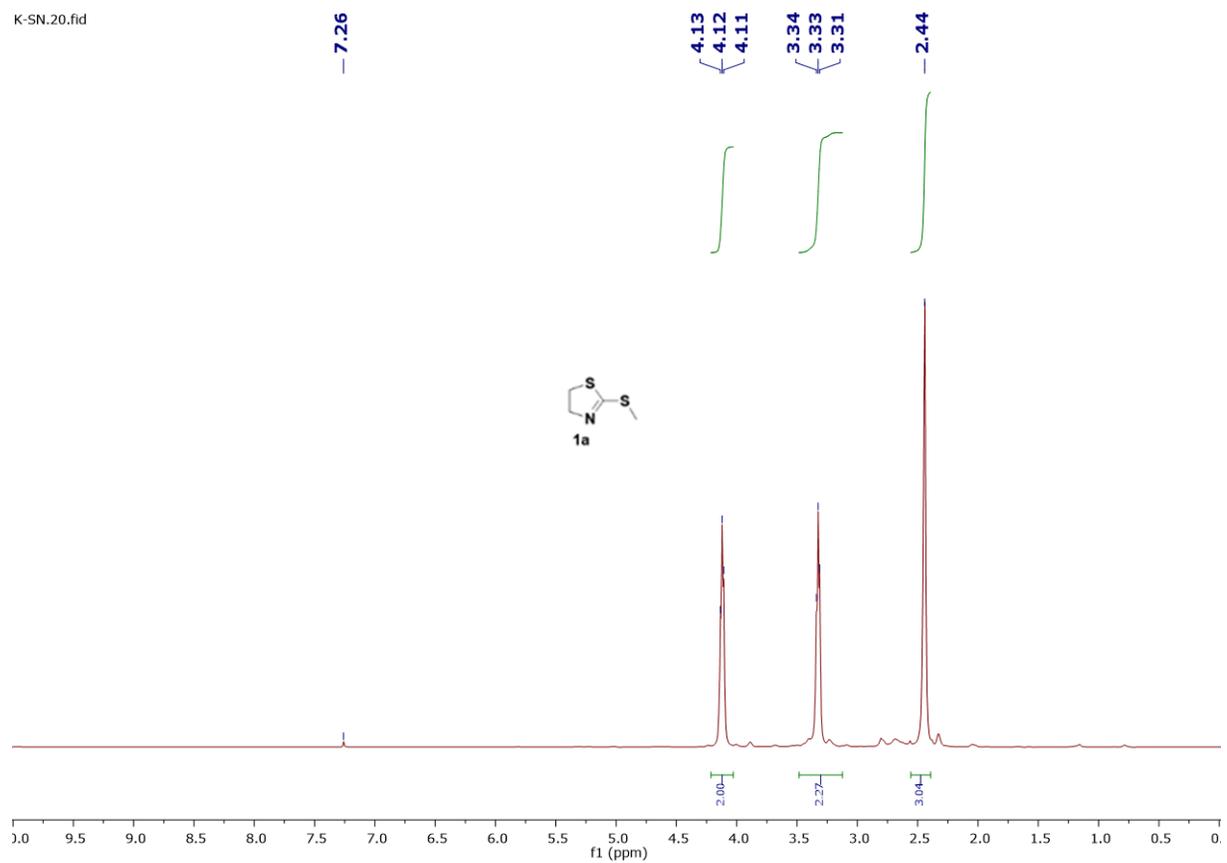


Synthesis of 1a



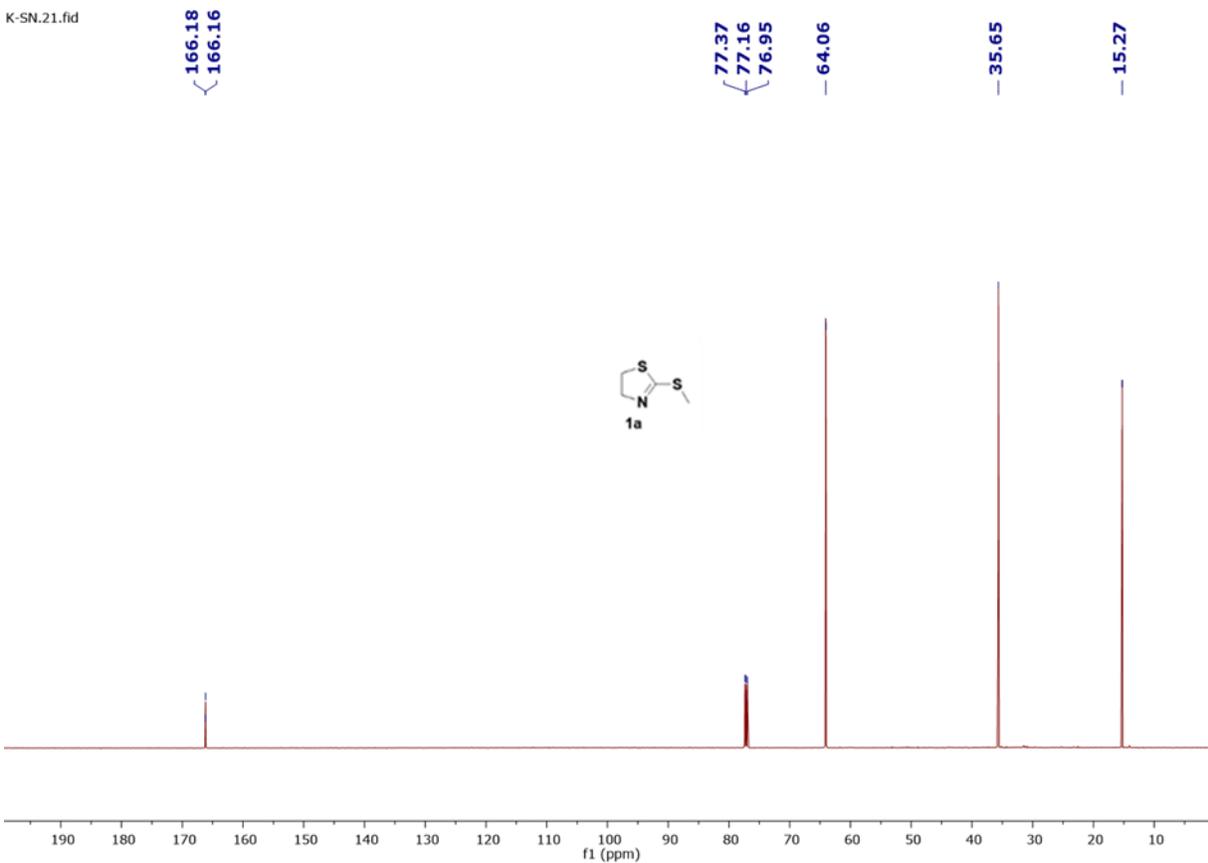
K-SN.20.fid

¹H NMR of 1a

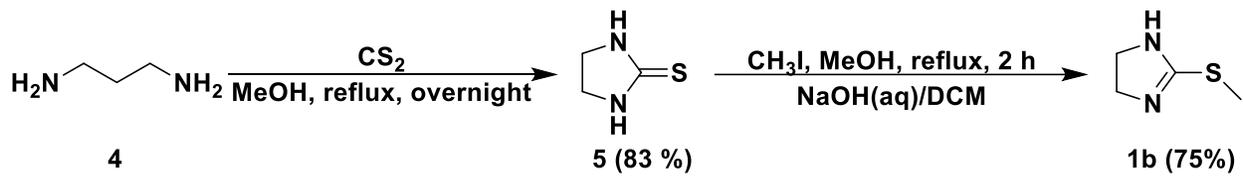


¹³C NMR of 1a

K-SN.21.fid



Synthesis of 1b



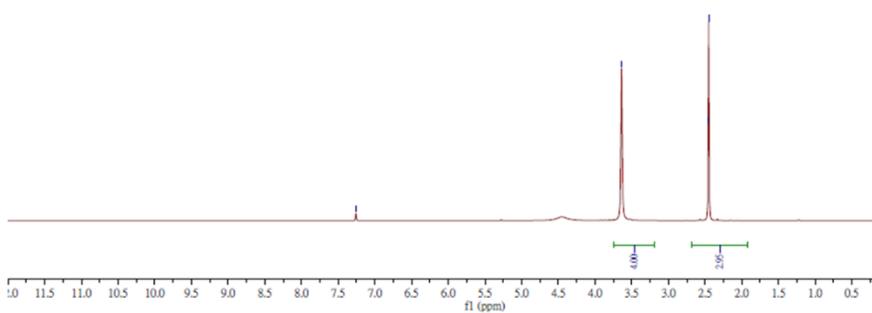
¹H NMR of 1b

K-NN10.fid

7.26

3.06

2.32



¹³C NMR of 1b

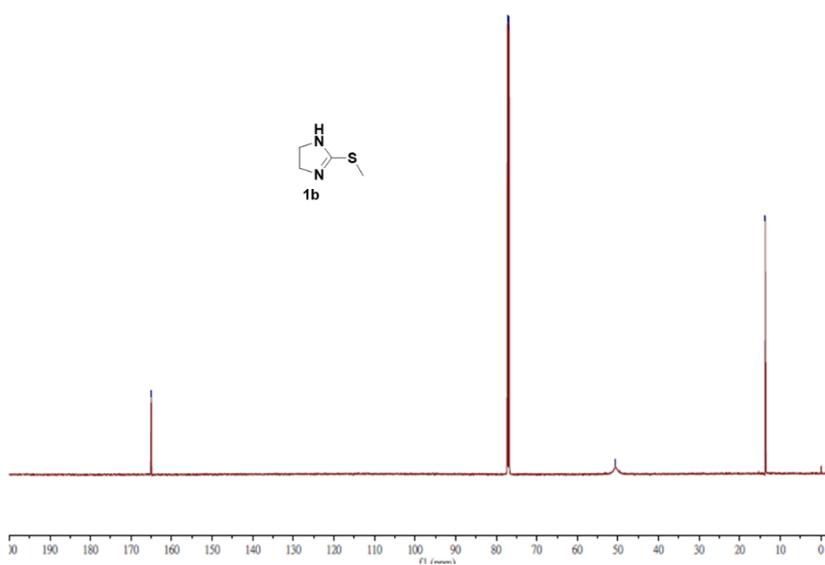
K-NN11.fid

164.59

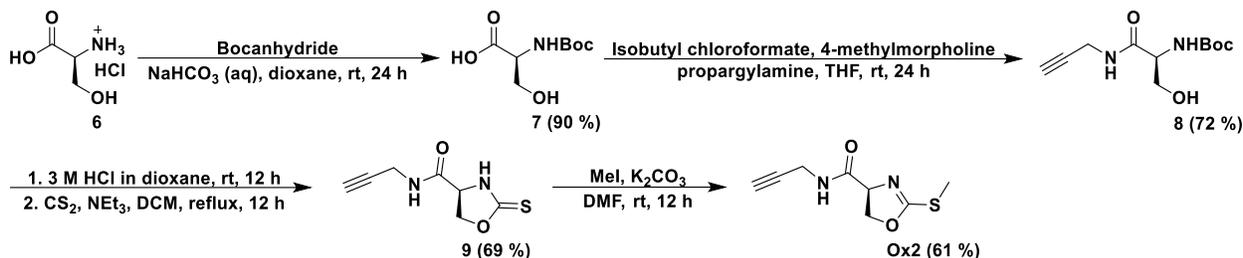
77.00

38.39

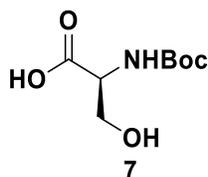
11.70



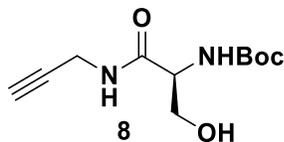
Synthesis of Ox2



The solution of L-serine **6** (3.5 g, 33.3 mmol) in saturated aqueous sodium bicarbonate solution (20 mL) was cooled to 0 °C. To this, Boc₂O (8.97 g, 40 mmol) in dioxane (20 mL) was added. The reaction mixture was warmed to room temperature and stirred for 24 h. The pH of aqueous layer was adjusted to 2 by 1M HCl (aq), followed by addition of brine. The solution was extracted with ethyl acetate. The organic layers were dried over magnesium sulfate. Filtration and concentration under reduced pressure afforded **7** as a colorless syrup (90%, 6.1g), which was used without further purification. ¹H and ¹³C NMR spectrum was consistent with that previously reported³.

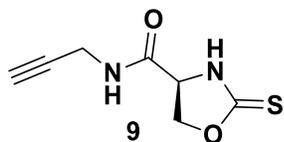


N-tert-butoxycarbonyl-L-serine **7** (1 g, 4.88 mmol), isobutylchloroformate (638 μL, 4.88 mmol), and 4-methylmorpholine (532 μL, 4.88 mmol) were added sequentially into dry THF (20 mL) and stirred at room temperature for 10 mins. Then, propargylamine (310 μL, 4.88 mmol) was added into the reaction mixture stirred at 25 °C for 24 h. The reaction mixture was concentrated under reduced pressure, then wash with ethyl acetate and brine. The organic layer was dried over anhydrous MgSO₄ and concentrated by rotary evaporation. The crude was purified by recrystallization from hexane and DCM to obtain compound **8** as white solid (72%, 850mg).



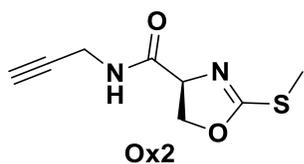
¹H NMR (600 MHz, CDCl₃): δ 7.03 (s, 1H), 5.59 (s, 1H), 4.17-4.00 (m, 4H), 3.66 (m, 1H), 3.10 (s, 1H), 2.23 (s, 1H), 1.46 (s, 3 H). **¹³C NMR** (150 MHz, d-DMSO): δ 170.42, 157.14, 156.26, 80.44, 80.00, 79.08, 71.59, 55.47, 42.12, 31.56, 29.08.

N-tert-butoxycarbonyl-D-serine-N'-propargylamide **8** (500 mg, 2.06 mmol) was treated with 3M HCl in dioxane 10 mL for 12 h. The reaction solution was removed under reduced pressure. The intermediate was used in next step without further purification. 2-amino-3-hydroxy-N-(prop-2-yn-1-yl) propanamide hydrochloride (2.06 mmol) was dissolved in 20 mL DCM, and triethylamine (344 μL, 2.47 mmol), CS₂ (150 μL, 2.47 mmol) were sequentially added. The reaction mixture was refluxed for 12 hours. The reaction solution was concentrated by rotary evaporation, and the residue was purified by column chromatography (hexane: ethyl acetate 1:1) to afford compound **9** as yellow powder (69%, 261 mg).



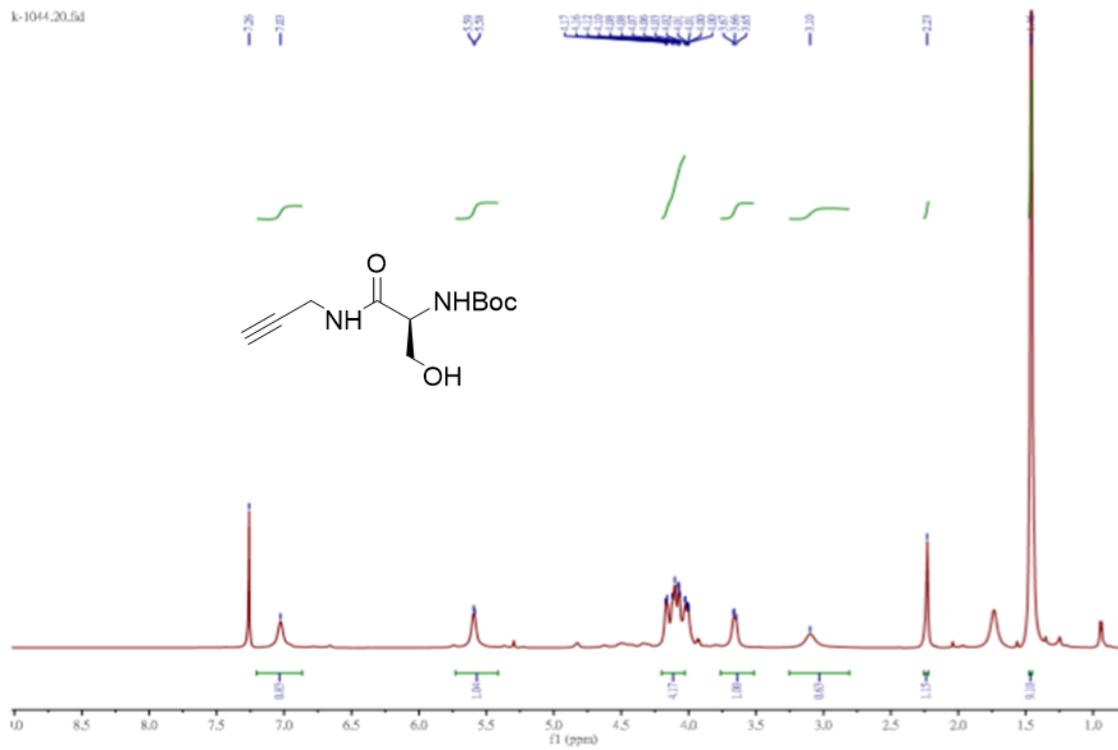
¹H NMR (600 MHz, d-DMSO): δ 10.24 (s, 1H), 8.69 (t, *J* = 5.4 Hz 1H), 4.73 (t, *J* = 9.2 Hz 1H), 4.52-4.45 (m, 2H), 3.92-3.89 (m, 2H), 3.32 (s, 1H), 3.19 (s, 1H). **¹³C NMR** (150 MHz, d-DMSO): δ 188.78, 168.47, 80.31, 73.63, 72.65, 57.66, 28.31.

To a solution of (S)-N-(prop-2-yn-1-yl)-2-thioxooxazolidine-4-carboxamide **9** (200 mg, 1.08 mmol) in DMF (10 mL), K₂CO₃ (179 mg, 1.30 mmol) and methyl iodide (80.6 μL, 1.30 mmol) were added sequentially. The reaction mixture was allowed to stir at room temperature for 12 hours. The reaction solution was washed with ethyl acetate and brine for 3 times. The residue was concentrated under reduced pressure, and purified by the column chromatography (hexane: ethyl acetate 5:1) to obtain compound **Ox2** as yellow oil (130 mg, 61%).



¹H NMR (600 MHz, CDCl₃): δ 6.81 (s, 1H), 4.70-4.53 (m, 4H), 4.12-4.05 (m, 2H), 2.50 (s, 3H), 2.26 (m, 1H). **¹³C NMR** (150 MHz, CDCl₃): δ 171.14, 170.23, 79.17, 72.26, 71.93, 68.67, 29.04, 68.67.

¹H NMR spectra of 8



¹³C NMR spectrum of 8

K-494.11.f.d

170.42

155.24
155.06

81.44
81.00
77.06
76.62
76.18
75.74
75.30
74.86

55.87

43.22

32.95

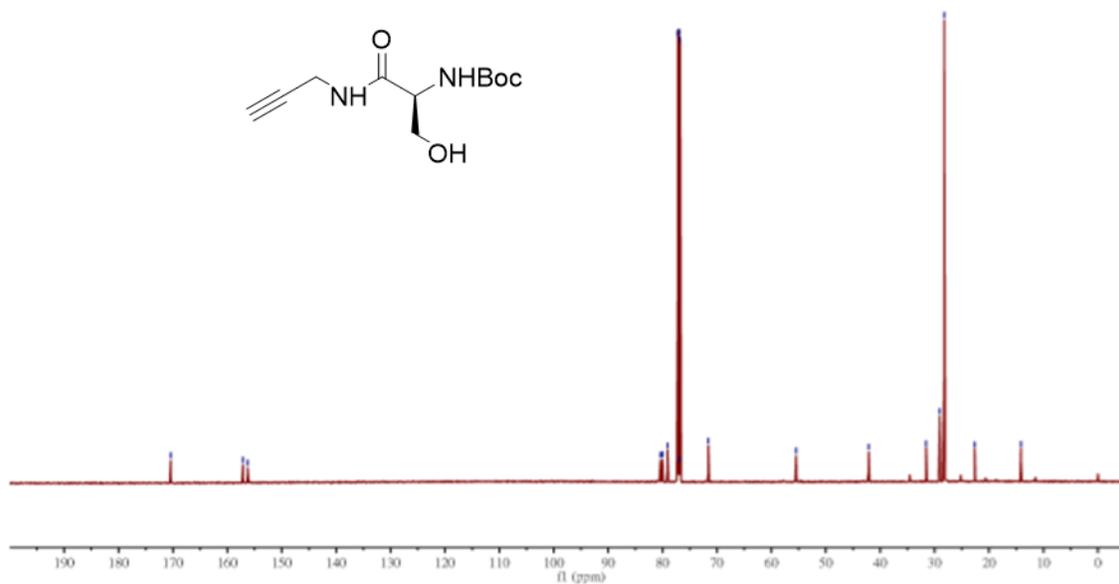
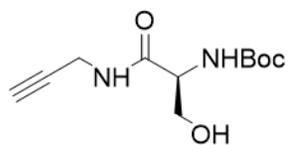
32.51

32.07

31.63

23.84

14.23



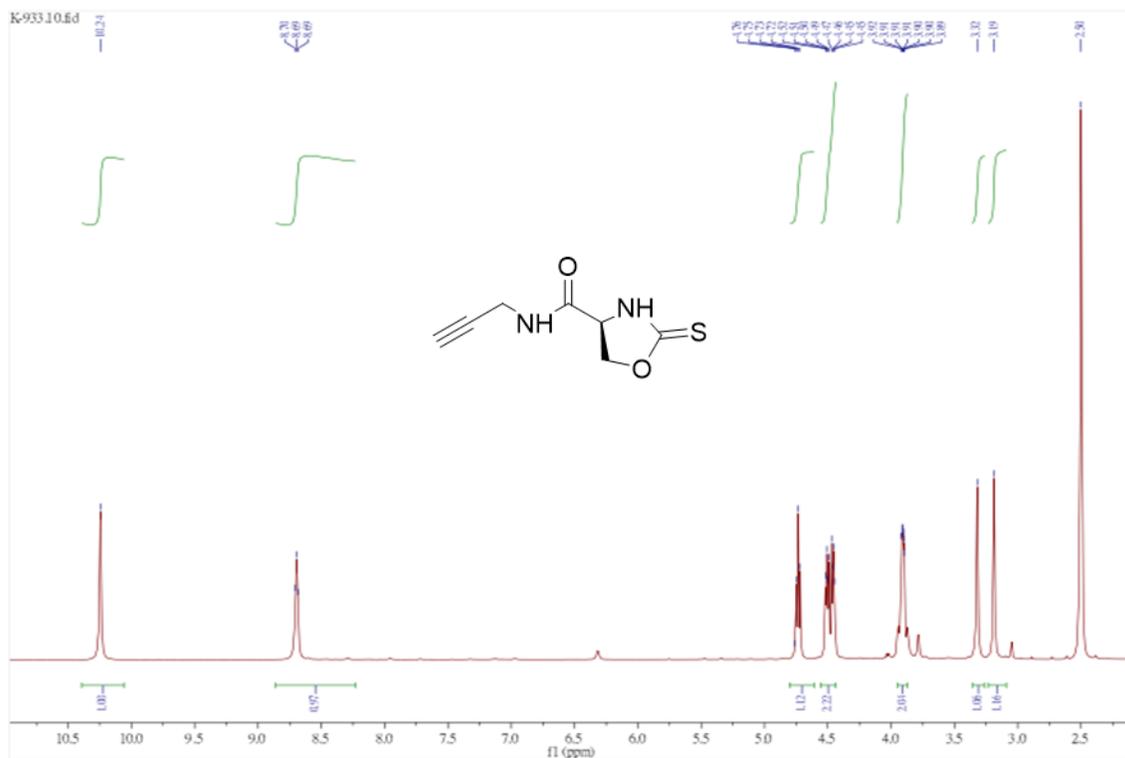
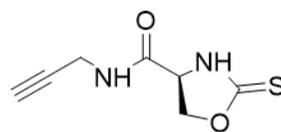
¹H NMR spectra of 9

K-933.10.f.d

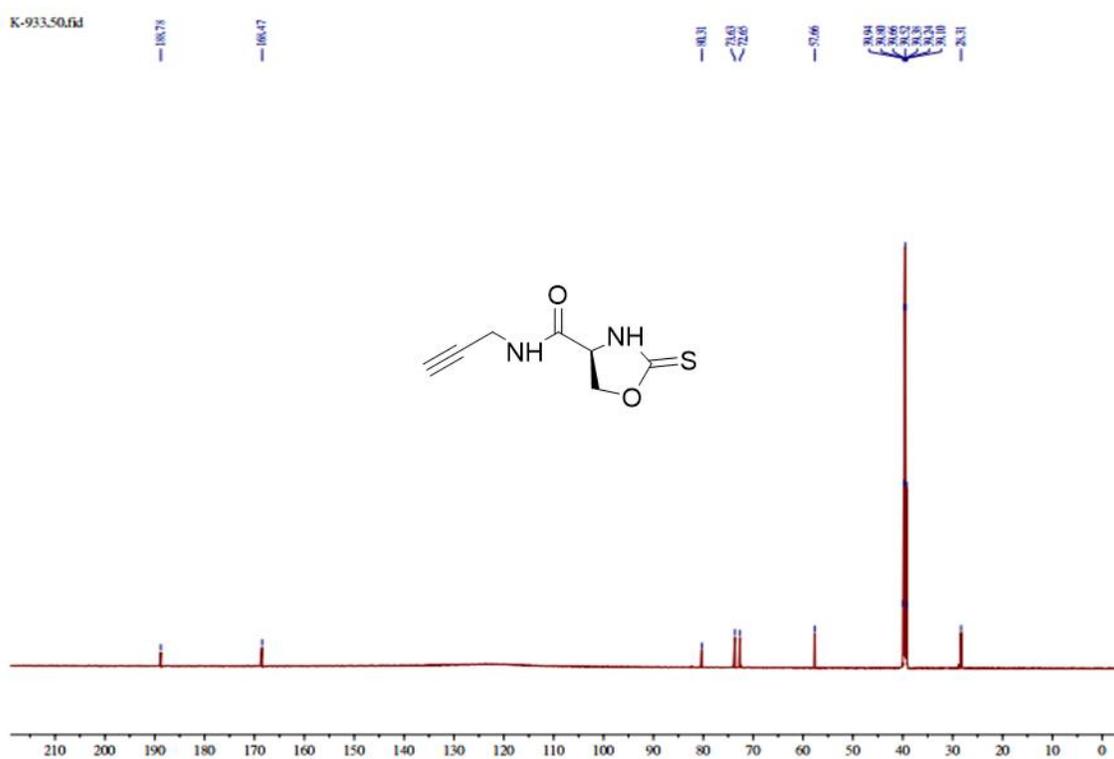
10.24

8.71
8.69

4.72
4.70
4.68
4.66
4.64
4.62
4.60
4.58
4.56
4.54
4.52
4.50
4.48
4.46
4.44
4.42
4.40
4.38
4.36
4.34
4.32
4.30
4.28
4.26
4.24
4.22
4.20
4.18
4.16
4.14
4.12
4.10
4.08
4.06
4.04
4.02
4.00
3.98
3.96
3.94
3.92
3.90
3.88
3.86
3.84
3.82
3.80
3.78
3.76
3.74
3.72
3.70
3.68
3.66
3.64
3.62
3.60
3.58
3.56
3.54
3.52
3.50
3.48
3.46
3.44
3.42
3.40
3.38
3.36
3.34
3.32
3.30
3.28
3.26
3.24
3.22
3.20
3.18
3.16
3.14
3.12
3.10
3.08
3.06
3.04
3.02
3.00
2.98
2.96
2.94
2.92
2.90
2.88
2.86
2.84
2.82
2.80
2.78
2.76
2.74
2.72
2.70
2.68
2.66
2.64
2.62
2.60
2.58
2.56
2.54
2.52
2.50
2.48
2.46
2.44
2.42
2.40
2.38
2.36
2.34
2.32
2.30
2.28
2.26
2.24
2.22
2.20
2.18
2.16
2.14
2.12
2.10
2.08
2.06
2.04
2.02
2.00
1.98
1.96
1.94
1.92
1.90
1.88
1.86
1.84
1.82
1.80
1.78
1.76
1.74
1.72
1.70
1.68
1.66
1.64
1.62
1.60
1.58
1.56
1.54
1.52
1.50
1.48
1.46
1.44
1.42
1.40
1.38
1.36
1.34
1.32
1.30
1.28
1.26
1.24
1.22
1.20
1.18
1.16
1.14
1.12
1.10
1.08
1.06
1.04
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0.72
0.70
0.68
0.66
0.64
0.62
0.60
0.58
0.56
0.54
0.52
0.50
0.48
0.46
0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00

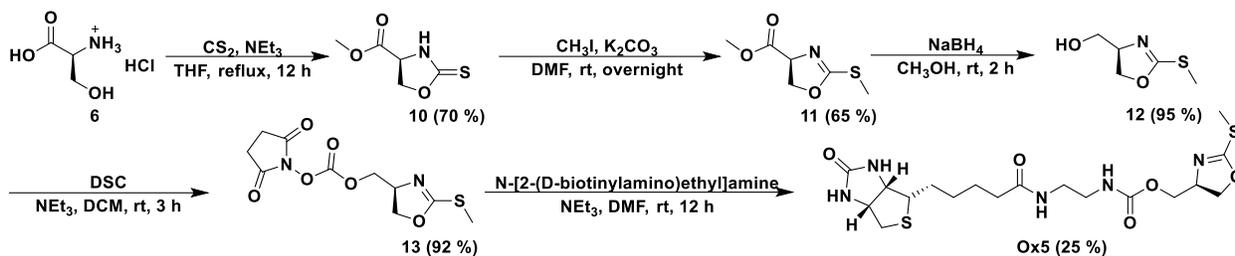


¹³C NMR spectrum of 9

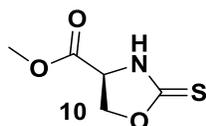


¹H NMR spectrum of Ox2

Synthesis of Ox5



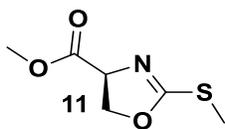
To a suspension of L-serine methyl ester hydrochloride **6** (6.5 g, 41.8 mmol) in 60 mL of tetrahydrofuran at 0 °C was added triethylamine (5.8 mL, 41.8 mmol), followed by the addition of 25 mL of carbon disulfide. The reaction mixture was refluxed for 12 h. Dichloromethane was added, and the organic solution was washed with aqueous sodium bicarbonate, dried over MgSO₄, filtered, and concentrated to provide crude as an orange-yellow oil. Column chromatography on silica gel eluting with ethyl acetate-hexanes (1:2) provided compound **10** as orange oil (4.8 g, 70%).



¹H NMR (600 MHz, CDCl₃): δ 7.93 (s, 1H), 4.85 (m, 2H), 4.65 (dd, *J* = 6.2 3.6 Hz 1H), 3.83 (s, 3H). ¹³C NMR (150 MHz, CDCl₃): δ 189.95, 168.81, 72.09, 57.11, 53.41.

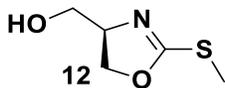
To a mixture of methyl (S)-2-thioxooxazolidine-4-carboxylate **10** (2 g, 12.4 mmol) and K₂CO₃ (2 g, 14.9 mmol) in dry DMF, CH₃I (924 μL, 14.9 mmol) was added. The reaction mixture was allowed to stir at room temperature for 14 h. The reaction solution was washed with ethyl acetate and brine for 3 times. The organic layer was collected, dried over anhydrous MgSO₄, filtered,

and concentrated under the reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 3:1) to yield compound **11** as colorless oil (1.4 g, 65%).



¹H NMR (600 MHz, CDCl₃): δ 4.58 (m, 1H), 4.42-4.51 (m, 2H), 3.68 (s, 3H), 2.38 (s, 3H). **¹³C NMR** (150 MHz, CDCl₃): δ 171.08, 169.51, 71.22, 67.95, 52.52, 52.50, 14.37.

To a solution of methyl (S)-2-(methylthio)-4,5-dihydrooxazole-4-carboxylate **11** (1.5 g, 8.6 mmol) in dried methanol (30 mL), NaBH₄ (976 mg, 25.8 mmol) was added over a period of 1 h under ice bath. The reaction solution was allowed to warm to room temperature and stir for 3 hours. The reaction mixture was quenched by addition of water, then washed with ethyl acetate and brine three times. The organic portion was collected, dried over anhydrous magnesium sulfate, filtered and concentrated by rotary evaporation to obtain compound **12** as colorless oil (1.2 g, 95%). Which was used without further purification.

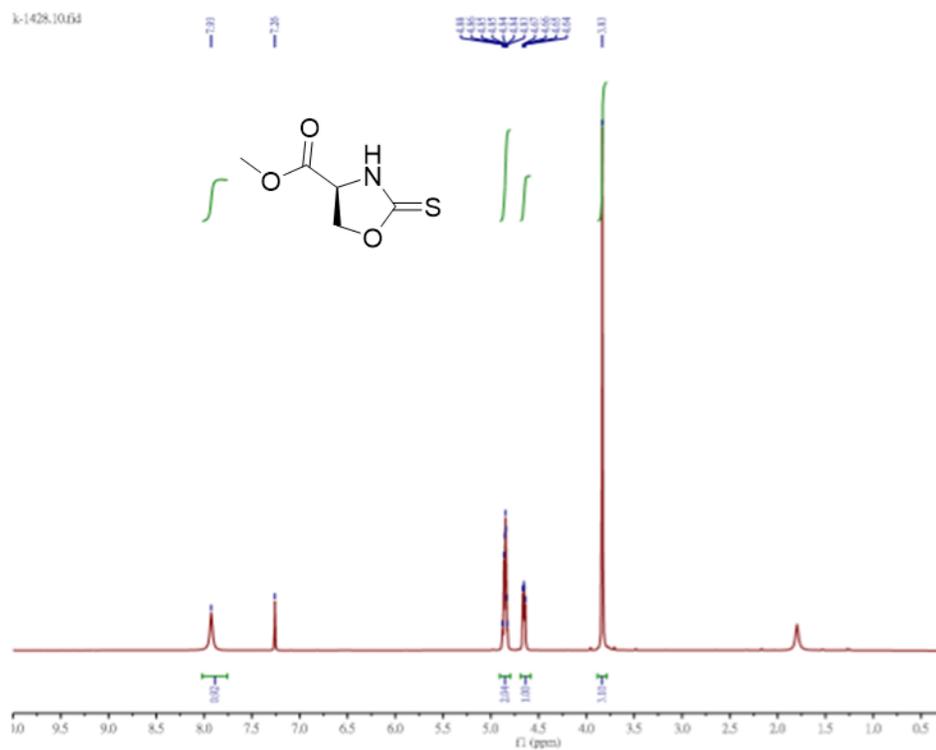


¹H NMR (600 MHz, CDCl₃): δ 4.42 (t, *J* = 6.9 Hz, 1H), 4.25 (m, 2H), 3.57 (d, *J* = 10.5 Hz, 1H), 3.55 (t, *J* = 10.5 Hz, 1H), 2.45 (s, 3H). **¹³C NMR** (150 MHz, CDCl₃): δ 171.30, 168.46, 71.26, 68.15, 63.91, 60.52, 21.6, 14.3.

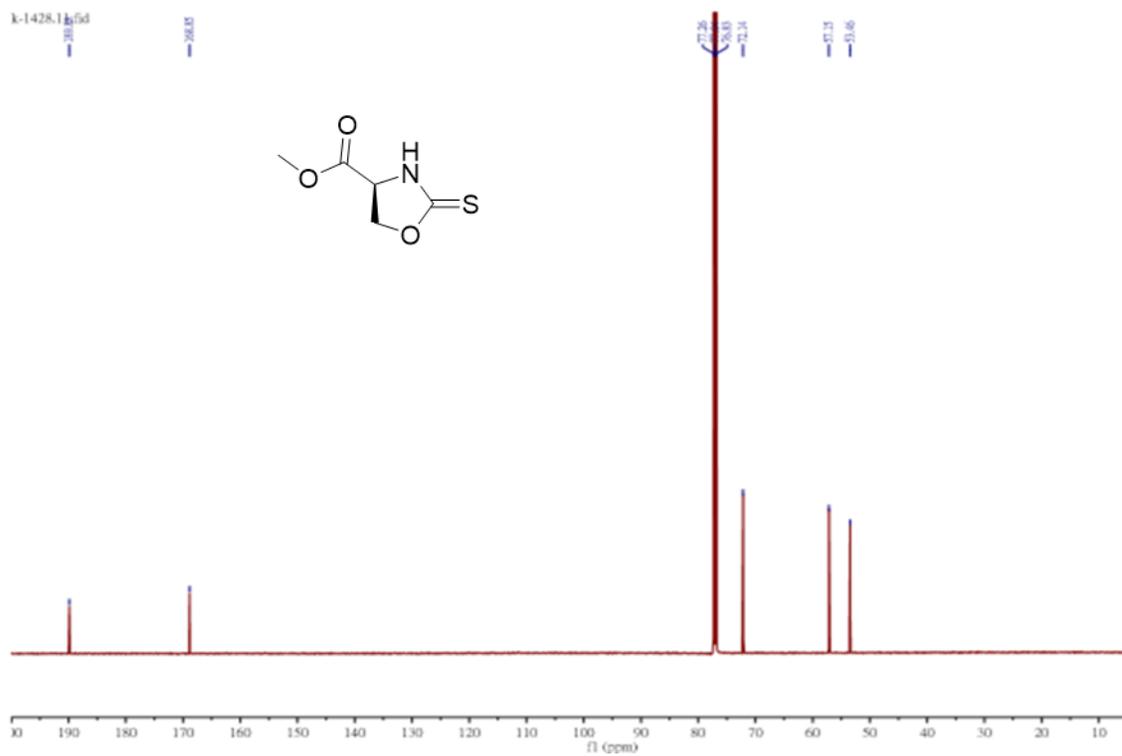
(R)-(2-(methylthio)-4,5-dihydrooxazol-4-yl) methanol **12** (142 mg, 0.97 mmol) was dissolved in 10 mL dry DCM. To the solution, triethylamine (201 μ L, 1.45 mmol) and disuccinimidyl carbonate (DSC, 371 mg, 1.45 mmol) was added. The reaction was stirred for 3h at room temperature and extracted with DCM and aqueous sodium bicarbonate solution. The organic part was collected and dried over anhydrous MgSO_4 , filtered, and concentrated to obtain compound **13** as colorless oil (257 mg, 92%). The NHS ester derivative **13** was used directly without further purification.

A mixture of **13** (140 mg, 0.49 mmol) and triethylamine (67 μ L, 0.49 mmol) in 5 mL of DMF was added to biotinylethylenediamine (139 mg, 0.49 mmol) dissolved in 5 mL of DMF. The solution was stirred under nitrogen at room temperature for 12 h, after which the solution was evaporated to dryness under reduced pressure to give a yellow oil. The crude was purified by column chromatography (MeOH: ethyl acetate 1: 50 to 1:7) to afford compound **Ox5** as colorless oil (56 mg, 25%).

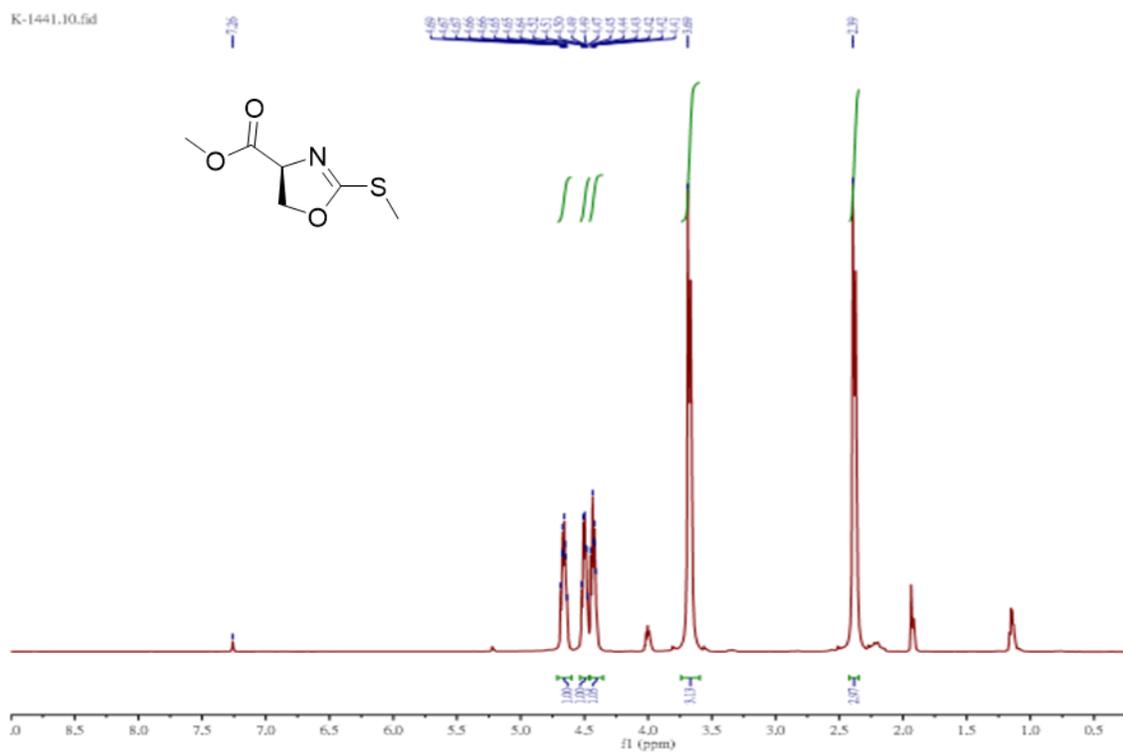
^1H NMR spectra of 10



^{13}C NMR spectrum of 10

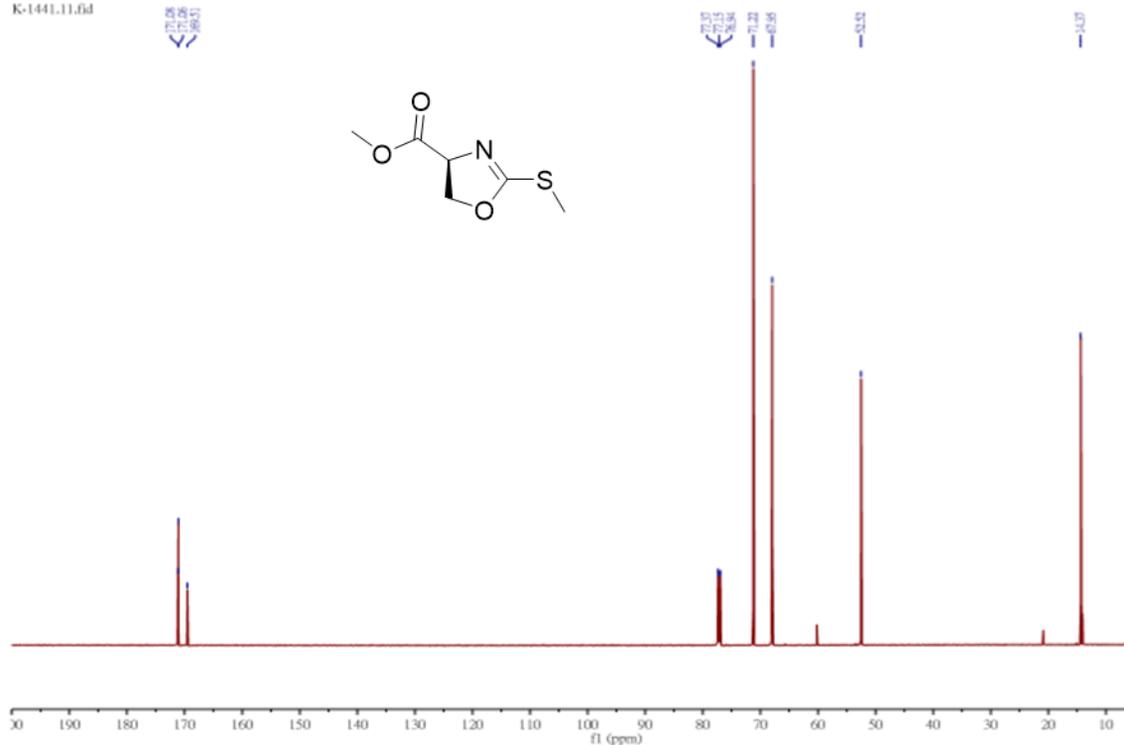


^1H NMR spectra of 11



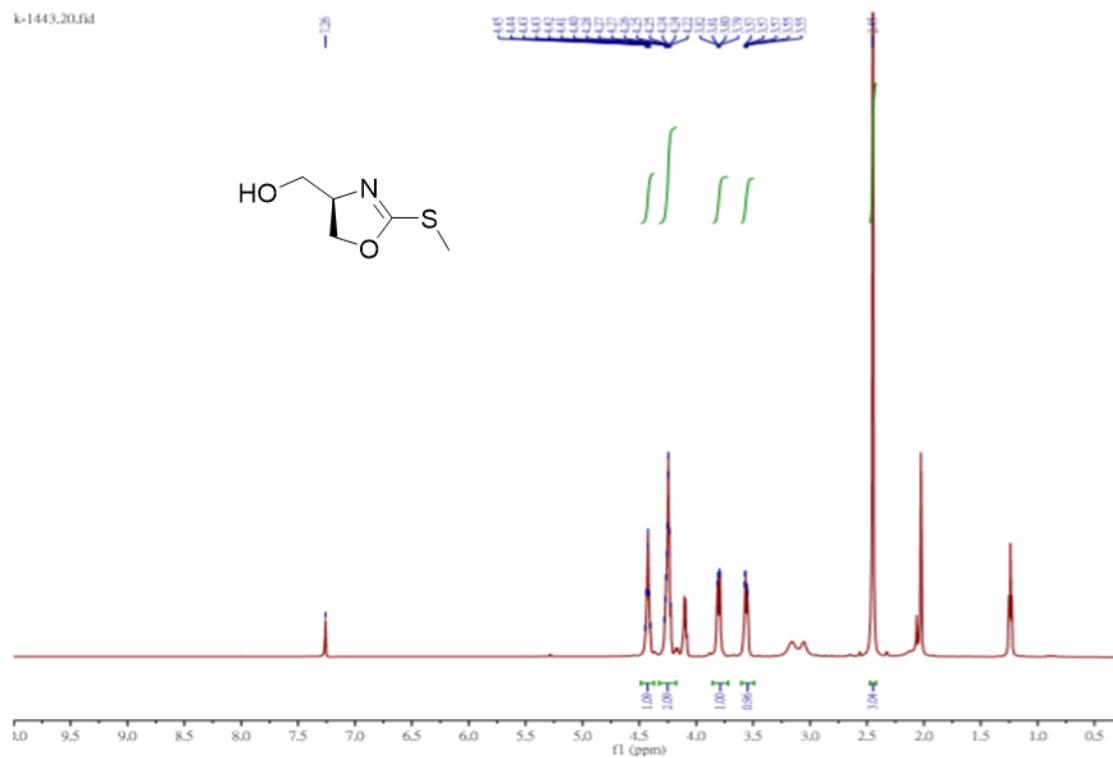
^{13}C NMR spectrum of 11

K-1441.11.fid

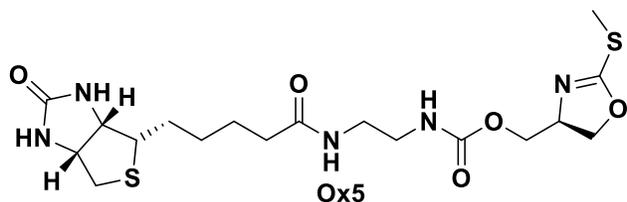
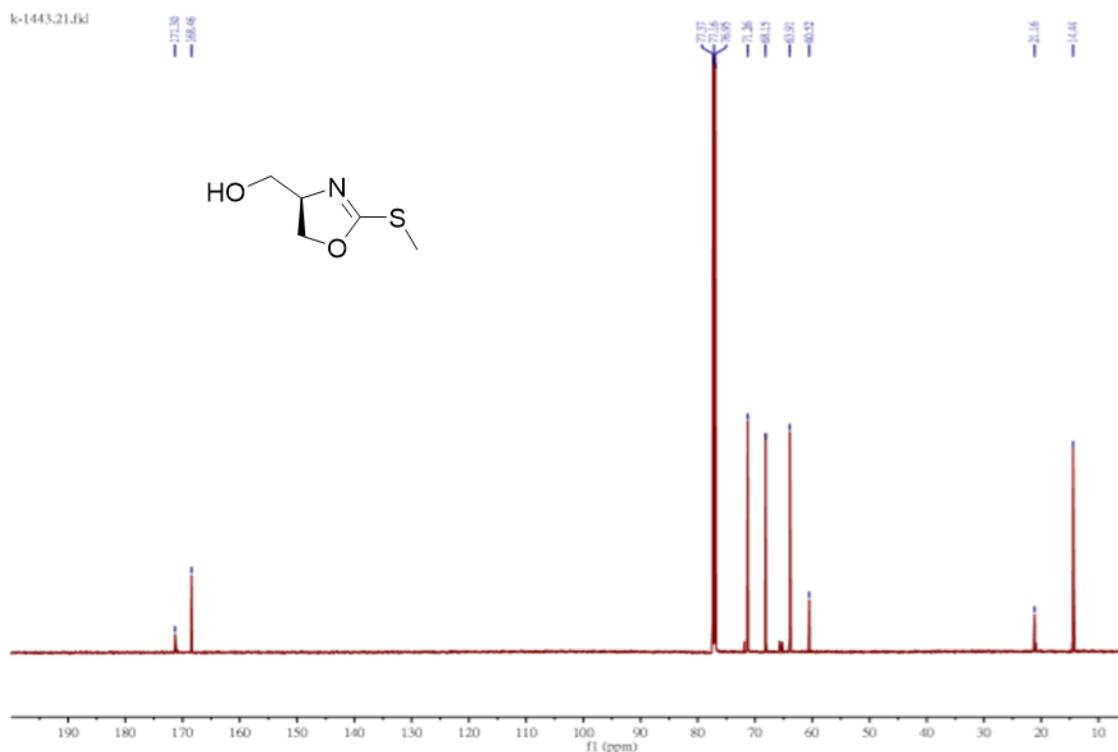


¹H NMR spectra of 12

k-1443.20.fid

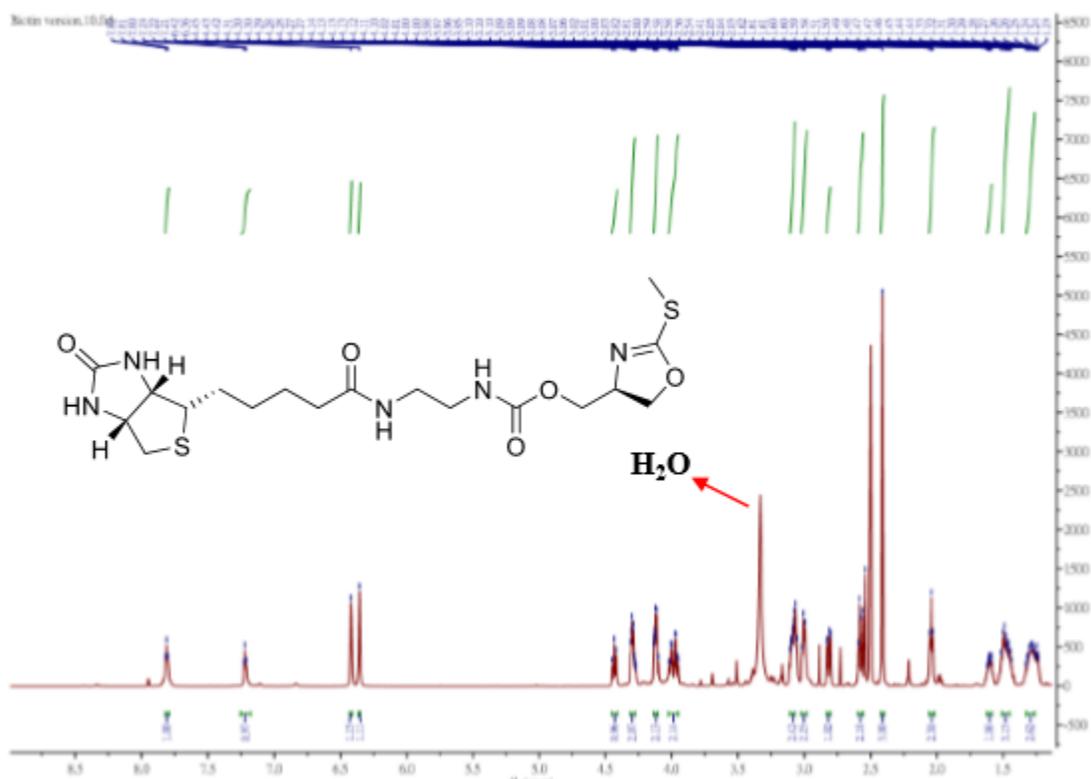


¹³C NMR spectrum 12

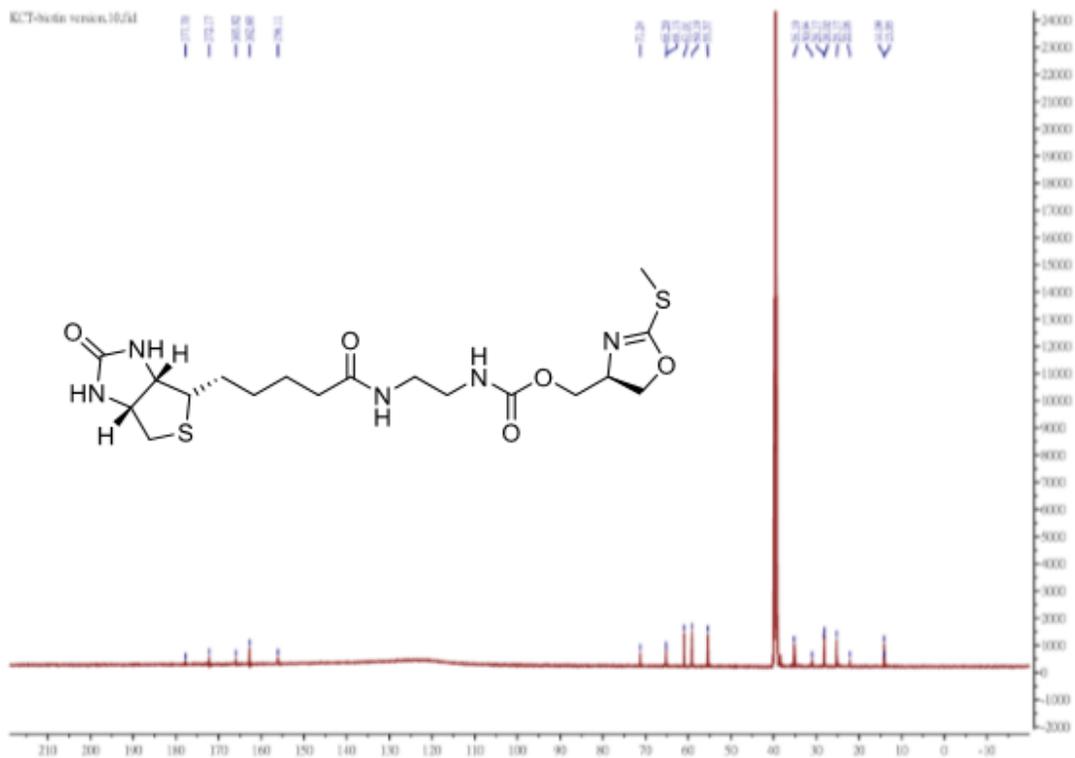


¹H NMR (600 MHz, DMSO): 7.80 (s, 1H), 7.22 (s, 1H), 6.42 (s, 1H), 6.36 (s, 1H), 4.43 (t, *J* = 8.6 Hz, 1H), 4.31-4.29 (m, 2H), 4.13-4.11, 4.13-4.11 (m, 2H), 4.02-3.95 (m, *J* = 2H), 3.10-3.06 (m, 2H), 3.02-2.98 (m, 2H), 2.81 (dd, *J* = 7.3 5.1 Hz, 1H), 2.59-2.56 (m, 2H), 2.41 (s, 3H), 2.04 (t, *J* = 7.4 Hz, 2H), 2.36 (s, 3H) 1.63-1.44 (m, 4H) 1.33-1.23 (m, 2H). ¹³C NMR (150 MHz, DMSO): δ 177.70, 172.17, 165.92, 162.68, 156.11, 71.24, 65.20, 65.13, 61.01, 59.19, 55.37, 35.19, 30.94, 28.17, 28.02, 25.17, 22.06, 14.09.

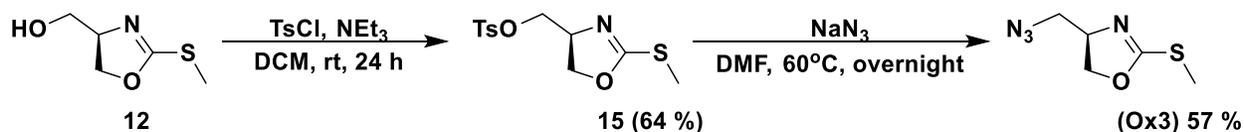
¹H NMR spectra of Ox5



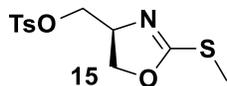
¹³C NMR spectrum of Ox5



Synthesis of Ox3

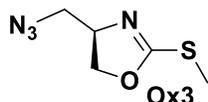


To (R)-(2-(methylthio)-4,5-dihydrooxazol-4-yl)methanol **12** (1g, 6.8 mmol) in triethylamine (1.9 mL, 13.6 mmol) p-toluenesulfonyl chloride (1.95 g, 10.2 mmol) was added at 0 °C. The reaction mixture was stirred at room temperature for 24 hours. Then, the reaction mixture was washed with aqueous sodium bicarbonate solution and extracted with DCM for 3 times. The combined organic phase was dried over anhydrous MgSO₄. The solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (hexanes: ethyl acetate 3:1) to afford compound **15** as yellow oil (1.3g, 64%).



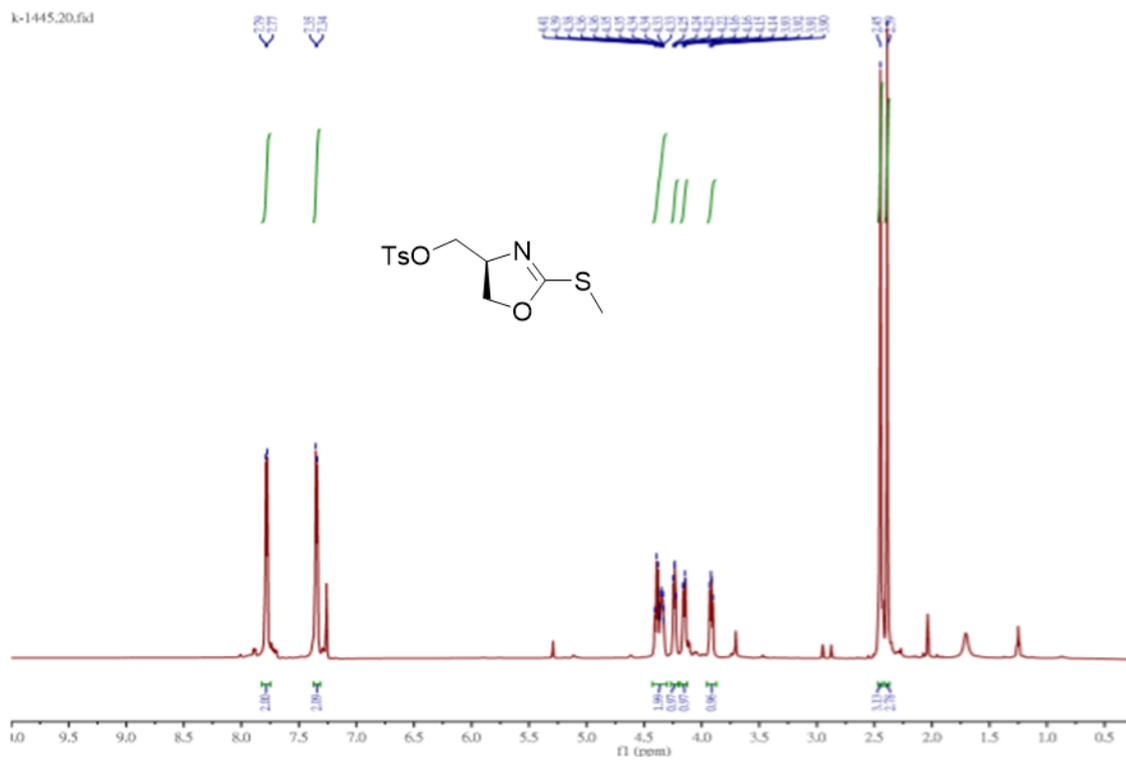
¹H NMR (600 MHz, CDCl₃): δ 7.78 (d, *J* = 8.0 Hz, 2H), 7.34 (d, *J* = 8.0 Hz, 2H), 4.41-4.33 (m, 2H), 4.24 (t, *J* = 7.0 Hz, 1H), 4.16 (dd, *J* = 10.3 3.6 Hz, 1H), 3.92 (t, *J* = 8.5 Hz, 1H), 2.45 (s, 3H), 2.39 (s, 3H). **¹³C NMR** (150 MHz, CDCl₃): δ 169.39, 145.24, 132.67, 130.08, 128.12, 71.82, 70.056, 64.96, 21.79, 14.55.

To a solution of (S)-2-(2-(methylthio)-4,5-dihydrooxazol-4-yl)methyl 4-methylbenzenesulfonate **15** (580 mg, 3.2 mmol) in dry DMF (20 mL), sodium azide (1.3 g, 32.2 mmol) was added at room temperature. The reaction mixture was stirred at 40 °C for 24 h. Then, the reaction was quenched by addition of water and extracted with ethyl acetate. The combined organic phase was washed with brine and dried over anhydrous MgSO₄. The yellow oil was purified by column chromatography on silica gel (hexanes: ethyl acetate 5:1) to afford compound **Ox3** as colorless oil (313 mg, 57%).

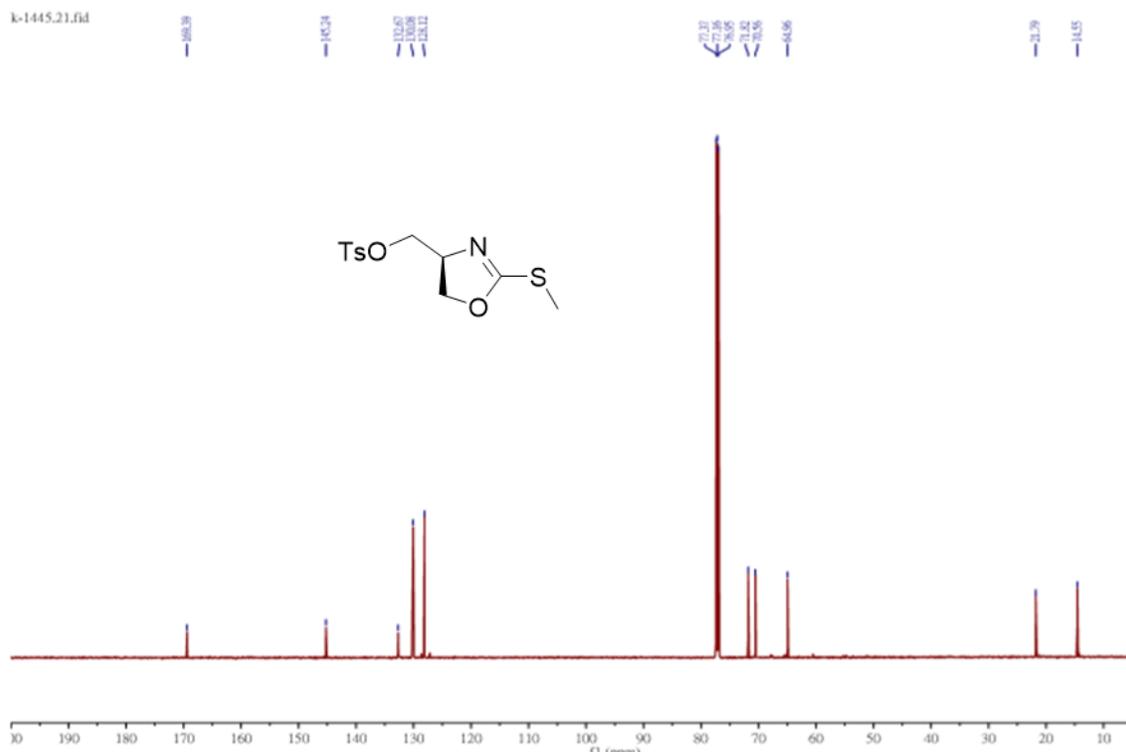


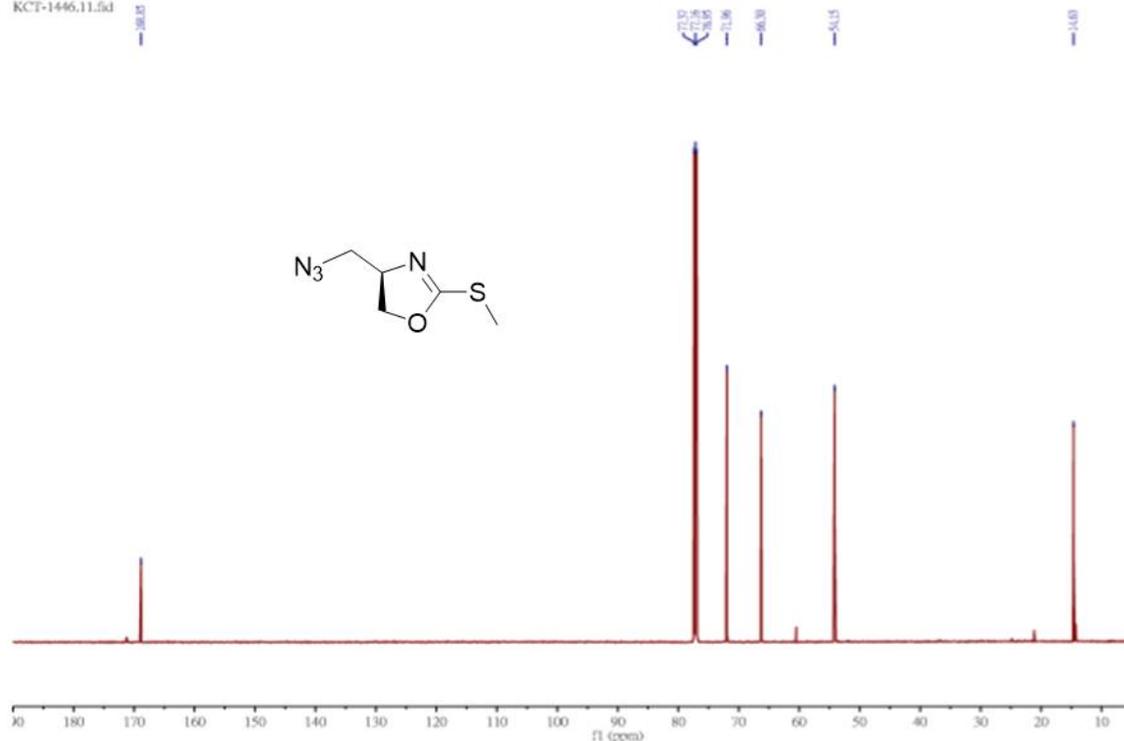
¹H NMR (600 MHz, CDCl₃): δ 4.41 (t, *J* = 8.9 Hz, 1H), 4.36-4.31 (m, 1H), 3.46 (dd, *J* = 12.7 5.0 Hz, 1H), 3.31 (dd, *J* = 12.7 5.0 Hz, 1H), 2.46 (s, 3H). **¹³C NMR** (150 MHz, CDCl₃): δ 168.85, 71.96, 66.30, 54.15, 14.6

¹H NMR spectra of 15

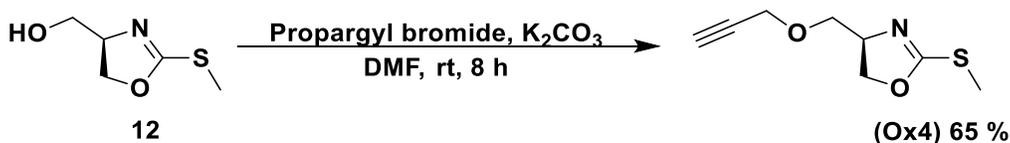


¹³C NMR spectrum of 15



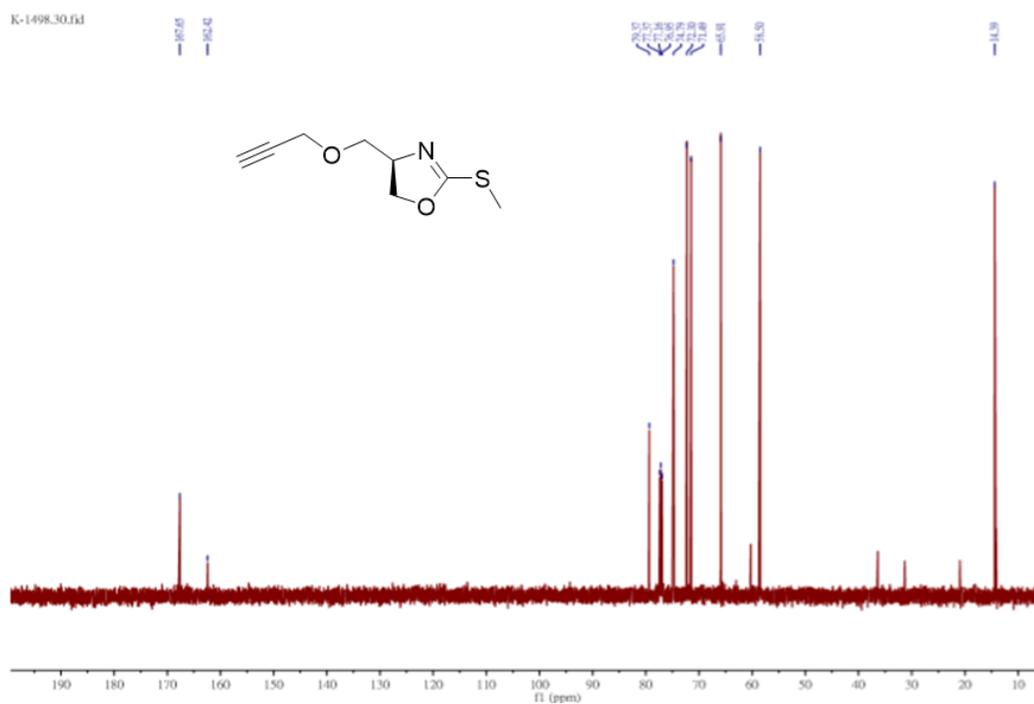


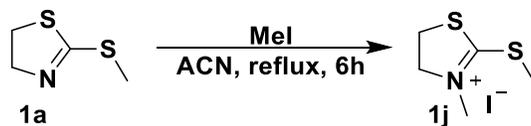
Synthesis of Ox4



To a solution of (R)-2-(2-(methylthio)-4,5-dihydrooxazol-4-yl) methanol **12** (200 mg, 1.36 mmol), in 5 mL of dry DMF at 0 °C, NaH (40 mg, 1.36 mmol) was added. After 10 mins propargyl bromide (123 μL , 1.63 mmol) was added drop wise and the reaction mixture was allowed to warm to room temperature. After 8 h reaction mixture was washed with ethyl acetate and brine. The organic part was dried over anhydrous MgSO_4 , filtered, and concentrated. The crude was purified by flash column (hexane: ethyl acetate 3:1) to afford compound **Ox4** as yellow oil (163 mg, 65%).

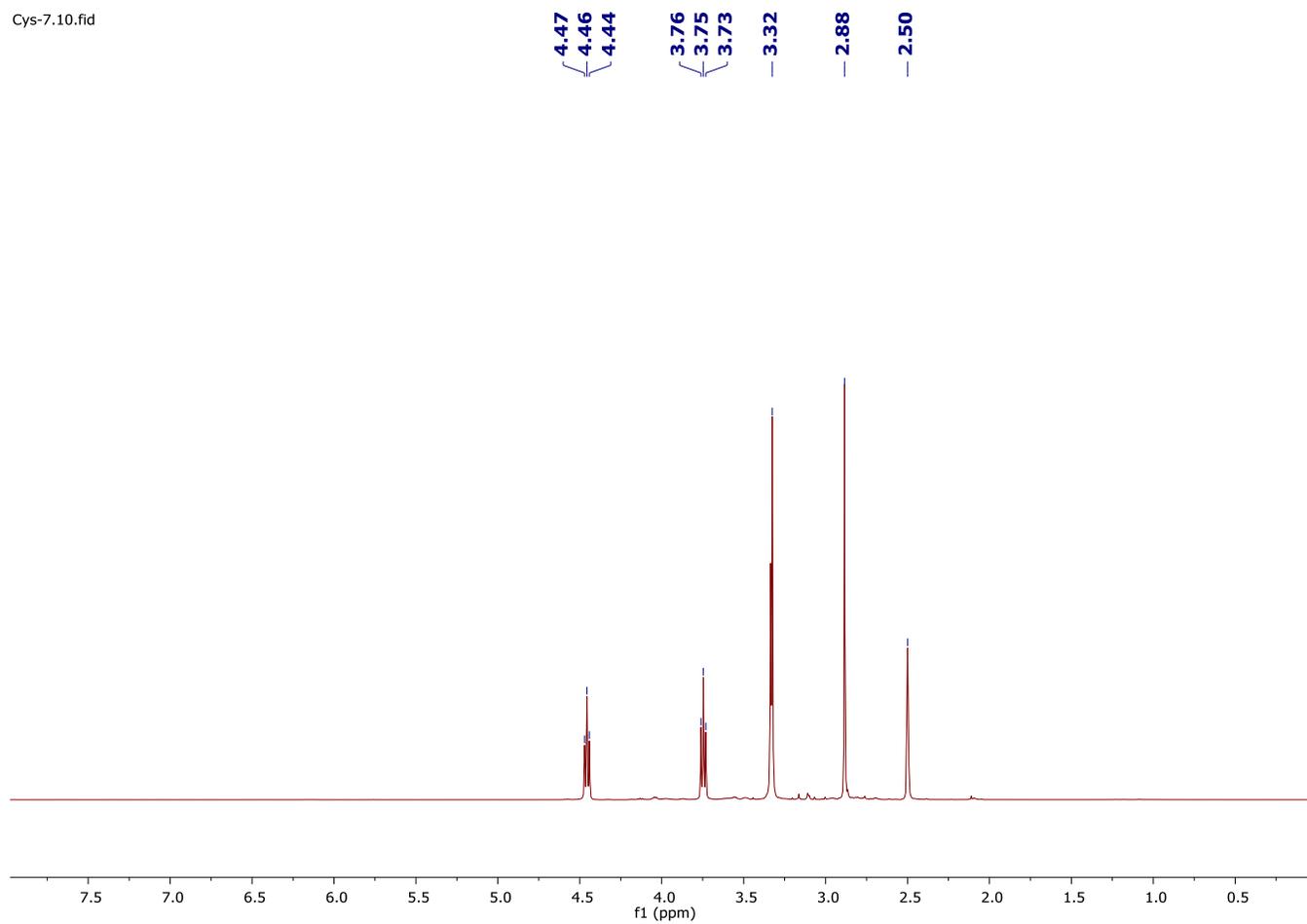
¹³C NMR spectrum of Ox4



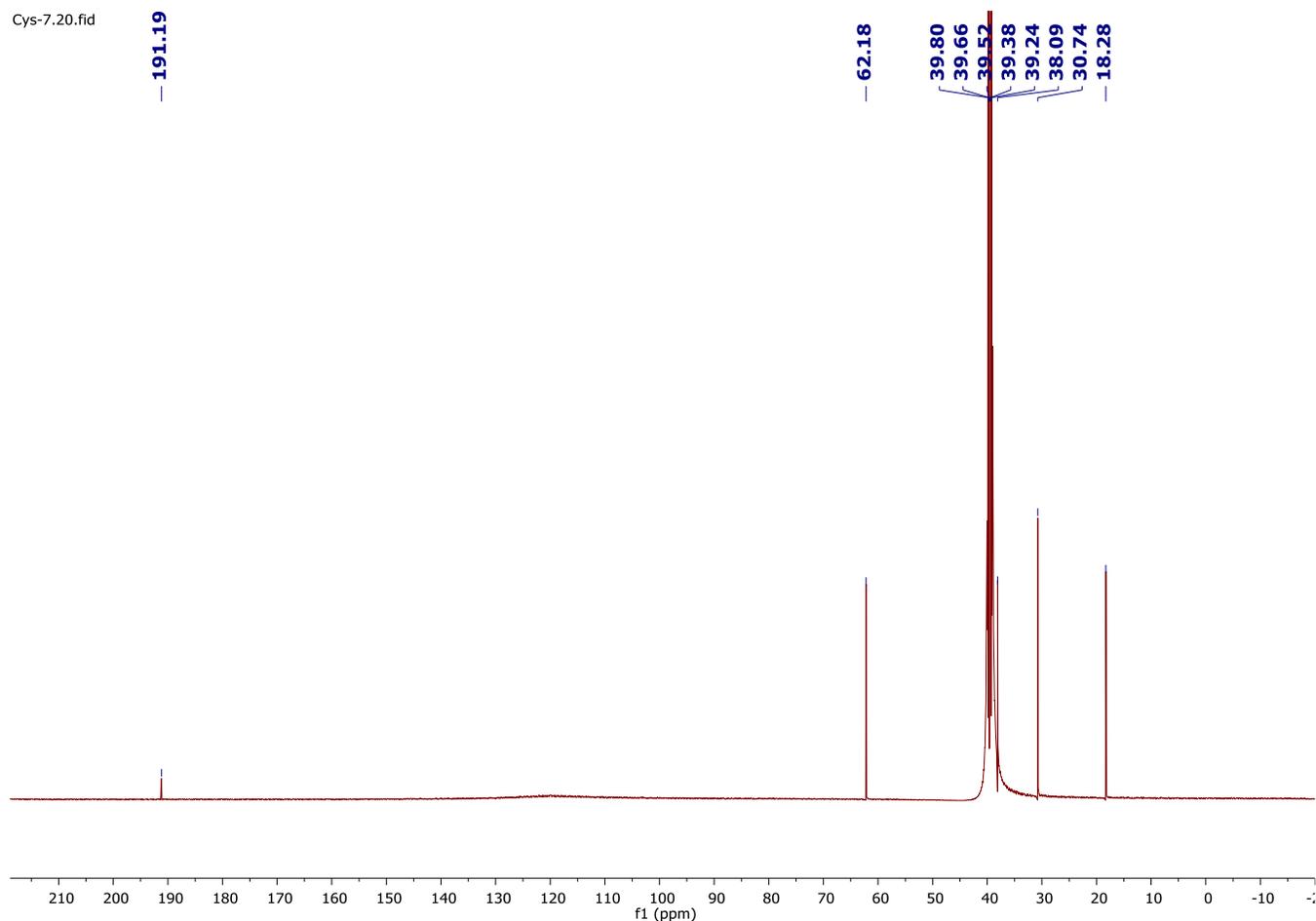


To a solution of 2-(methylthio)-4,5-dihydrothiazole **1a** (300 mg, 2.20 mmol) in ACN (5 mL), methyl iodide (137 μL , 2.20 mmol) was added. The reaction mixture was refluxed for 6 hours. The reaction solution was cooled down to room temperature and concentrated by rotary evaporation to obtain the yellow powder. The crude powder was purified by recrystallization with ethanol / hexane (1:10) to afford pure **1j** as yellow powder (91 %, 295 mg). **^1H NMR** (600 MHz, $\text{DMSO-}d_6$) δ 4.47 (t, $J = 8.8$ Hz, 2H), 3.76 (t, $J = 8.8$ Hz, 2H), 3.34 (s, 3H), 2.90 (s, 3H). **^{13}C NMR** (150 MHz, $\text{DMSO-}d_6$): δ 191.72, 62.70, 38.62, 31.27, 18.81.

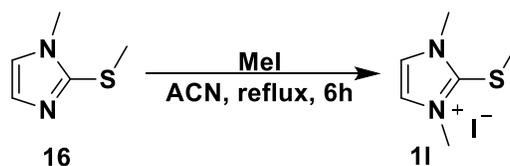
^1H NMR of 1j



¹³C NMR of 1j



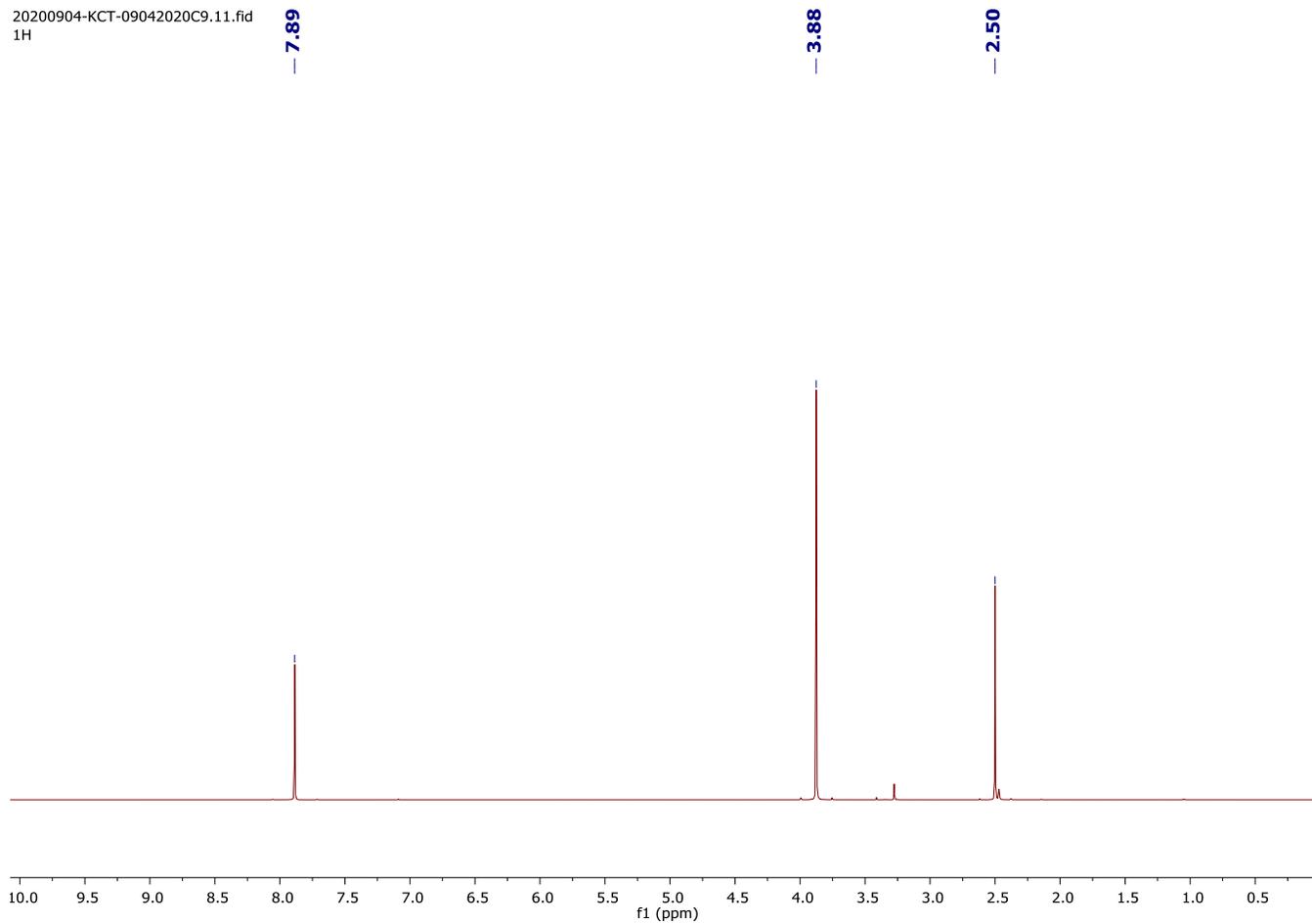
Synthesis of compound 1I



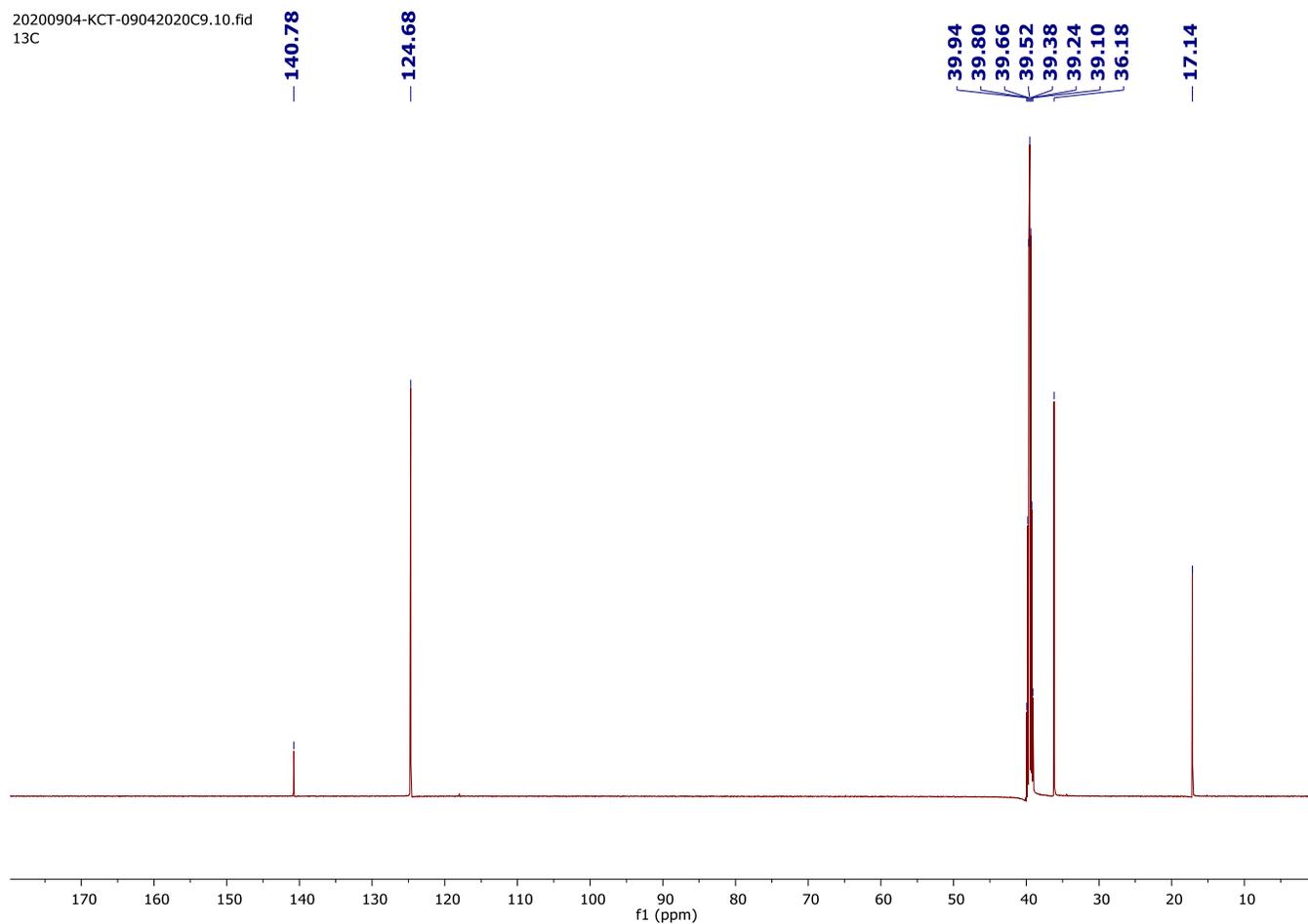
To a solution of 1-Methyl-2-(methylthio)imidazole **16** (200 mg, 1.56 mmol) in dry acetonitrile (5 mL), CH₃I (486 μ L, 7.8 mmol) was added. The reaction mixture was refluxed for 6 hours. The reaction solution was cooled to room temperature and concentrated by rotary evaporation to afford the white solid. The crude was purified by recrystallization with ethanol / hexane (1:10) to afford pure **1I** as white powder (95 %, 211.9 mg). ¹H NMR (600 MHz, DMSO-*d*₆) δ 7.89 (s, 2H), 3.88 (s, 6H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 140.78, 124.68, 36.18, 17.14.

¹H NMR spectra of 1l

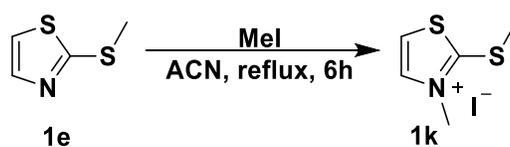
20200904-KCT-09042020C9.11.fid
1H



¹³C NMR spectra of 1l



Synthesis of compound 1k

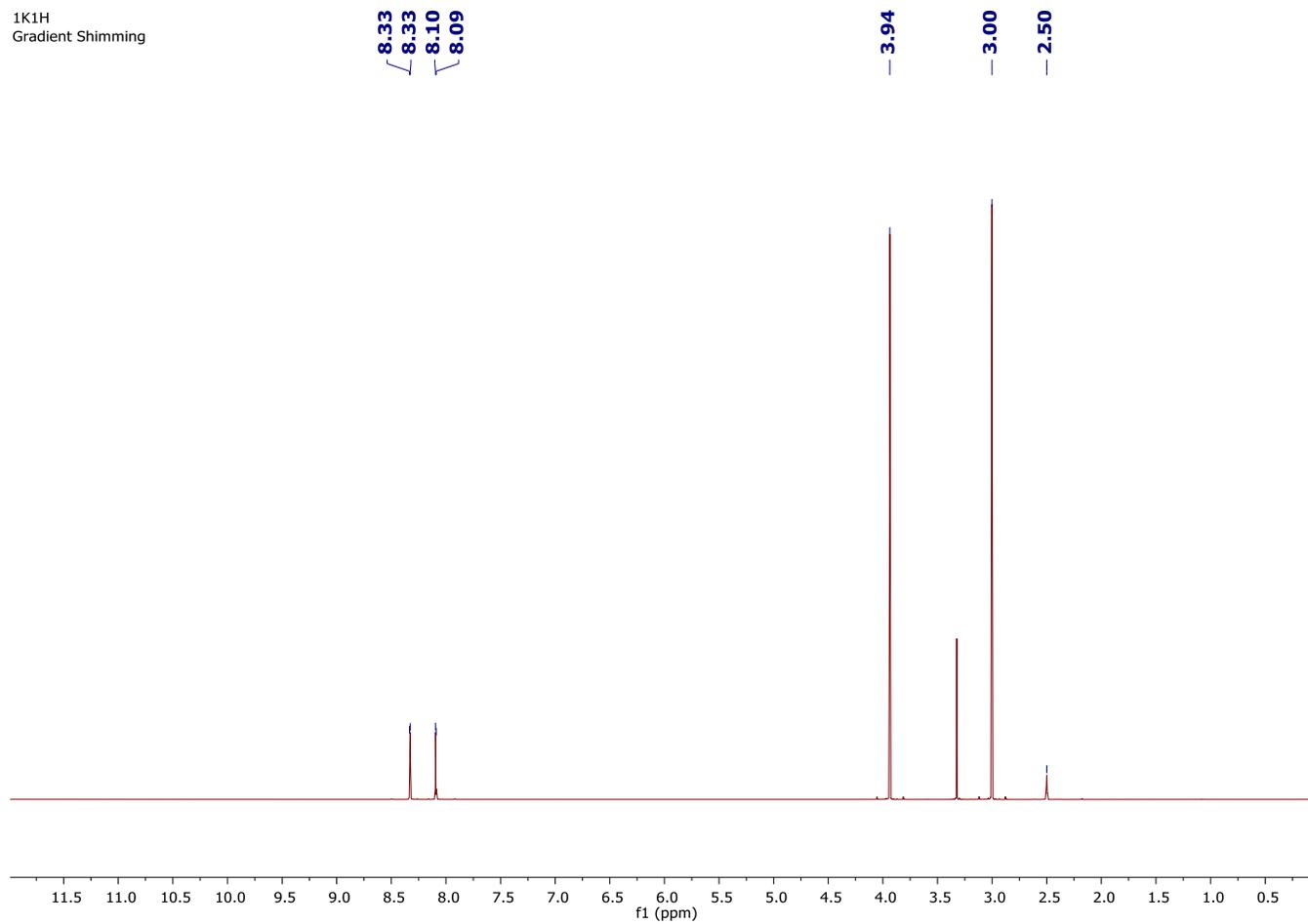


To a solution of 2-(methylthio)thiazole **1e** (203 mg, 1.56 mmol) in dry acetonitrile (5 mL), CH_3I (486 μL , 7.8 mmol) was added. The reaction mixture was refluxed for 6 hours. The reaction solution was cooled to room temperature and concentrated by rotary evaporation to afford the white solid. The crude was purified by recrystallization with ethanol / hexane (1:10) to afford pure **1k** as white powder (91 %, 386 mg). $^1\text{H NMR}$ (600 MHz, $\text{DMSO-}d_6$) δ 8.34 (d, $J = 4.0$ Hz,

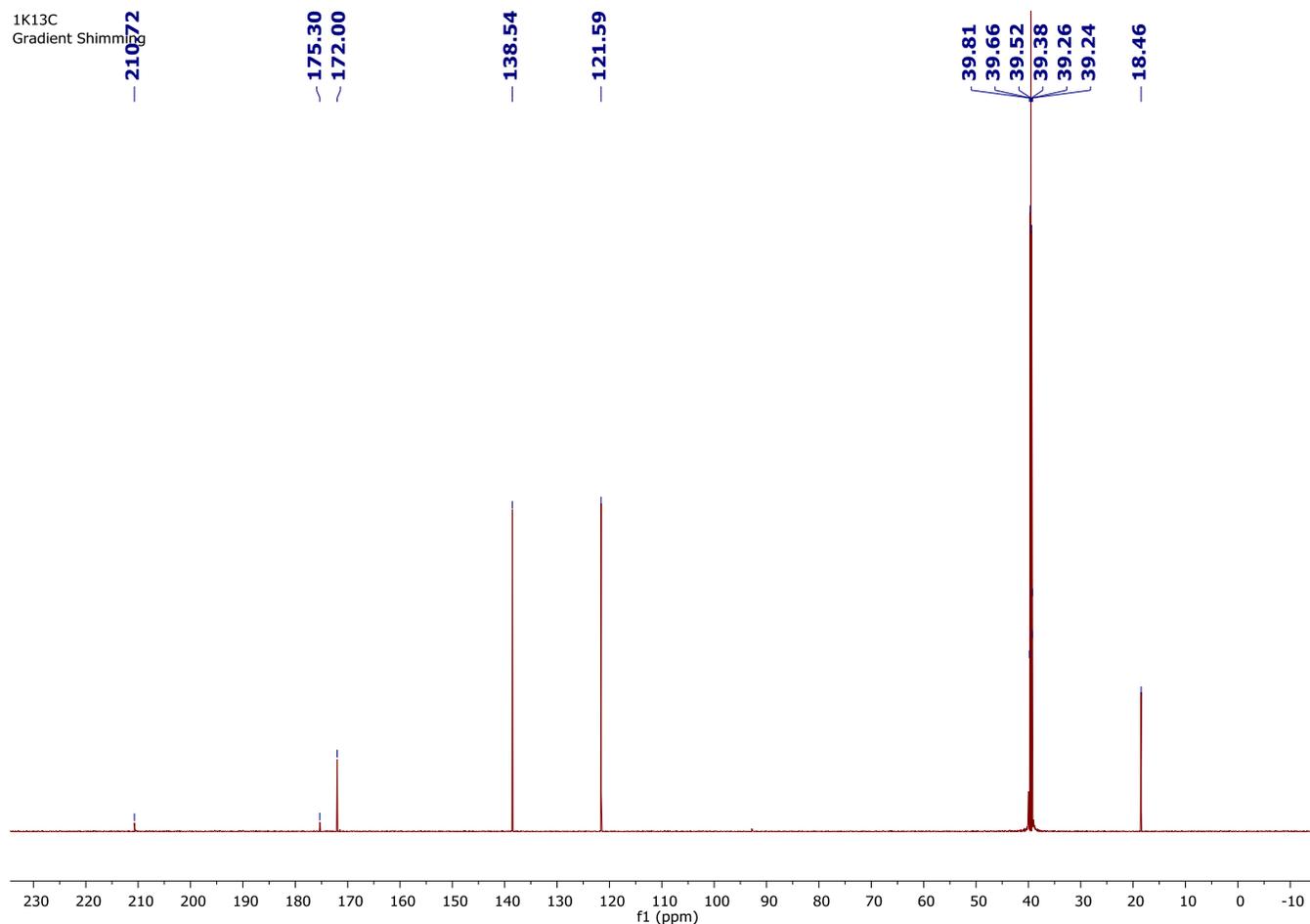
1H), 8.10 (d, $J = 4.0$ Hz, 1H), 3.92 (s, 3H), 2.99 (s, 3H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 211.47, 175.09, 172.00, 138.36, 121.86, 18.85.

^1H NMR spectra of 1K

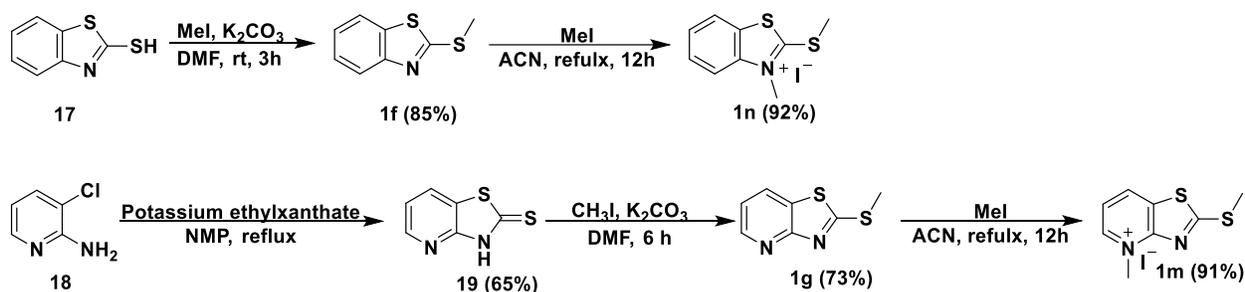
1K1H
Gradient Shimming



^{13}C NMR spectra of 1K



Synthesis of 1n and 1m



Synthesis of compound 1f

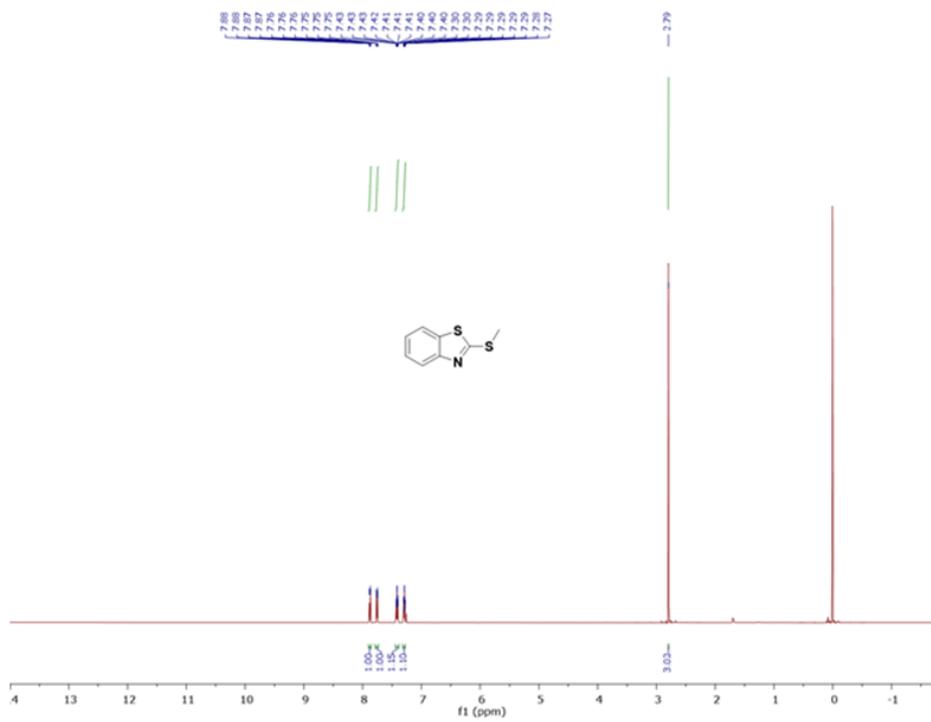
Mercaptobenzothiazole **17** (3 g, 18.0 mmol), potassium carbonate (2.48 g, 18.0 mmol), and methyl iodide (1.1 mL, 18.0 mmol) were added sequential into dry DMF (20 mL) at 0 °C. The mixture was warmed to room temperature and stirred for 3 hours. The reaction mixture was

washed with ethyl acetate and brine for 3 times. The organic layer was collected, dried over anhydrous MgSO_4 , filtered, and concentrated under the reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 3:1) to yield compound **1f** as colorless crystal (2.77 g, 85 %). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.88 (d, $J = 8$ Hz, 1 H), 7.75 (d, $J = 8$ Hz, 1 H), 7.42 (t, $J = 8$ Hz, 1 H), 7.29 (d, $J = 8$ Hz, 1 H), 2.97 (s, 3 H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 167.94, 153.29, 135.09, 125.97, 124.00, 121.30, 120.88, 15.85.

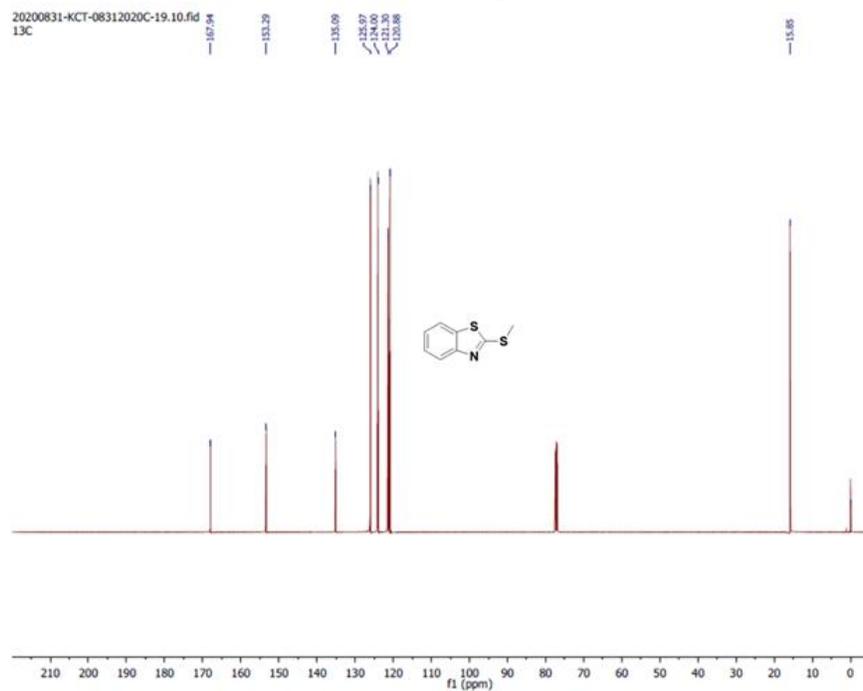
Synthesis of compound **1n**

The mixture of 2-(Methylthio)benzothiazole **1f** (1.5 g, 8.28 mmol) and MeI (2.6 mL, 41.43 mmol) in dry ACN (10 mL) was refluxed for 12 hours. The reaction mixture was cooled to room temperature and solvent was removed by rotary evaporation. The recrystallization was carried out in EtOH : Hexane (1:20) to obtain pure compound **1n** as yellow powder (92 %, 1.49 g). $^1\text{H NMR}$ (600 MHz, $\text{DMSO}-d_6$) δ 8.42 (d, $J = 8.1$ Hz, 1H), 8.21 (d, $J = 8.5$ Hz, 1H), 7.88 – 7.82 (m, 1H), 7.74 (t, $J = 7.7$ Hz, 1H), 4.11 (s, 3H), 3.13 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, $\text{DMSO}-d_6$) δ 181.13, 142.45, 129.10, 128.18, 126.93, 123.94, 115.66, 36.53, 18.20.

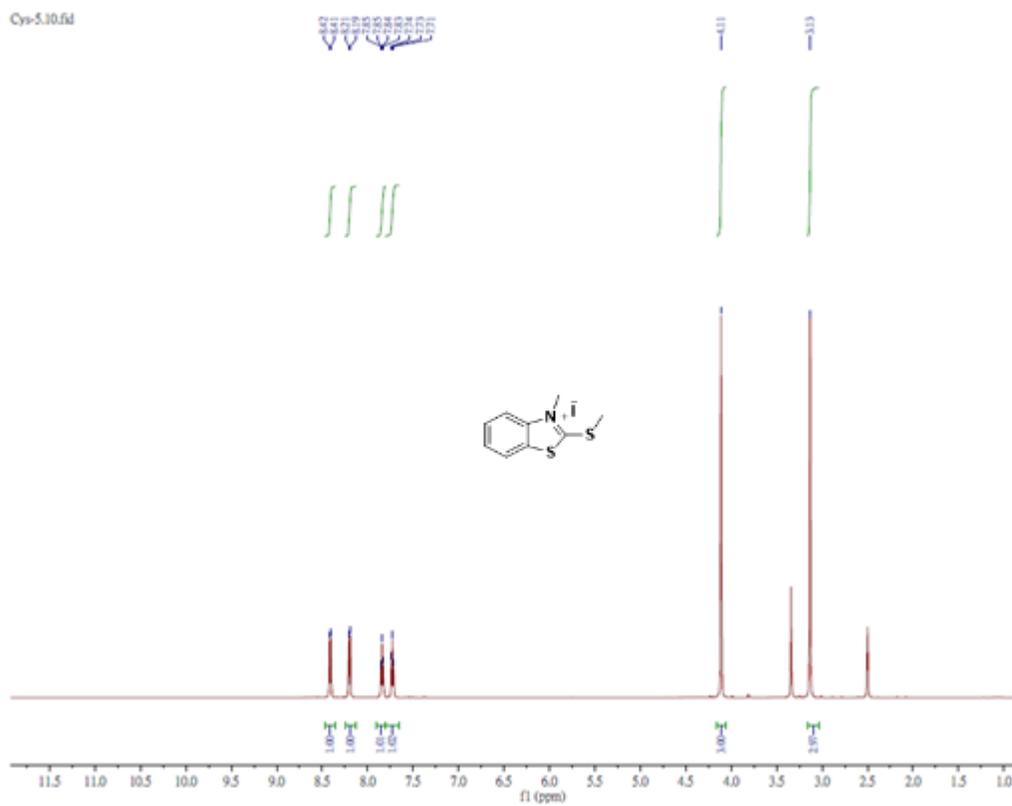
$^1\text{H NMR}$ of compound **1f**



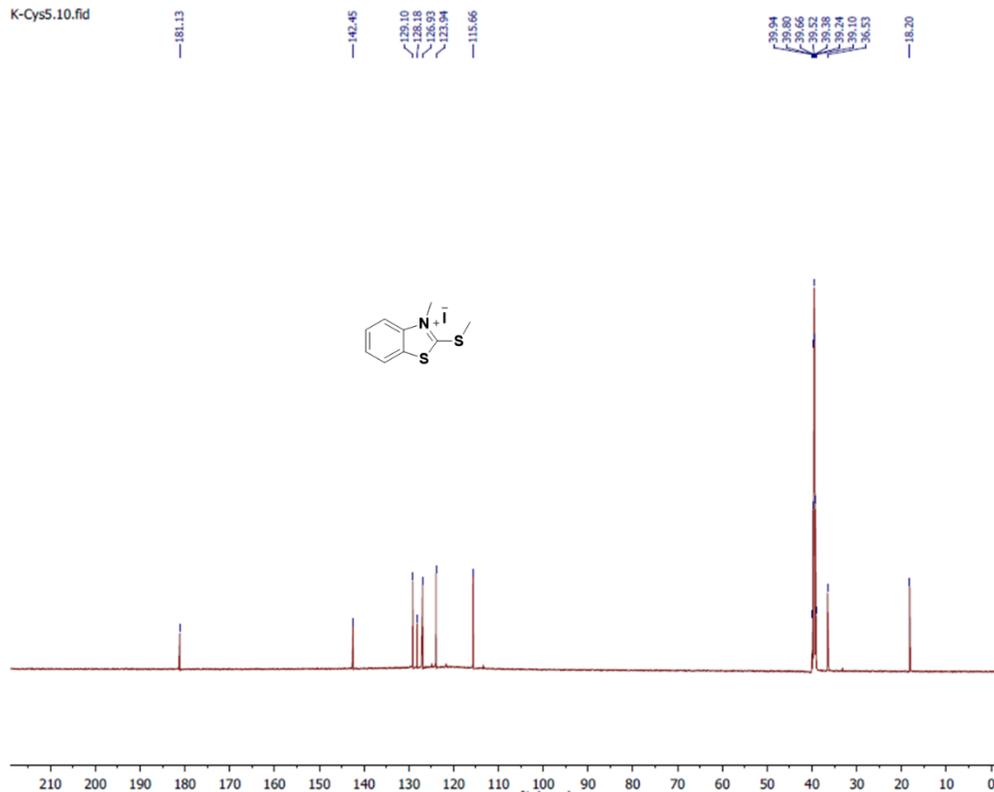
¹³C NMR of compound 1f



¹H NMR of compound 1n



¹³C NMR of compound 1n



Synthesis of compound **19**

2-Amino-3-chloro-pyridine **18** (1 g, 7.78 mmol) was dissolved in NMP (15 mL) and potassium ethyl xanthate (1.87 g, 11.6 mmol) was added. The solution was heated to 160 °C for 12 h. The solution was then cooled to RT and treated with glacial acetic acid (5 mL) and diluted with water (100 mL). The resulting precipitate was filtered off and washed with diethyl ether for three times. The off-white precipitate was dried under high vacuum to obtain compound **19** as off-white powder (65 %, 850 mg). ¹H NMR (600 MHz, DMSO-*d*₆) δ 14.28 (s, 1H), 8.35 (dd, *J* = 4.9, 1.5 Hz, 1H), 8.12 (dd, *J* = 7.9, 1.5 Hz, 1H), 7.30 (dd, *J* = 7.9, 4.9 Hz, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 191.01, 153.73, 146.80, 130.52, 124.17, 119.58, 39.94.

Synthesis of compound **1g**

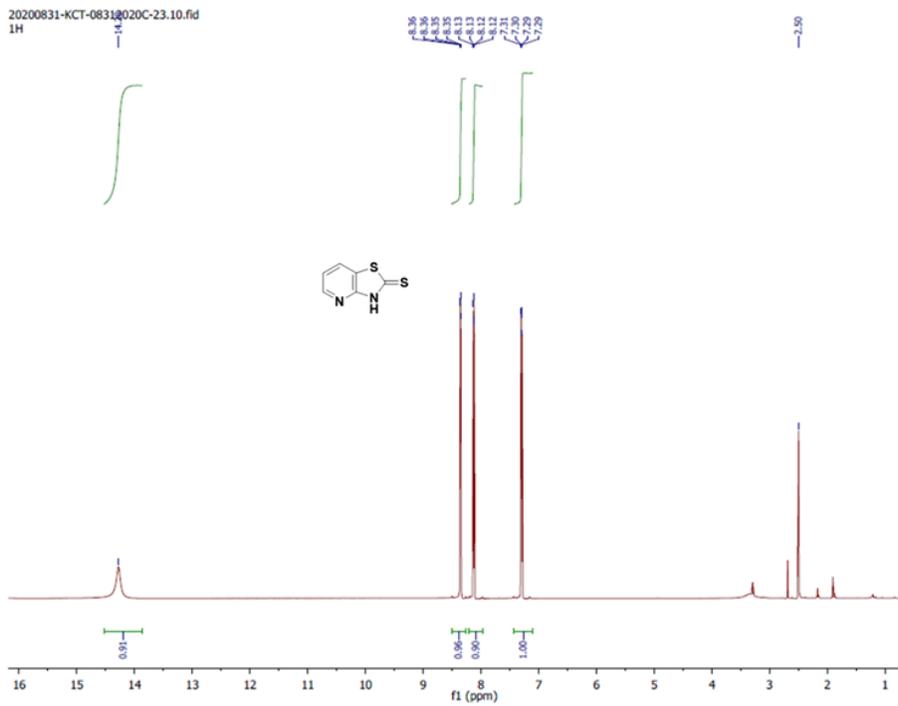
To solution of thiazolo[4,5-*b*]pyridine-2(3H)-thione **H** (400 mg, 2.38 mmol) in dry DMF (10 mL) was added K₂CO₃ (328 mg, 2.38 mmol), MeI (149 μL, 2.38 mmol) sequentially at 0 °C,

then the reaction was warmed to room temperature and stirred for 6 h. The reaction mixture was washed with ethyl acetate and brine for three times. The organic layer was collected and dried over anhydrous MgSO_4 , filtered, and concentrated. The crude was purified by flash chromatography (hexane: ethyl acetate 2:1) to afford compound **1g** as colorless crystal (73 %, 312 mg). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 8.58 (dd, $J = 4.7, 1.7$ Hz, 1H), 8.07 (dd, $J = 7.9, 1.7$ Hz, 1H), 7.19 (dd, $J = 7.9, 4.7$ Hz, 1H), 2.84 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 172.47, 163.86, 147.50, 129.68, 128.77, 118.83, 15.88.

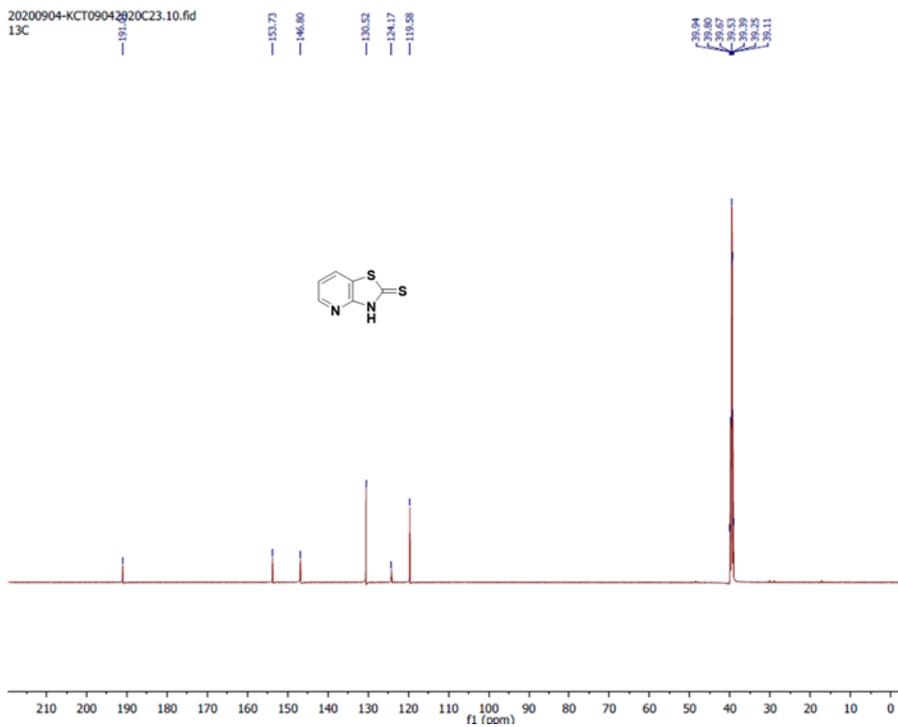
Synthesis of compound **1m**

2-(Methylthio)thiazolo[4,5-b]pyridine **1g** (150 mg, 0.82 mmol) was dissolved in dry ACN (10 mL), and treated with methyl iodide (258 μL , 4.12 mmol). The reaction mixture was refluxed for 12 h. The mixture was cooled to room temperature and ACN was removed by rotary evaporation. The crude was purified by recrystallization with methanol / hexane (1:20) to obtain compound **1m** as bright yellow powder (91 %, 242 mg). $^1\text{H NMR}$ (600 MHz, $\text{DMSO}-d_6$) δ 9.18 (d, $J = 8.1$ Hz, 1H), 9.00 (d, $J = 6.1$ Hz, 1H), 7.92 (d, $J = 1.9$ Hz, 1H), 4.48 (s, 3H), 2.98 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, $\text{DMSO}-d_6$) δ 210.72, 177.17, 171.96, 150.60, 147.50, 139.71, 125.23, 120.49, 92.78, 41.69, 15.02.

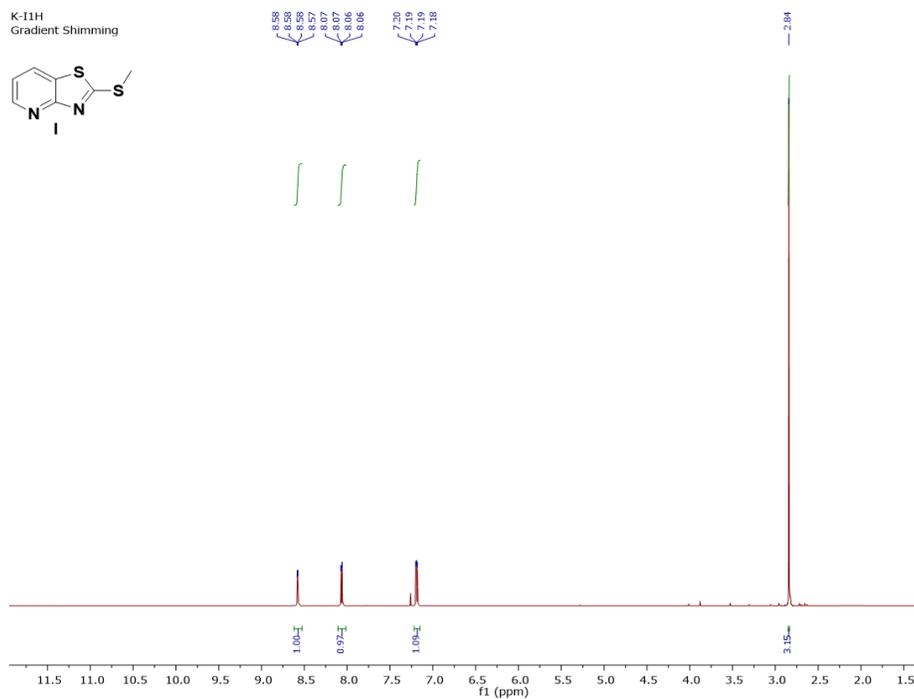
$^1\text{H NMR}$ of compound **19**



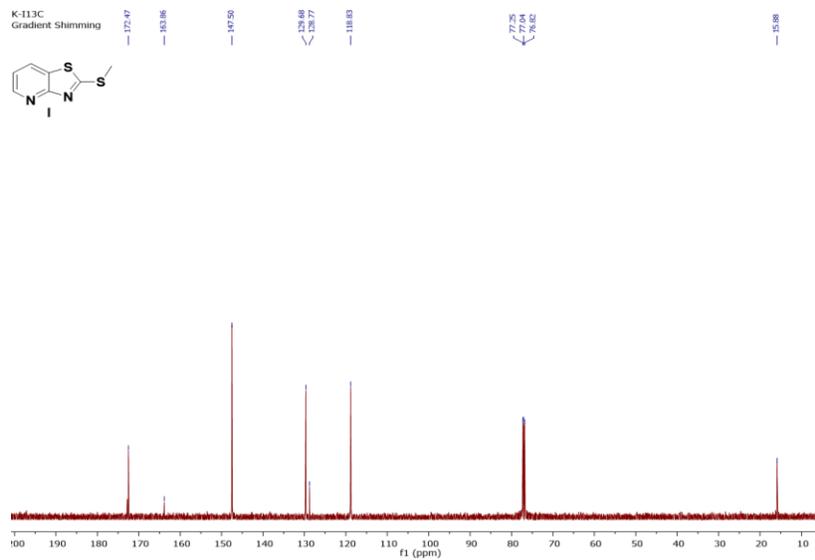
¹³C NMR of compound 19



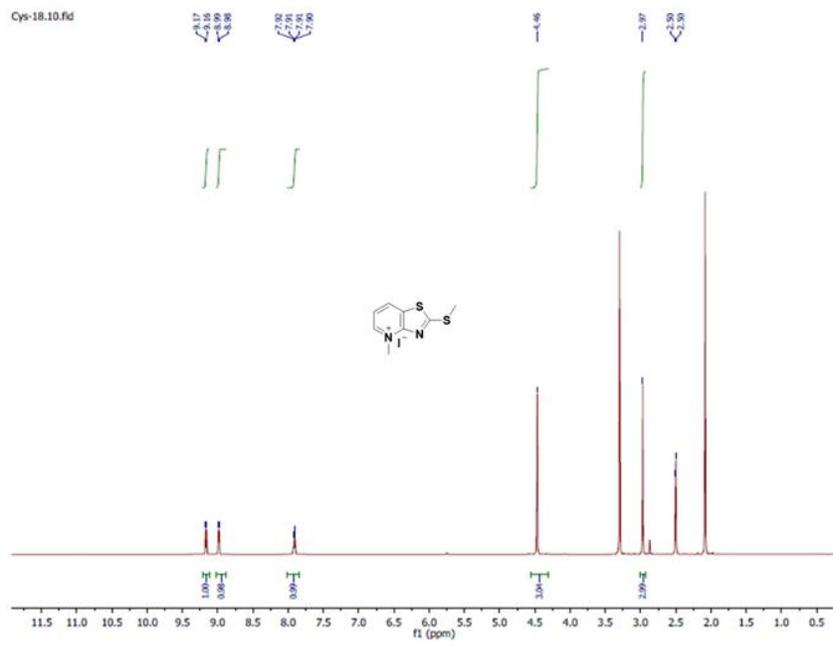
¹H NMR of compound 1g



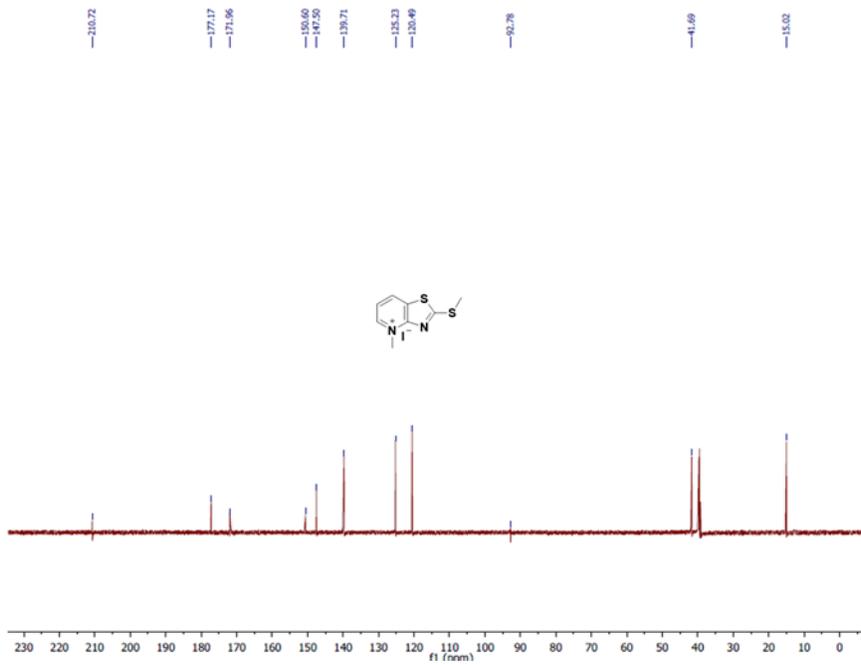
¹³C NMR of compound 1g



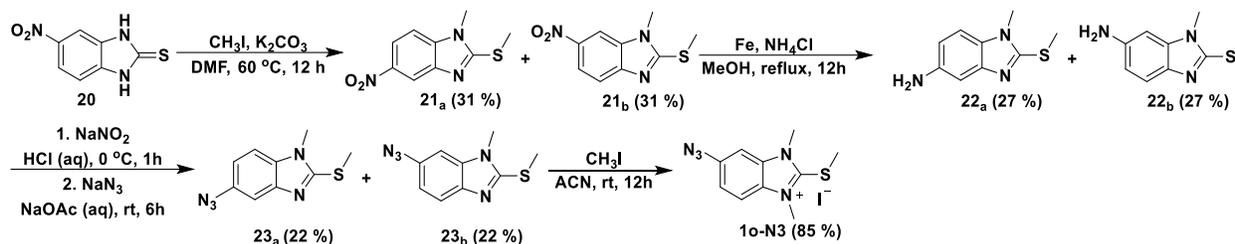
¹H NMR of compound 1m



¹³C NMR of compound 1m



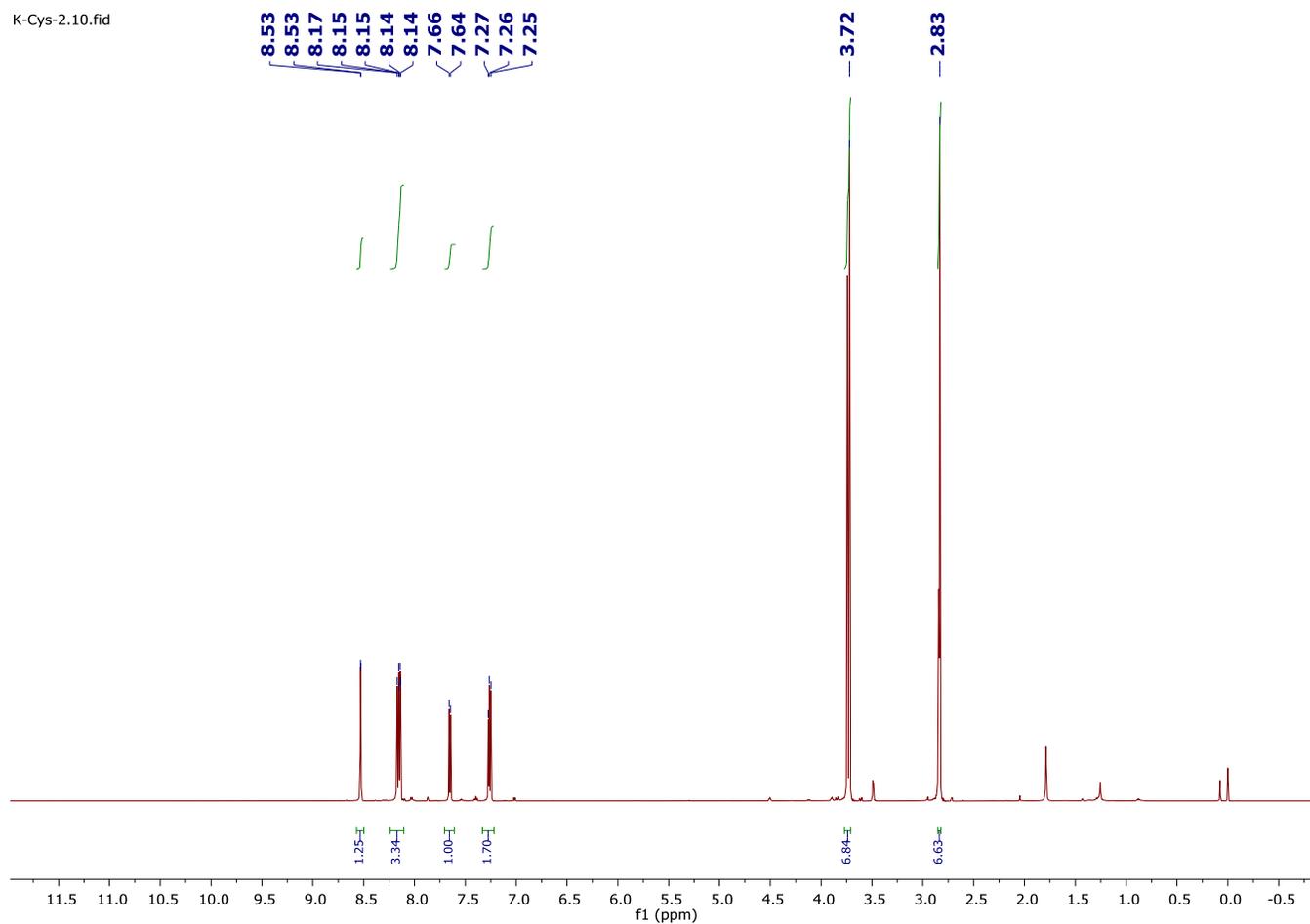
Synthesis of 1o-N3



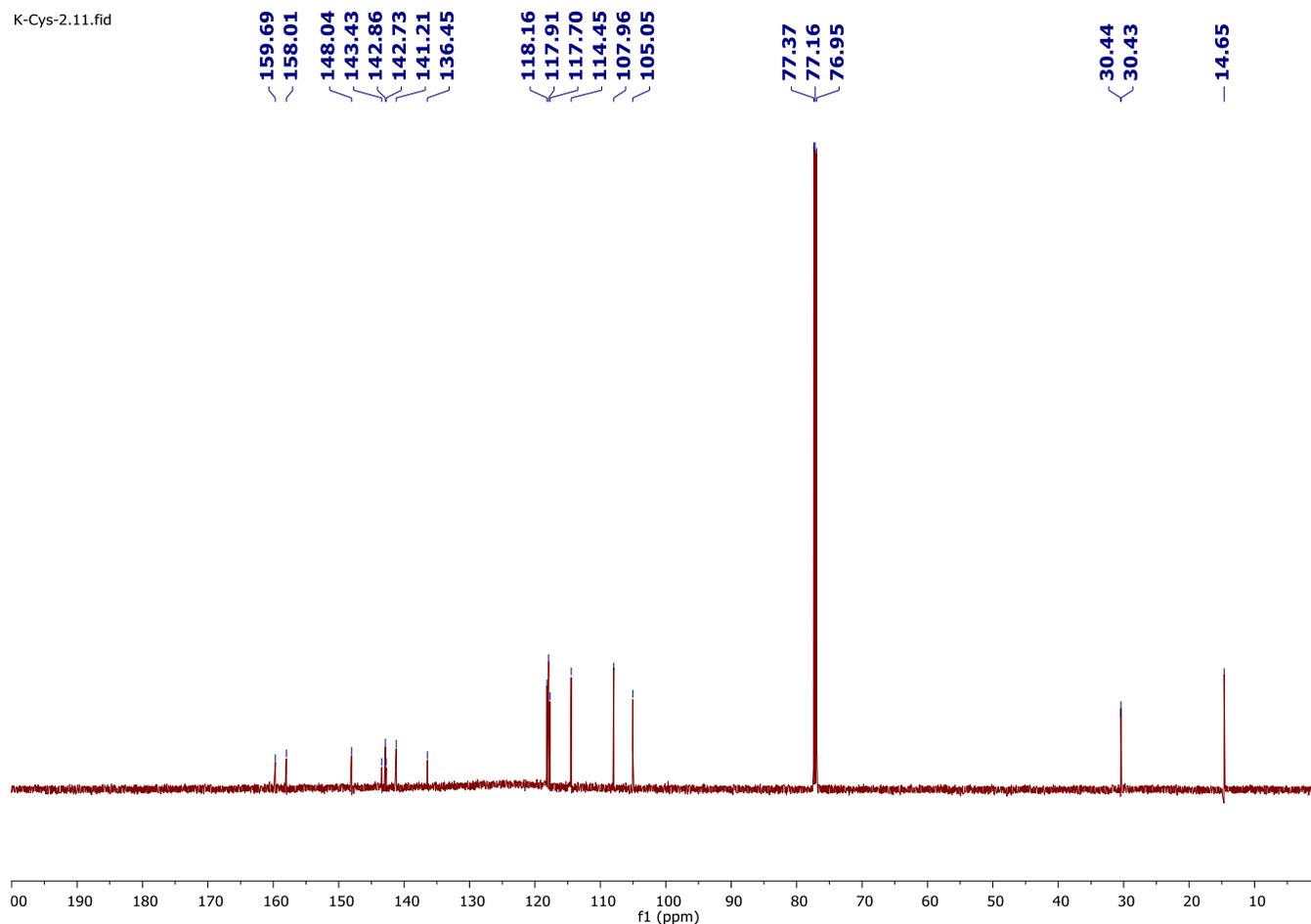
Synthesis of isomers 21_{a,b}

To a stirred solution of 2-Mercapto-5-nitrobenzimidazole **20** (1 g, 5.12 mmol) in dry DMF (15 mL), K_2CO_3 (707.7 mg, 5.12 mmol) and MeI (320 μL , 5.12 mmol) were added consecutively at room temperature, then the reaction was heated to $60\text{ }^\circ\text{C}$ for 12 h. The mixture was cooled to room temperature and extracted with ethyl acetate and brine for three times. The organic portion was collected and dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The crude was purified by flash chromatography (hexane: ethyl acetate 1:1) to afford isomers **21_a** and **21_b** as yellow powder (62 %, 707.9 mg). **^1H NMR** (600 MHz, Chloroform-*d*) δ 8.53 (d, $J = 2.1$ Hz, 1H), 8.19 – 8.12 (m, 3H), 7.65 (d, $J = 8.7$ Hz, 1H), 7.26 (d, $J = 8.8$ Hz, 1H), 3.73 (d, $J = 12.7$ Hz, 6H), 2.84 (d, $J = 5.1$ Hz, 6H). **^{13}C NMR** (151 MHz, Chloroform-*d*) δ 159.58, 157.90, 147.92, 143.32, 142.75, 142.62, 141.09, 136.34, 118.05, 117.79, 117.59, 114.34, 107.84, 104.94, 30.33, 30.31, 14.53.

^1H NMR spectra of isomers 21_{a,b}



13C NMR spectra of isomers 21_{a,b}



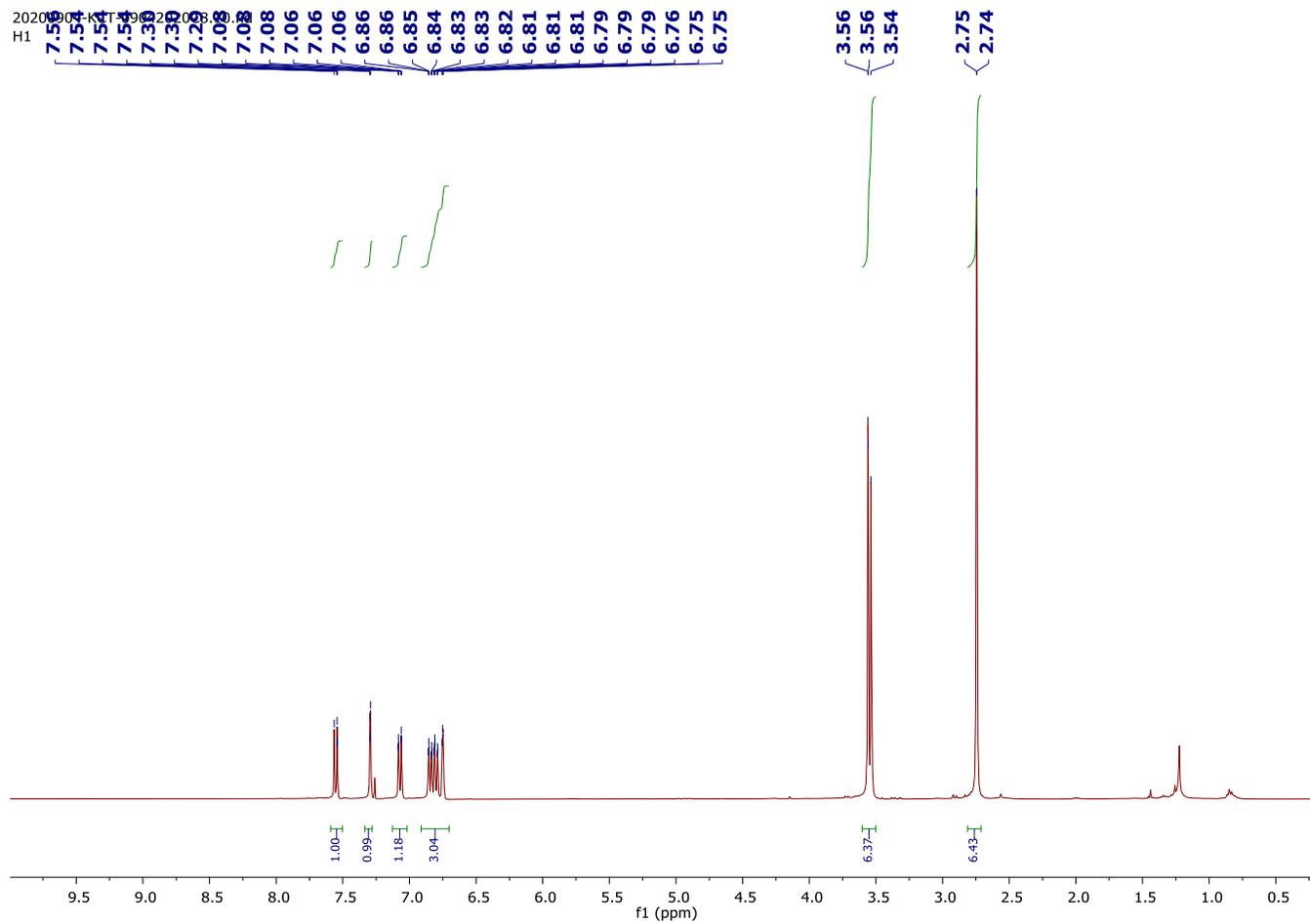
Synthesis of isomers **22_a**,**22_b**

A mixture of the isomers **21_a** and **21_b** (500 mg, 2.24 mmol), iron powder (125 mg, 2.24 mmol) and ammonia chloride (600 mg, 11.2 mmol) in methanol (50 mL) was vigorously stirred and refluxed for 12 h. The suspension was cooled and filtered through a pad of Celite. The filtrate was diluted with water and extracted with dichloromethane for 3 times. The combined organic layers were washed with saturated aqueous NaHCO₃ solution for 3 times, dried over sodium sulfate, and concentrated in vacuo to afford isomers **22_a** and **22_b** as purple oil (54 %, 233 mg). The isomers **22_a** and **22_b** were used directly without further purification.

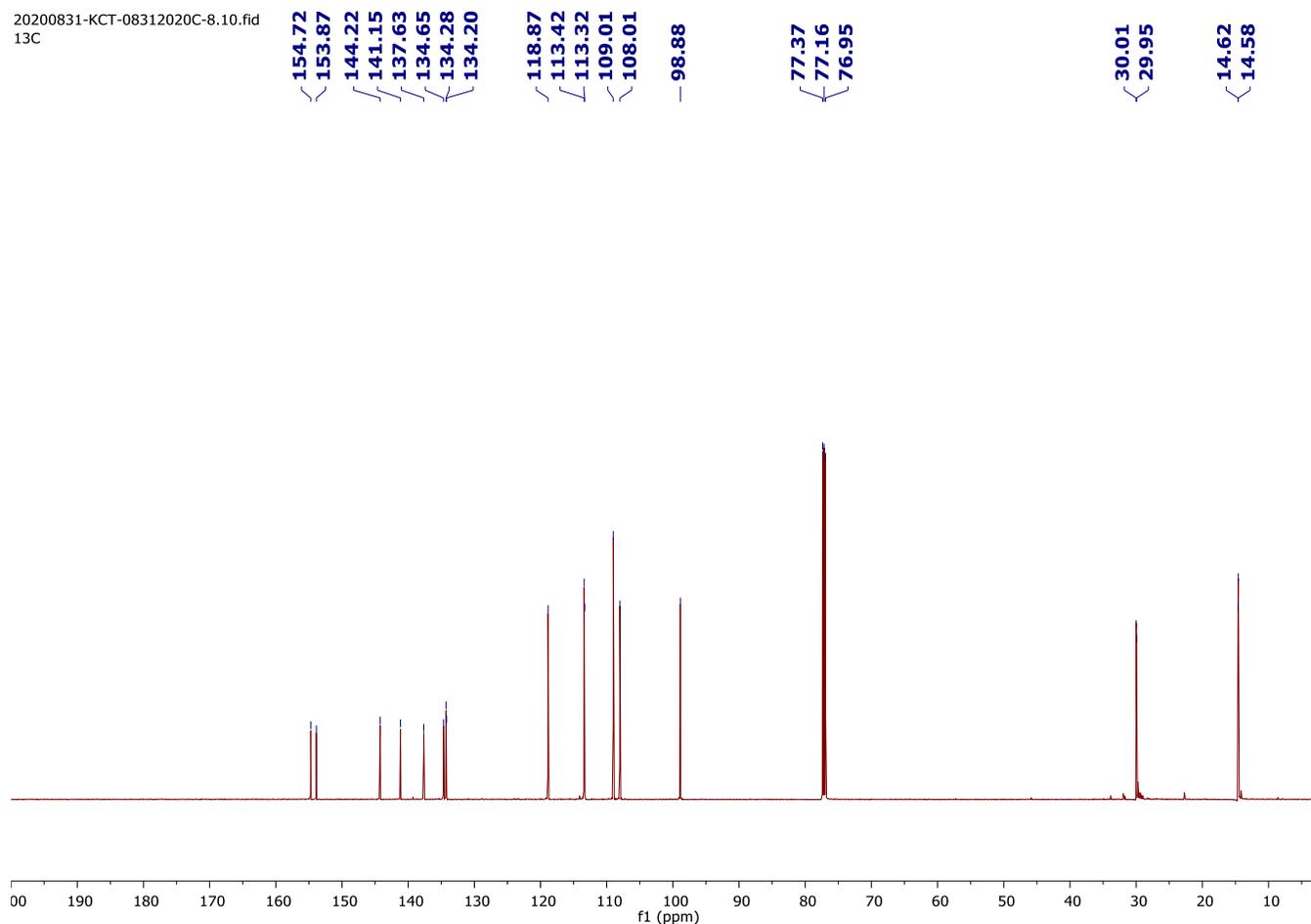
Synthesis of isomers **23**_{a,b}

To a solution of isomers **22**_a and **22**_b (200 mg, 1.04 mmol) in 3M HCl (2.5 mL) was added a solution of NaNO₂ (71.7 mg, 1.04 mmol) in H₂O (2.5 mL) at 0°C. After 60 min at 0°C a solution of NaN₃ (202.7 mg, 10.5 mmol) in saturated NaOAc solution (2.5 mL) was added dropwise and the mixture was stirred for 60 min at 0°C. The reaction was warmed to room temperature and allowed to stir for 6 h. The mixture was extracted with ethyl acetate and brine for three times, dried over MgSO₄, filtered and the solvent was removed under reduced pressure. The crude mixture was purified by flash chromatography (hexane : ethyl acetate 5:1) to get isomers **23**_a and **23**_b as bright yellow oil (45 %, 102.5 mg). **¹H NMR** (400 MHz, Chloroform-*d*) δ 7.55 (d, *J* = 8.5 Hz, 1H), 7.29 (d, *J* = 1.9 Hz, 1H), 7.07 (ddd, *J* = 8.5, 2.1, 1.3 Hz, 1H), 6.82 (dddd, *J* = 17.9, 8.5, 2.1, 0.8 Hz, 2H), 6.75 (t, *J* = 1.7 Hz, 1H), 3.57 – 3.55 (m, 3H), 3.55 – 3.52 (m, 3H), 2.74 (d, *J* = 1.0 Hz, 6H). **¹³C NMR** (151 MHz, Chloroform-*d*) δ 154.18, 153.33, 143.68, 140.61, 137.10, 134.11, 133.74, 133.66, 118.33, 112.88, 112.78, 108.47, 107.47, 98.34, 76.83, 76.62, 76.41, 29.47, 29.41, 14.08, 14.04.

¹H NMR isomers of **23_{a,b}**



^{13}C NMR isomers of $23_{a,b}$

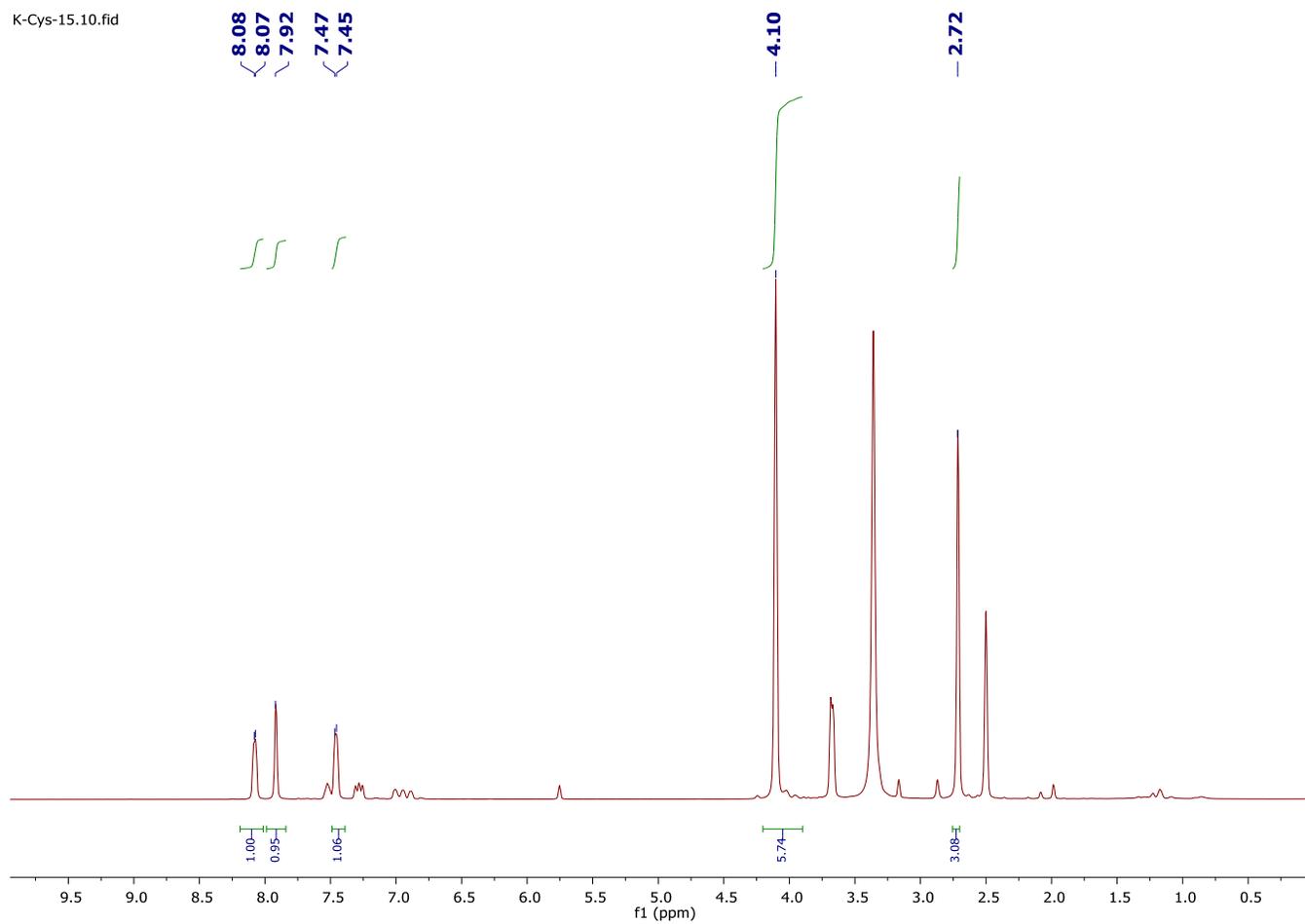


Synthesis of compound N₃-1o

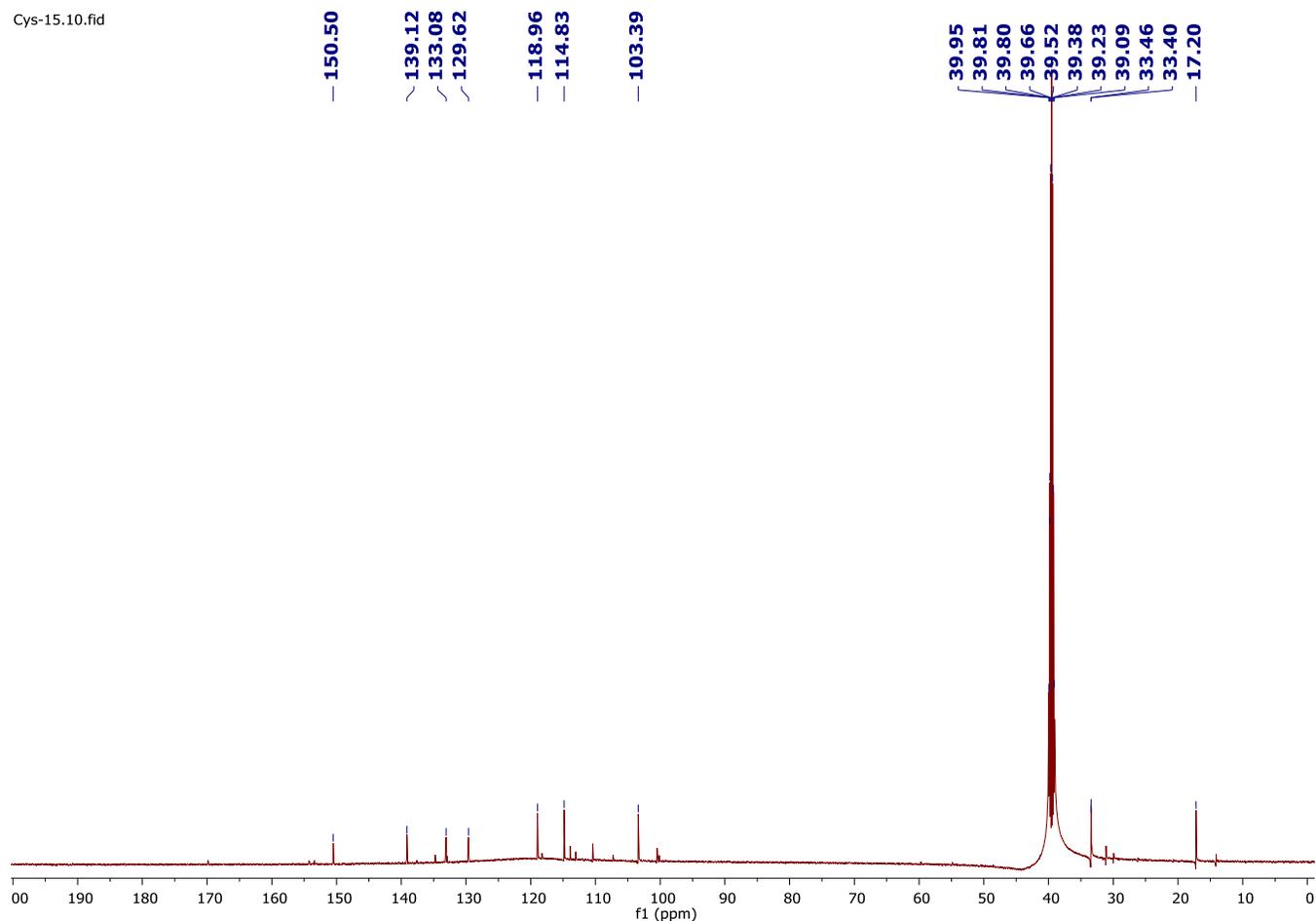
To the isomers of **23_a** and **23_b** (50 mg, 0.22 mmol) in the dry acetonitrile (7 mL), was added MeI (71.3 μ L, 1.14 mmol) under nitrogen at room temperature and reaction was stirred for 12 h. The solvent was removed by rotary evaporation to afford crude powder. The crude was washed with diethyl ether three times, dried under vacuum to obtain pure compound **N₃-1o** as bright yellow powder (85 %, 257.7 mg). **¹H NMR** (500 MHz, DMSO-*d*₆) δ 8.12 – 8.03 (m, 1H), 7.92 (d, *J* = 4.5 Hz, 1H), 7.49 – 7.40 (m, 1H), 4.10 (s, 6H), 2.72 (s, 3H). **¹³C NMR** (151 MHz, DMSO-*d*₆) δ 150.51, 139.12, 133.08, 129.63, 118.97, 114.84, 103.40, 33.46, 33.40, 17.20.

¹H NMR spectra of N₃-1o

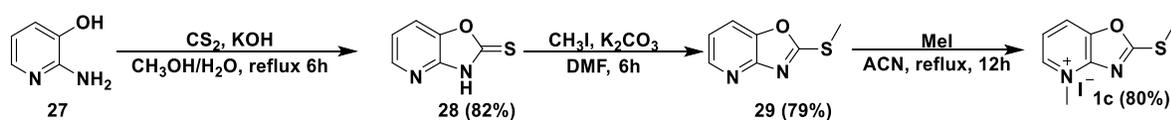
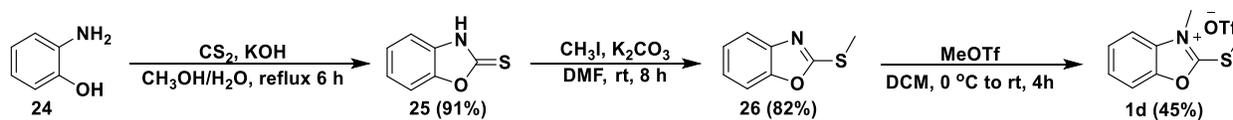
K-Cys-15.10.fid



^{13}C NMR spectra of N₃-10



Synthesis of 1d and 1c



Synthesis of compound 25

2-aminophenol **24** (1 g, 18.3 mmol) was dissolved in methanol / water (9:1, 90 mL) and KOH (1 g, 18.3 mmol) and CS₂ (3.3 mL, 55.5 mmol) were added sequentially. The reaction mixture was refluxed for 6 h. The reaction solution was washed with ethyl acetate and brine thrice. The

residue was concentrated under reduced pressure and purified by the column chromatography (hexane: ethyl acetate 2:1) to obtain compound **25** as off-white powder (91 %, 2.1 g). **¹H NMR** (600 MHz, DMSO-*d*₆) δ 13.8 (br s, 1H), 7.45 (d, *J* = 7.6 Hz, 1H), 7.27-7.20 (m, 3H). **¹³C NMR** (151 MHz, DMSO-*d*₆) δ 180.11, 148.10, 131.15, 125.06, 123.69, 110.42, 109.91.

Synthesis of compound **26**

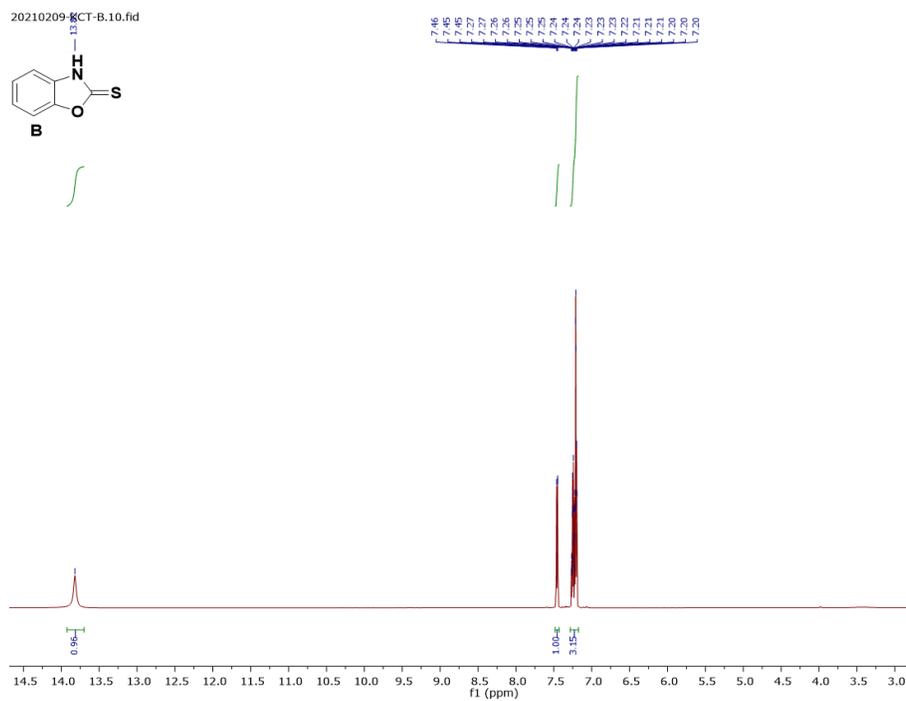
2-Mercaptobenzoxazole **25** (1 g, 6.62 mmol), potassium carbonate (914 mg, 6.62 mmol), and methyl iodide (412 μL, 6.62 mmol) were sequentially added into dry DMF (13.5 mL) at 0 °C. The mixture was warmed to room temperature and stirred for 8 hours. The reaction mixture was washed with ethyl acetate and brine for 3 times. The organic layer was collected, dried over anhydrous MgSO₄, filtered, and concentrated under the reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 5:1) to yield compound **26** as yellow oil (895.7 mg, 82 %). **¹H NMR** (600 MHz, CDCl₃) δ 7.54 (dd, *J* = 8.0, 4.5 Hz, 1H), 7.38 – 7.30 (m, 1H), 7.24 – 7.06 (m, 2H), 2.67 (s, 3H). **¹³C NMR** (151 MHz, CDCl₃) δ 165.53, 165.51, 151.84, 151.83, 141.84, 124.09, 124.07, 123.61, 118.16, 118.15, 109.66, 109.65, 14.37.

Synthesis of compound **1d**

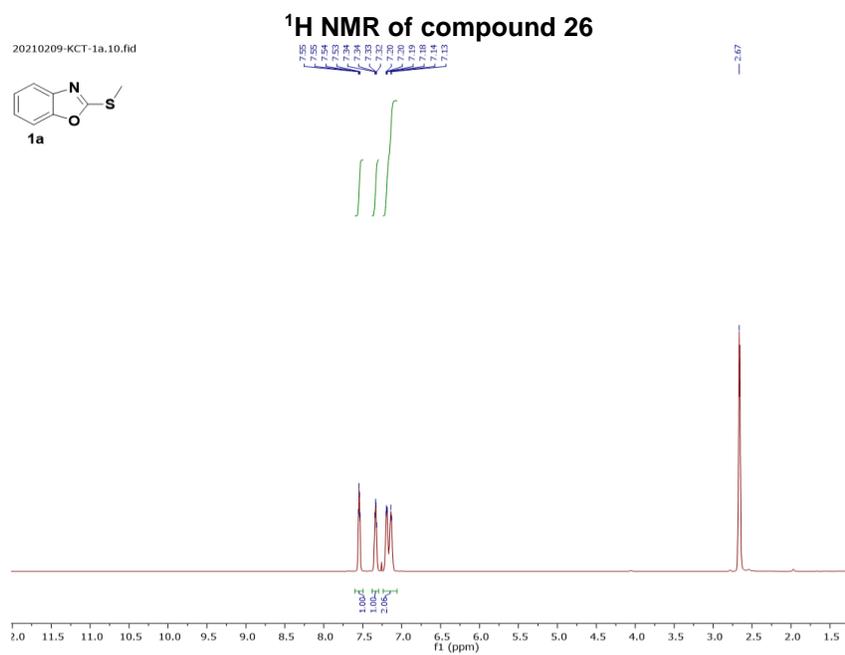
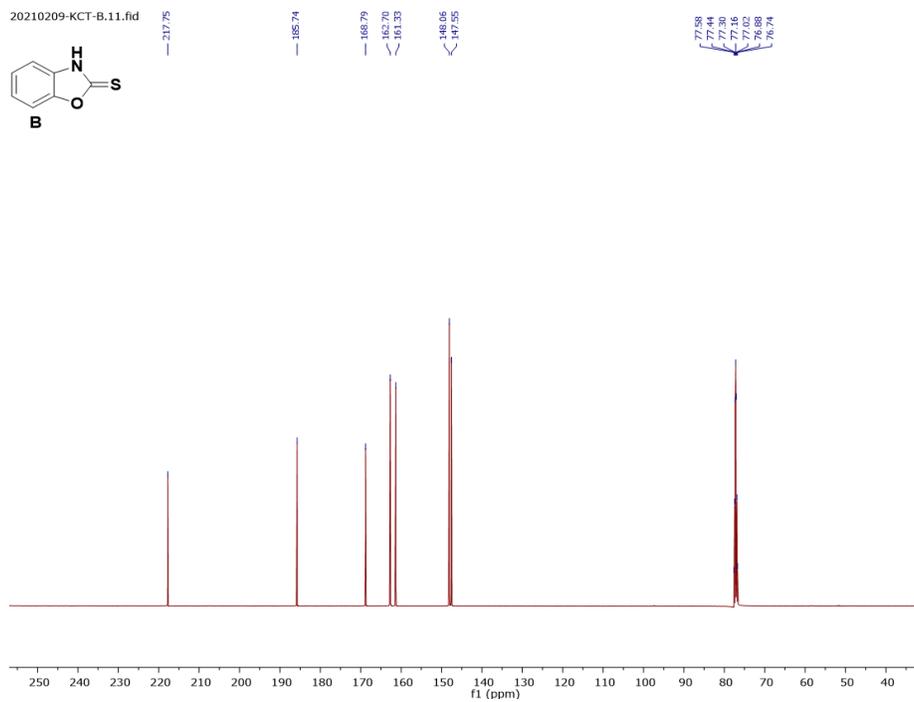
To a mixture of 2-(methylthio)benzoxazole **26** (86 mg, 0.52 mmol) in dry DCM (10 mL), MeOTf (70 μL, 0.625 mmol) was added at 0 °C. The reaction mixture was warmed to room temperature. After 4 h, solvent was removed by rotary evaporation to afford crude compound. The crude sample was purified by recrystallization in hexane to obtain pure compound **1d** as white powder (214.6 mg, 45 %). **¹H NMR** (600 MHz, DMSO-*d*₆) δ 8.08 (d, *J* = 8.6 Hz, 1H),

8.03 (d, $J = 7.5$ Hz, 1H), 7.76 – 7.66 (m, 2H), 3.94 (s, 3H), 3.05 (s, 3H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 172.04, 127.52, 127.49, 123.85, 122.17, 113.44, 112.28, 32.51, 14.

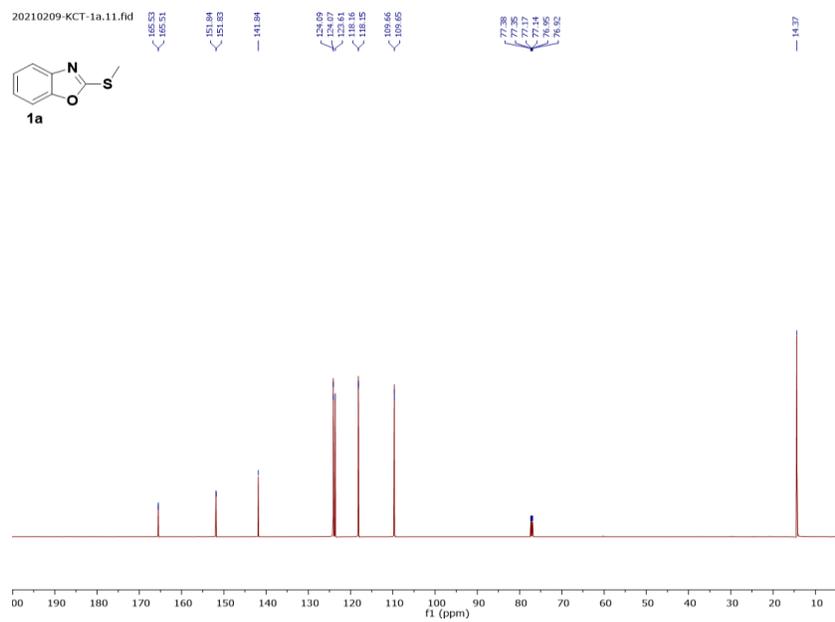
^1H NMR of compound 25



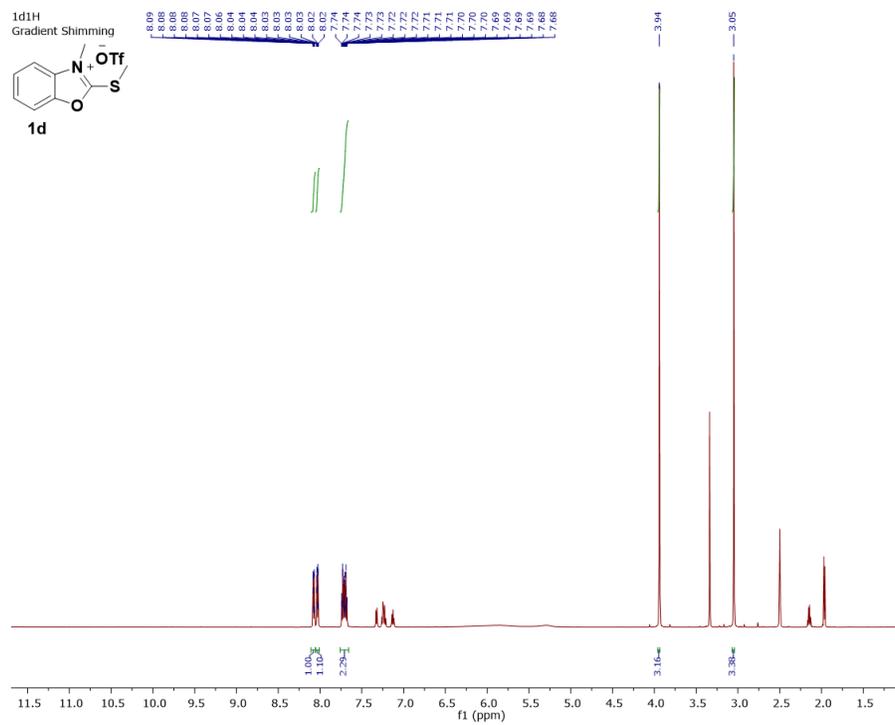
^{13}C NMR of compound 25



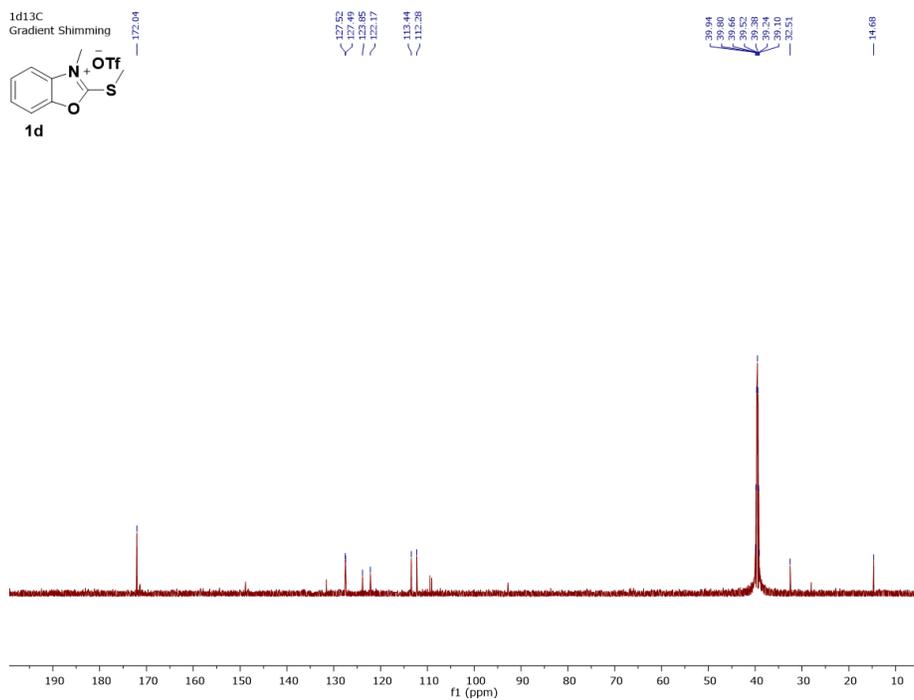
¹³C NMR of compound 26



¹H NMR of compound 1d



¹³C NMR of compound 1d



Synthesis of compound 28

2-Amino-3-hydroxypyridine **27** (1.1 g, 10 mmol) was dissolved in methanol / water (10:1, 81 mL) and KOH (1.38 g, 10 mmol) and CS₂ (1.8 mL, 30 mmol) were added sequentially. The reaction mixture was refluxed for 6 h. The solution was then cooled to RT and treated with glacial acetic acid (5 mL) and diluted with water (150 mL). The resulting precipitate was filtered off and washed with hexane for three times to obtain compound **28** as off-white powder (82 %, 1.24 g). ¹H NMR (600 MHz, DMSO-*d*₆) δ 14.47 (s, 1H), 8.19 (d, *J* = 5.0 Hz, 1H), 7.88 – 7.80 (m, 1H), 7.24 (d, *J* = 5.3 Hz, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 181.33, 146.97, 144.16, 141.59, 119.09, 117.02, 117.01.

Synthesis of compound **29**

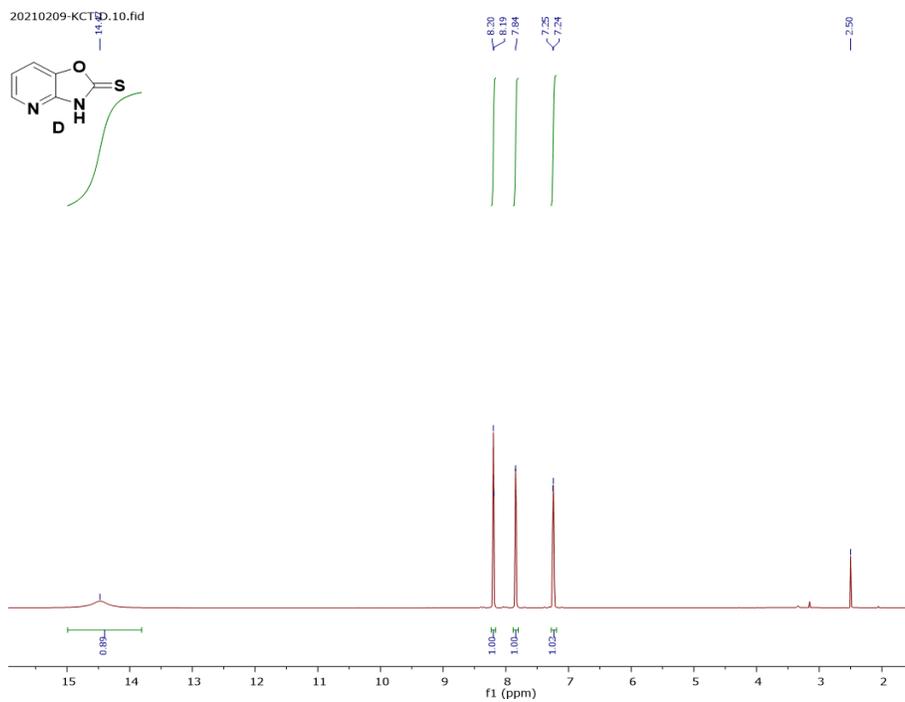
To solution of oxazolo[4,5-*b*]pyridine-2(3H)-thione **28** (500 mg, 3.29 mmol) in dry DMF (12 mL) was added K₂CO₃ (454 mg, 3.29 mmol), MeI (206 μL, 3.29 mmol) sequentially at 0 °C, then the reaction was warmed to room temperature and stirred for 6 h. The reaction mixture was washed with ethyl acetate and brine for three times. The organic layer was collected and dried over anhydrous MgSO₄, filtered, and concentrated. The crude was purified by flash chromatography (hexane: ethyl acetate 5:1) to afford compound **29** as white powder (79 %, 431 mg). ¹H NMR (600 MHz, CDCl₃) δ 8.47 – 8.36 (m, 1H), 7.74 – 7.60 (m, 1H), 7.20 – 7.07 (m, 1H), 2.78 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.35, 156.22, 145.89, 145.87, 144.25, 118.86, 117.08, 14.82.

Synthesis of compound **1c**

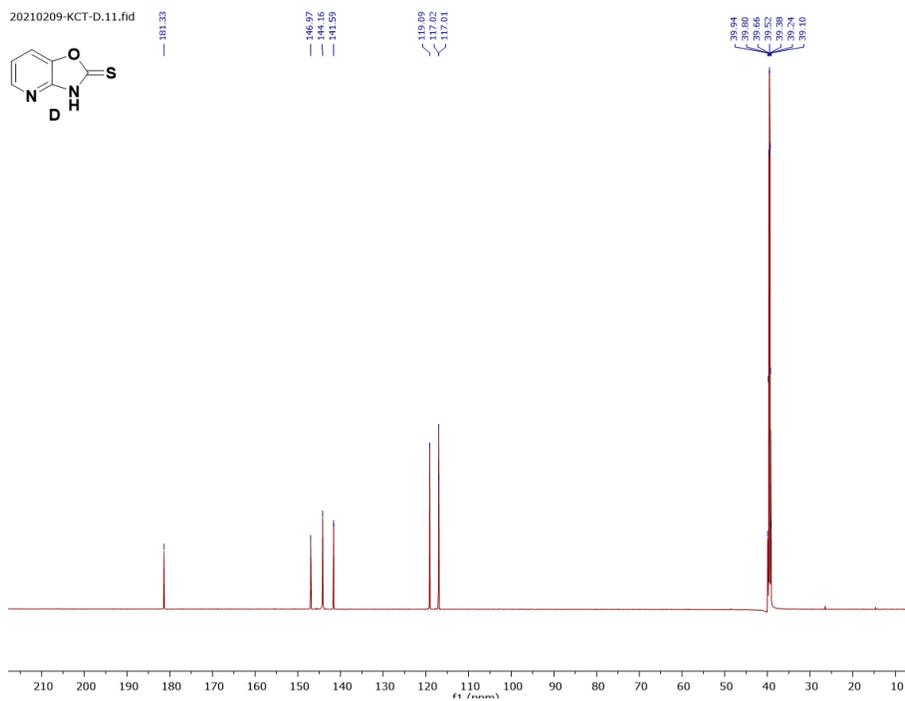
2-(Methylthio)oxazolo[4,5-*b*]pyridine **29** (166 mg, 1 mmol) was dissolved in dry ACN (10 mL), and treated with methyl iodide (313 μL, 5 mmol). The reaction mixture was refluxed for 12 h.

The mixture was cooled to room temperature and ACN was removed by rotary evaporation. The crude was purified by recrystallization with methanol / hexane (1:15) to obtain compound **1c** as bright yellow powder (80 %, 246 mg). **¹H NMR** (600 MHz, DMSO-*d*₆) δ 8.88 – 8.84 (m, 1H), 8.82 (d, *J* = 8.2 Hz, 1H), 8.07 – 7.78 (m, 1H), 4.37 (s, 3H), 2.94 (s, 3H). **¹³C NMR** (151 MHz, DMSO-*d*₆) 177.23, 171.99, 150.63, 147.54, 139.75, 125.25, 125.24, 120.52, 120.50, 41.69, 15.01.

¹H NMR of compound 28

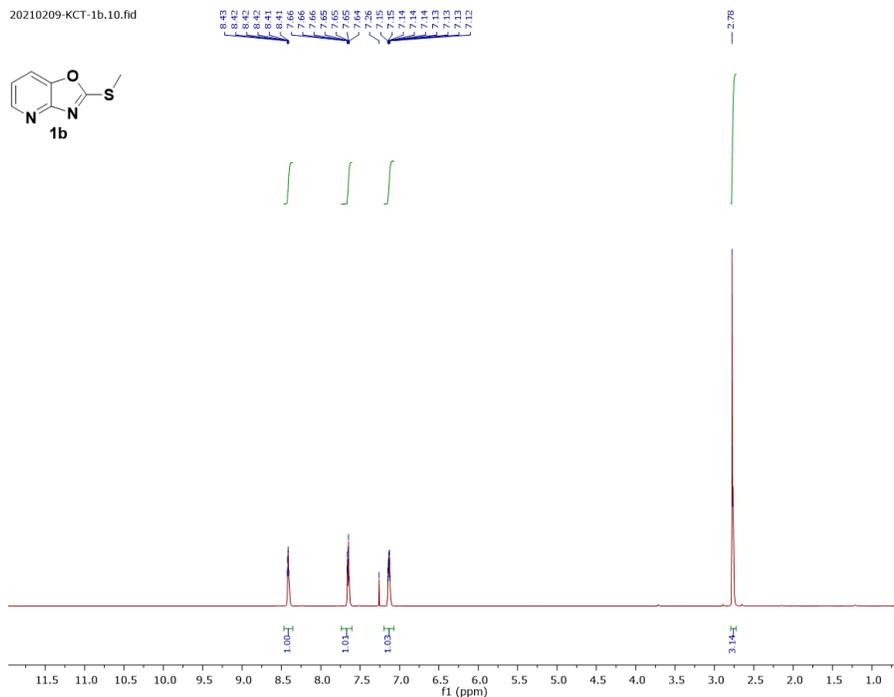
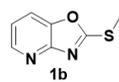


¹³C NMR of compound 28



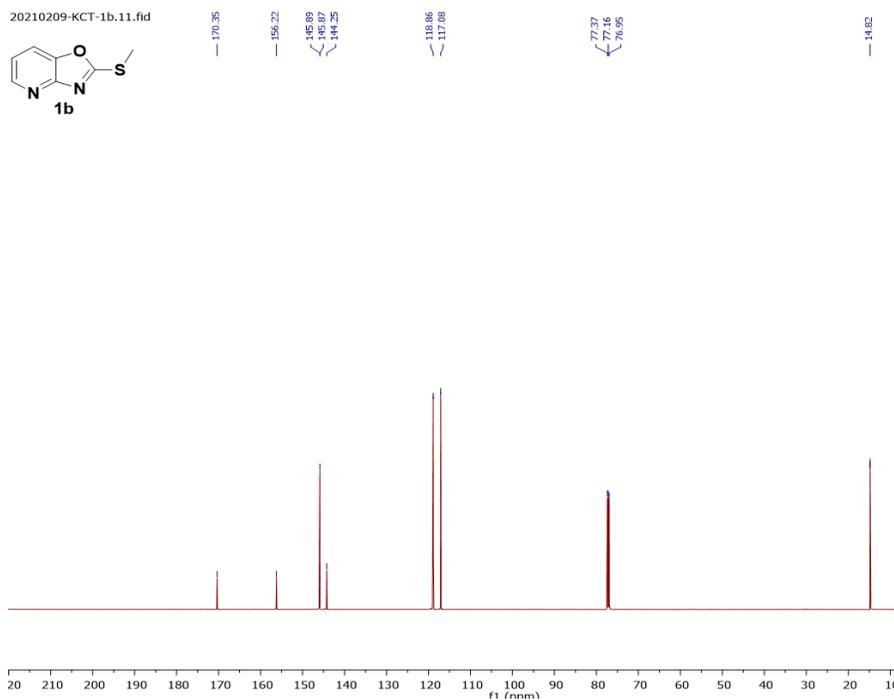
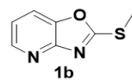
¹H NMR of compound 29

20210209-KCT-1b.10.fid

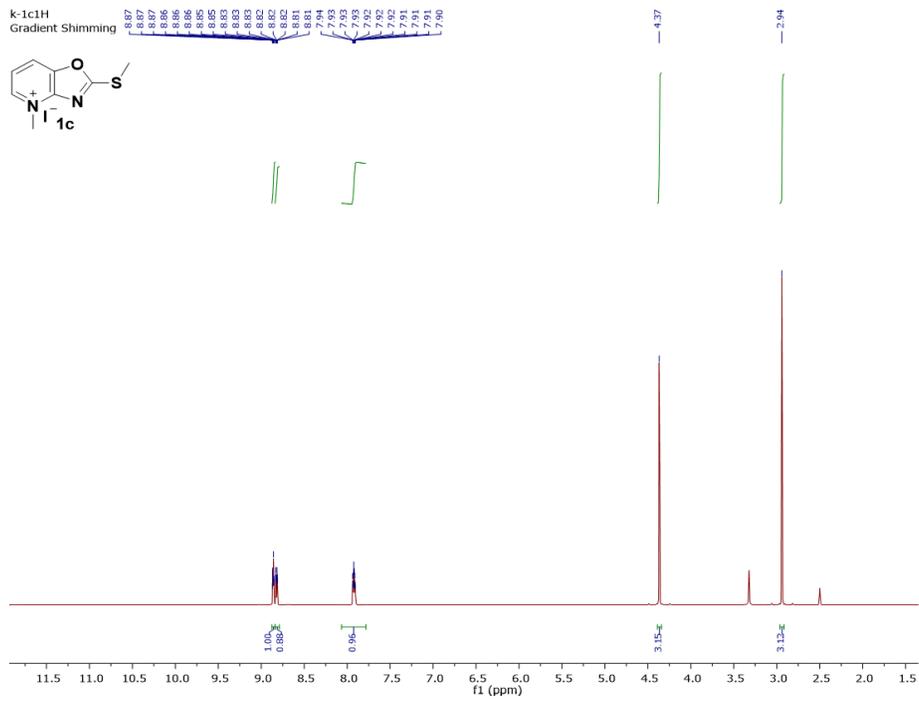


¹³C NMR of compound 29

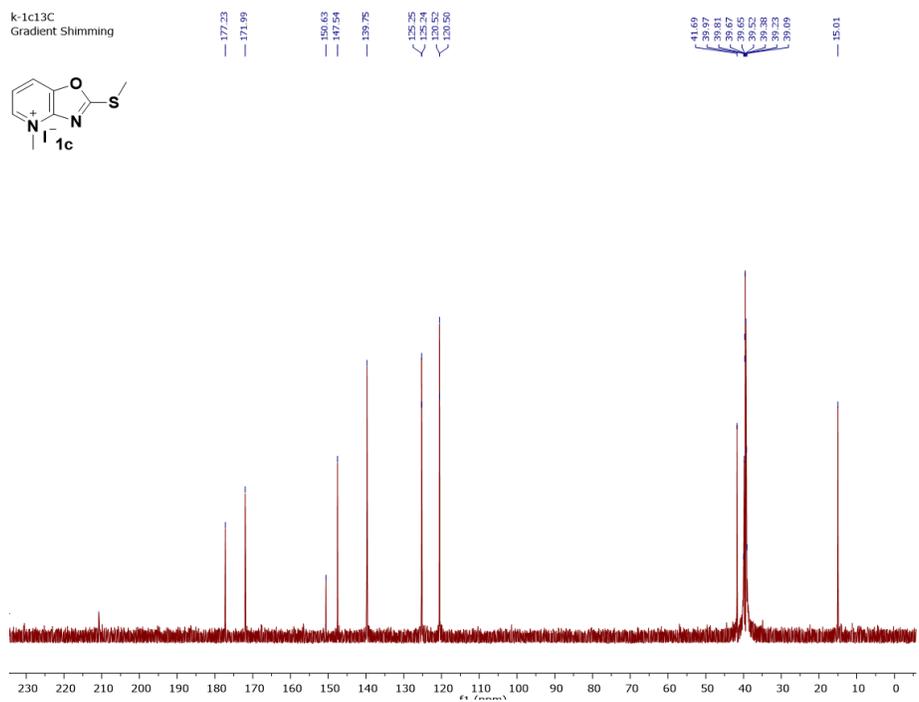
20210209-KCT-1b.11.fid



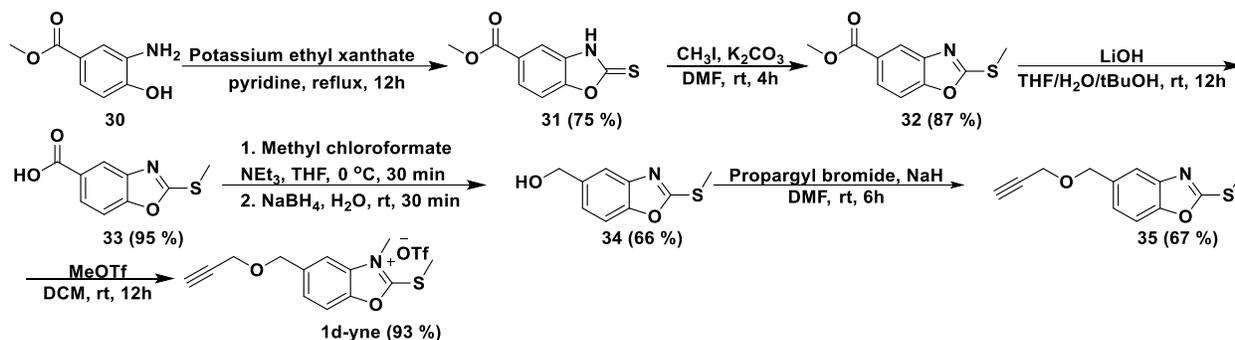
¹H NMR of compound 1c



¹³C NMR of compound 1c



Synthesis of 1d-yne



Synthesis of compound 31

Methyl 3-amino-4-hydroxybenzoate **30** (4 g, 24 mmol) was dissolved in dry pyridine (30 mL) and potassium ethyl xanthate (4.05 g, 25.27 mmol) was added. The solution was refluxed for 12 h. The solution was then cooled down to RT and treated with DI water (500 mL) and acetic acid (50 mL). The precipitate was filtered off and washed with hexane for three times. The off-white precipitate was dried under vacuum to obtain compound **31** as off-white powder (75 %, 3.8 g). $^1\text{H NMR}$ (600 MHz, DMSO- d_6) δ 7.96 (d, $J = 1.4$ Hz, 1H), 7.90 (dd, $J = 8.2, 1.6$ Hz, 1H), 7.31 (d, $J = 8.2$ Hz, 1H), 3.85 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, DMSO- d_6) δ 181.27, 165.54, 147.97, 135.43, 126.96, 124.98, 110.47, 110.31, 52.43.

Synthesis of compound 32

Methyl 2-thioxo-2,3-dihydrobenzo[d]oxazole-5-carboxylate **31** (2.35 g, 11.2 mmol), potassium carbonate (1.54 g, 11.2 mmol), and methyl iodide (700 μL , 11.2 mmol) were added sequentially into dry DMF (11.2 mL) at 0 $^\circ\text{C}$. The mixture was warmed to room temperature and stirred for 4 hours. The reaction mixture was washed with ethyl acetate and brine for 3 times. The organic portion was collected, dried over anhydrous MgSO_4 , filtered, and concentrated under reduced pressure. The residue was purified by the column chromatography (hexane: ethyl acetate 7:1) to

yield compound **32** as colorless oil (2.17 g, 87 %). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 8.12 (d, $J = 1.5$ Hz, 1H), 8.03 (dd, $J = 8.3, 1.5$ Hz, 1H), 7.61 (d, $J = 8.3$ Hz, 1H), 3.94 (s, 3H), 2.78 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 172.64, 169.30, 166.78, 151.83, 146.03, 126.49, 117.93, 111.50, 52.48, 14.73.

Synthesis of compound **33**

To a stirred solution of methyl 2-(methylthio)benzo[d]oxazole-6-carboxylate **32** (160 mg, 0.71 mmol) in THF (3.55 mL) were added t-BuOH (1.42 mL) and a solution of $\text{LiOH}\cdot\text{H}_2\text{O}$ (120 mg, 2.86 mmol) in water (1.42 mL), followed by stirring at rt for 12 h. The solvent was evaporated under reduced pressure. The mixture was extracted with ethyl acetate. The organic layer was washed with 0.5-N HCl solution, water, and brine, and dried over anhydrous MgSO_4 . After filtration, the solvent was evaporated under reduced pressure to afford compound **33** as white solid (141 mg, 95 %). Compound **33** was directly used without further purification.

Synthesis of compound **34**

To a solution of 2-(methylthio)benzo[d]oxazole-6-carboxylic acid **33** (150 mg, 0.71 mmol) in THF (10 mL) was added Et_3N (130 μL , 0.93 mmol), followed by dropwise addition of methyl chloroformate (66 μL , 0.86 mmol) at 0 °C for 30 min. The appeared salt was filtered off. To the resulting filtrate, NaBH_4 (108 mg, 2.86 mmol) in water (800 μL) was added dropwise. The reaction mixture was stirred at rt for 30 min, diluted with water and washed with ethyl acetate and brine, and dried over MgSO_4 . After filtration, the solvent was concentrated under reduced pressure. The residue was purified by silica gel column chromatography (hexane: ethyl acetate 3:1) to obtain compound **34** as white solid (92.2 mg, 66 %) $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.53

(d, $J = 0.6$ Hz, 1H), 7.46 (dd, $J = 1.6, 0.7$ Hz, 1H), 7.26 (d, $J = 8.9$ Hz, 1H), 4.78 (s, 2H), 2.76 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 172.63, 152.34, 141.54, 137.44, 123.42, 118.19, 108.62, 65.29, 14.66.

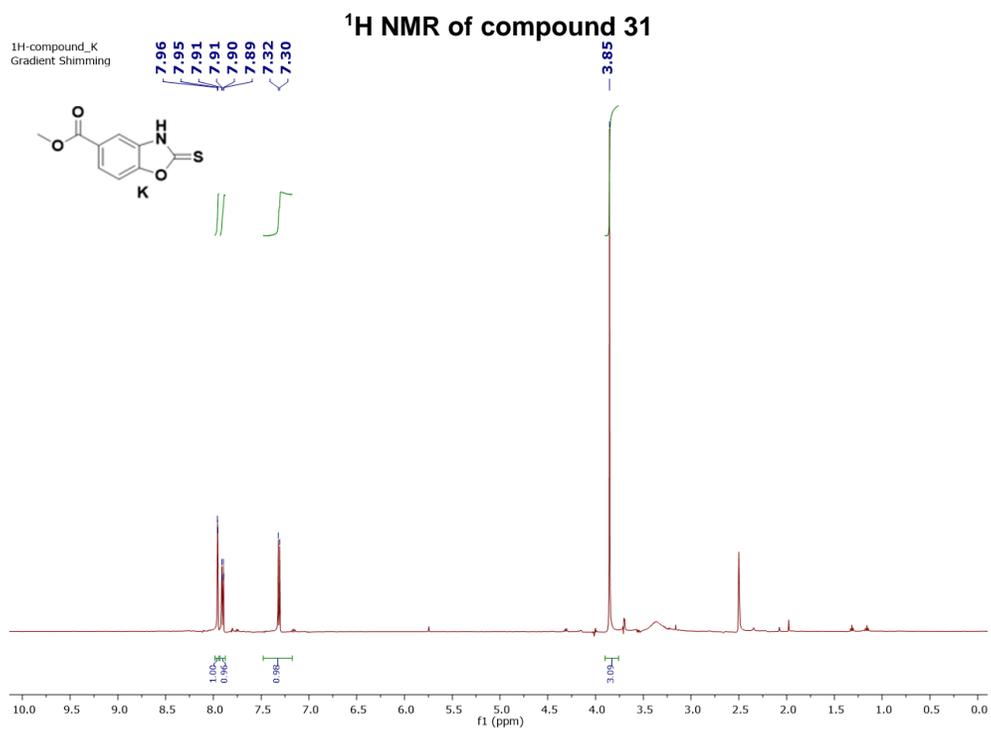
Synthesis of compound **35**

To a solution of (2-(methylthio)benzo[d]oxazol-6-yl)methanol **34** (36 mg, 0.184 mmol) in dry DMF (5 mL) at 0 °C, NaH (5.31 mg, 0.22 mmol) was added. After 10 mins propargyl bromide (20 μL , 0.22 mmol) was added dropwise and the reaction mixture was allowed to warm to room temperature. After 6 h reaction mixture was washed with ethyl acetate and brine. The organic part was dried over anhydrous MgSO_4 , filtered, and concentrated. The crude was purified by flash column (hexane: ethyl acetate 5:1) to afford compound **35** as yellow solid (28 mg, 67 %). ^1H NMR (400 MHz, CDCl_3) δ 7.59 (d, $J = 8.1$ Hz, 1H), 7.49 (dd, $J = 1.5, 0.7$ Hz, 1H), 7.31 (s, 1H), 4.72 (s, 2H), 4.21 (d, $J = 2.4$ Hz, 2H), 2.78 (s, 3H), 2.51 (s, 1H). ^{13}C NMR (101 MHz, CDCl_3) δ 166.36, 152.29, 141.90, 133.67, 124.60, 118.14, 109.78, 79.59, 74.96, 71.47, 57.19, 53.57, 14.67.

Synthesis of compound **1d-yne**

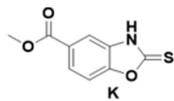
2-(methylthio)-5-((prop-2-yn-1-yloxy)methyl)benzo[d]oxazole **35** (86 mg, 0.52 mmol) and methyl trifluoromethanesulfonate (70 μL , 0.625 mmol) were mixed in 10 mL of DCM in a nitrogen environment at 0 °C. The reaction mixture was allowed to warm to 25 °C for 12 h. Precipitation with hexane was followed by filtration and yielded compound **1d-yne** as white powder (192 mg, 93 %). ^1H NMR (600 MHz, Methanol- d_4) δ 7.92 (s, 1H), 7.83 (d, $J = 8.4$ Hz, 1H), 7.68 (d, $J = 8.4$ Hz, 1H), 5.46 (dd, $J = 2.8, 1.4$ Hz, 2H), 4.77 (s, 2H), 3.95 (s, 3H), 3.31 (s,

1H), 3.08 (s, 3H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 172.04, 131.98, 123.62, 113.21, 111.03, 109.18, 108.70, 80.16, 77.47, 70.59, 56.71, 28.11, 14.71.

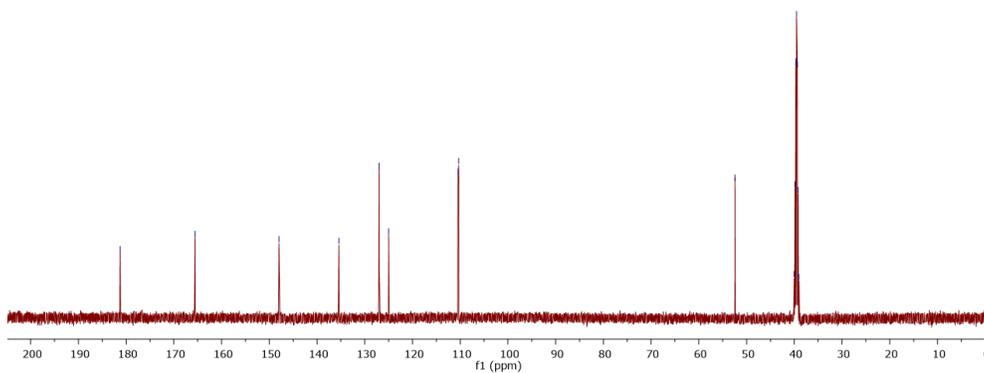


¹³C NMR of compound 31

13C-compound_K
Gradient Shimming

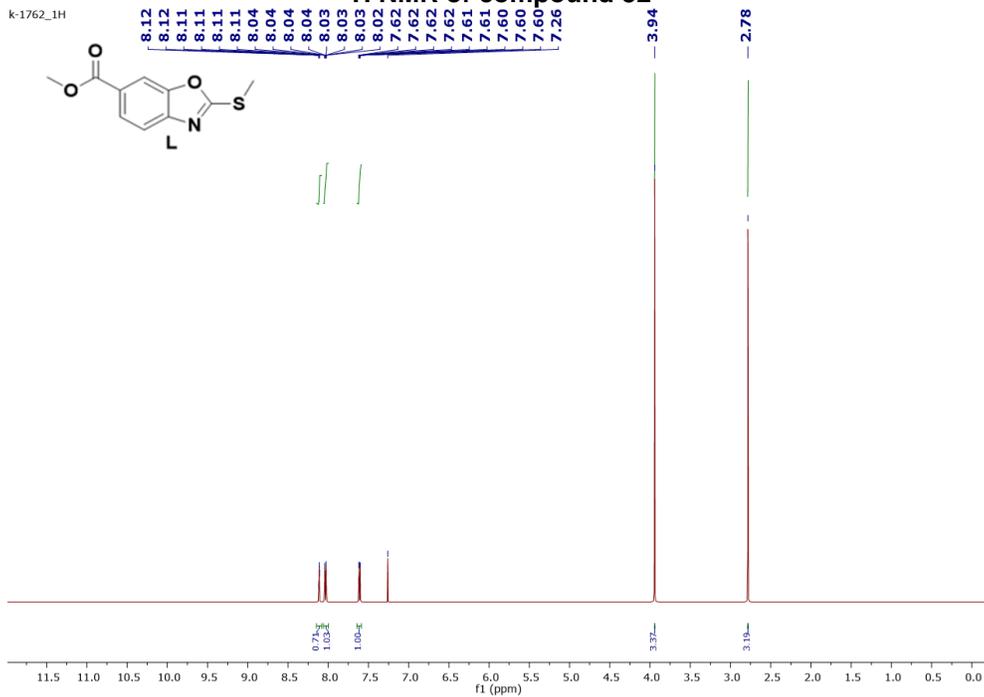
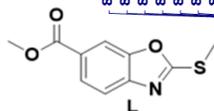


181.27
165.54
147.97
135.43
126.96
124.98
110.47
110.31
52.43
40.02
39.85
39.69
39.52
39.35
39.19
39.02



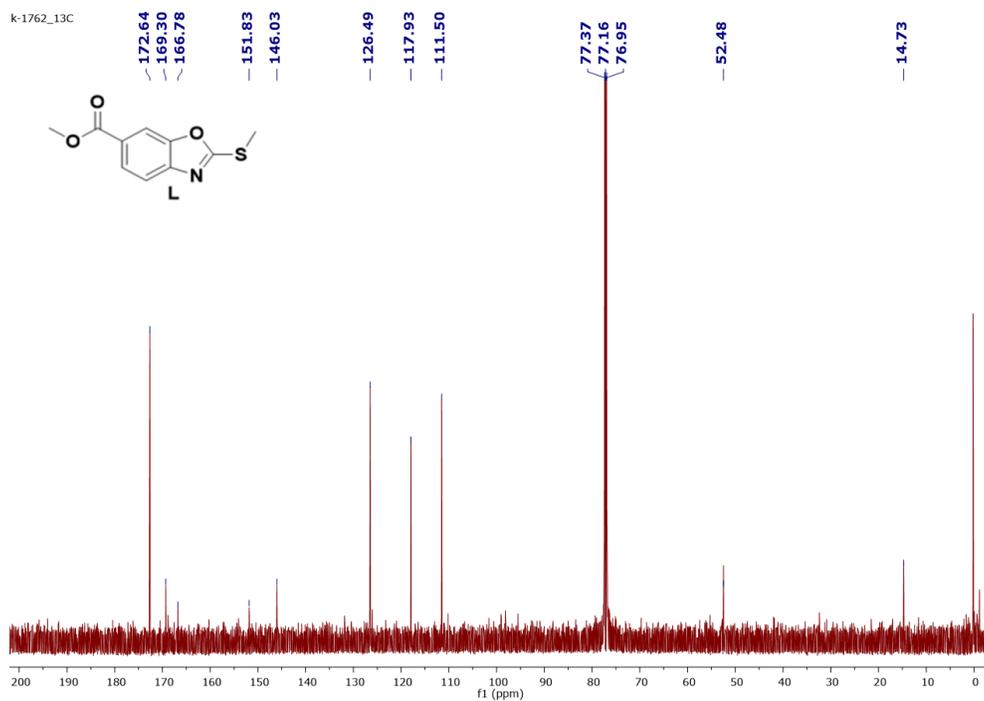
¹H NMR of compound 32

k-1762_1H



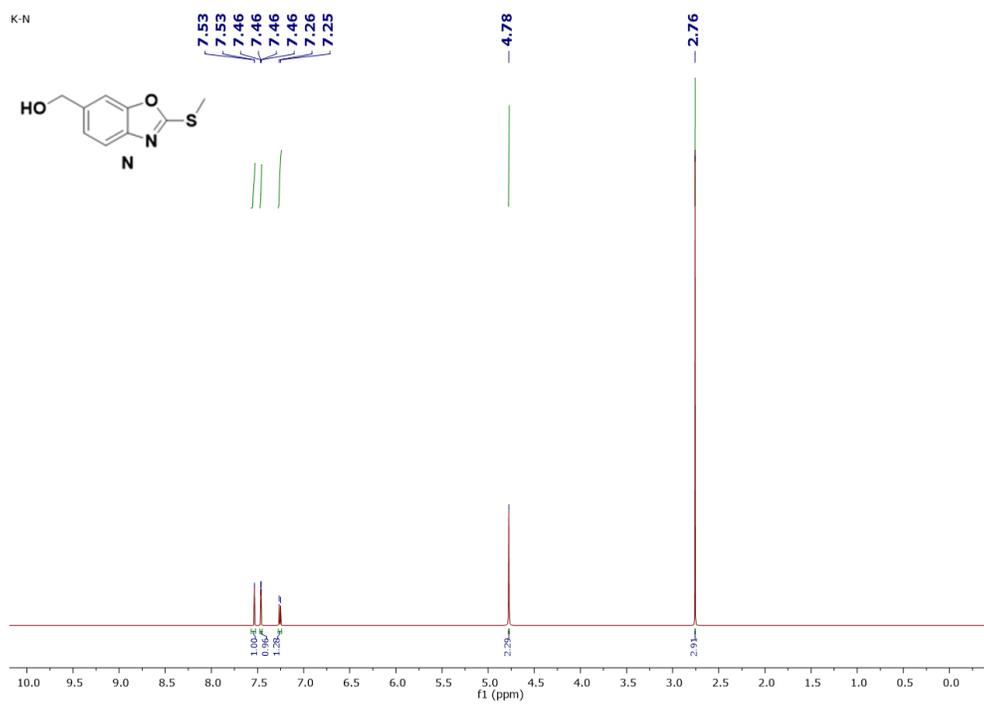
¹³C NMR of compound 32

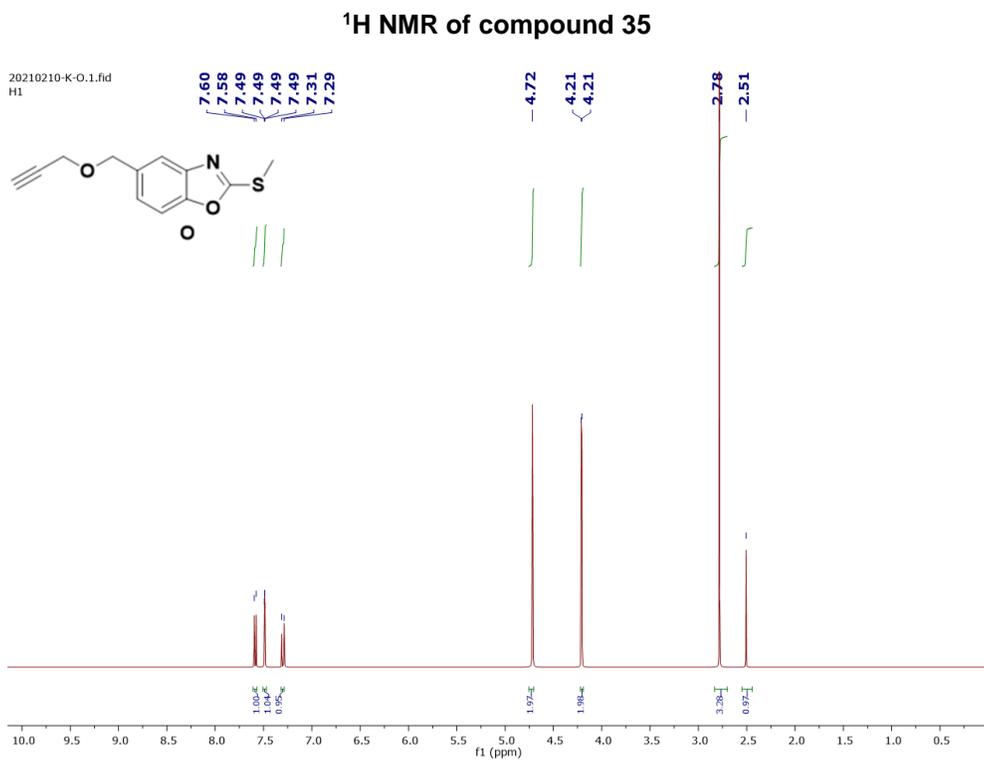
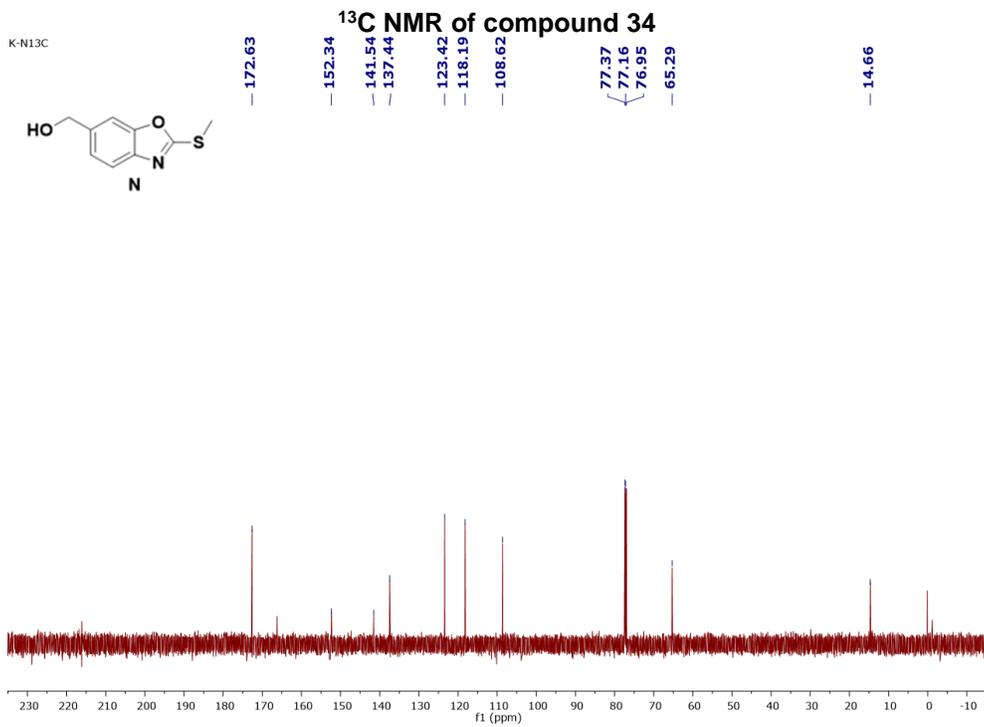
k-1762_13C

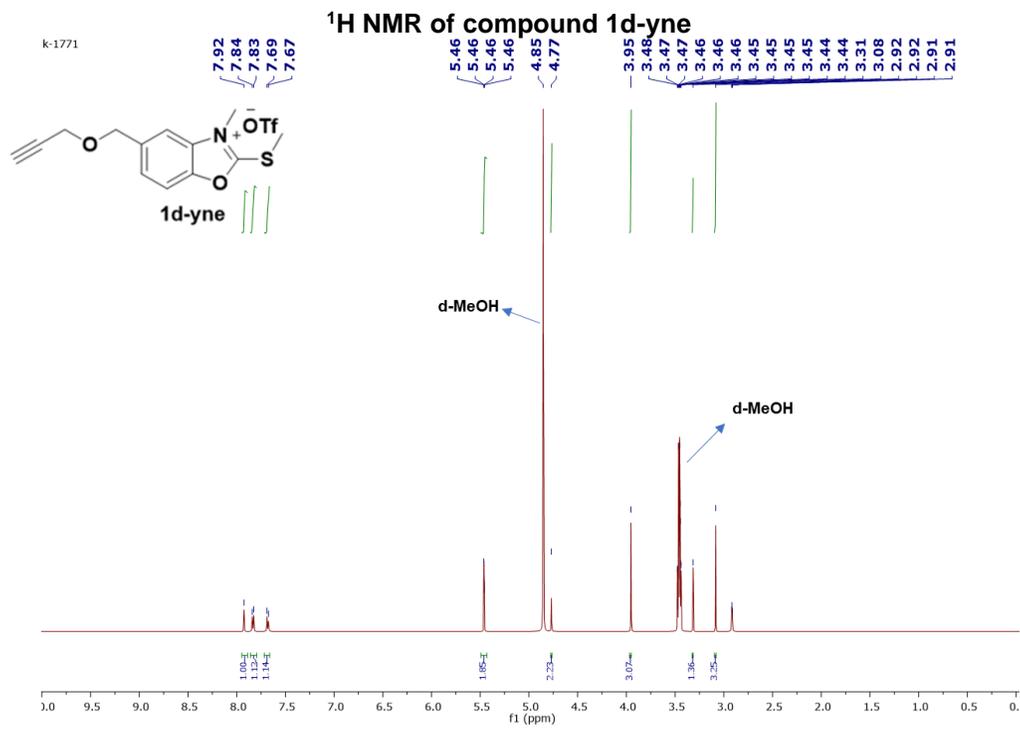
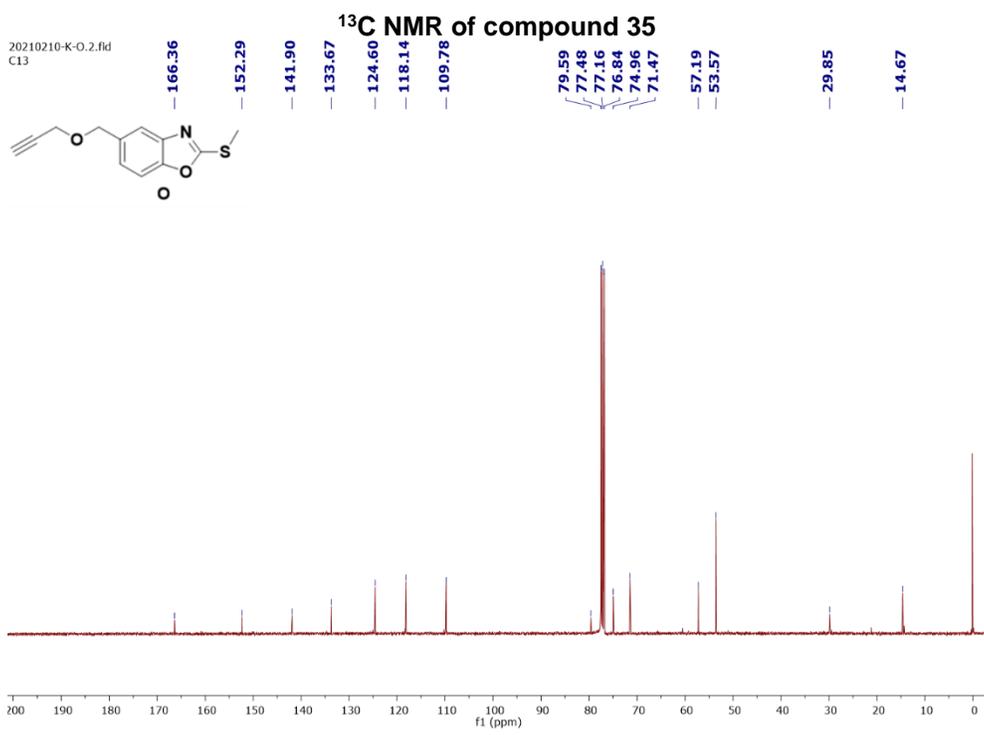


¹H NMR of compound 34

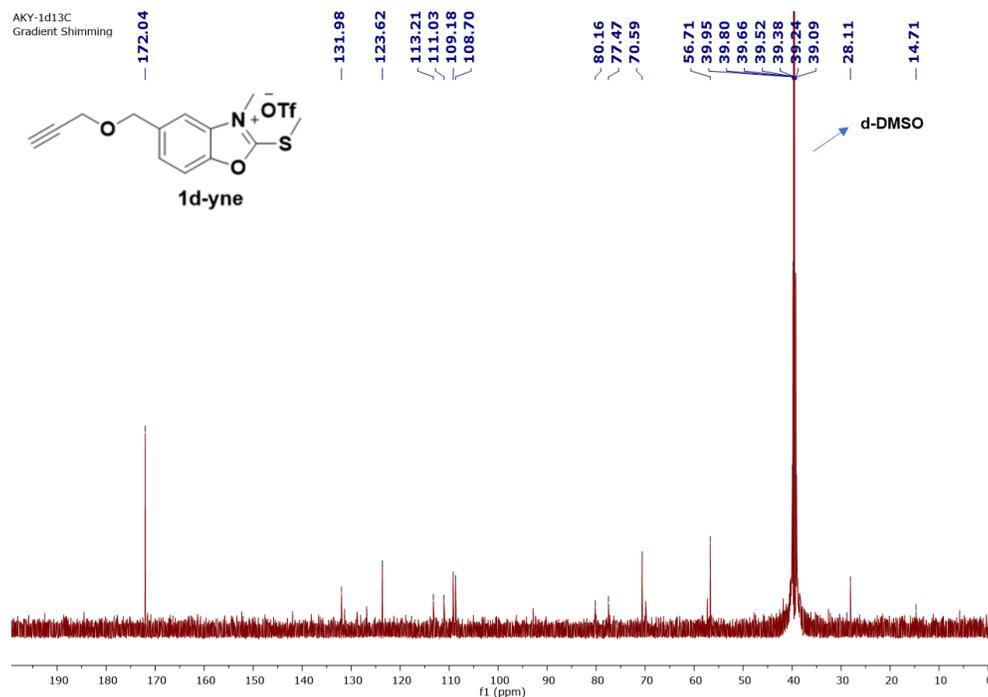
K-N



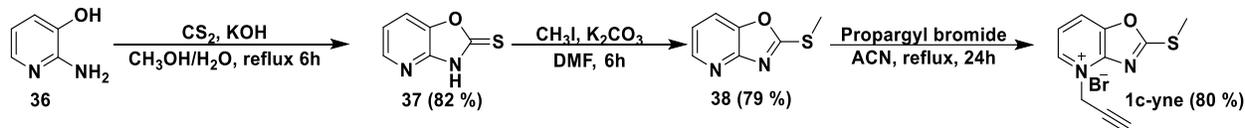




¹³C NMR of compound 1d-yne



Synthesis of 1c-yne



Synthesis of compound 37

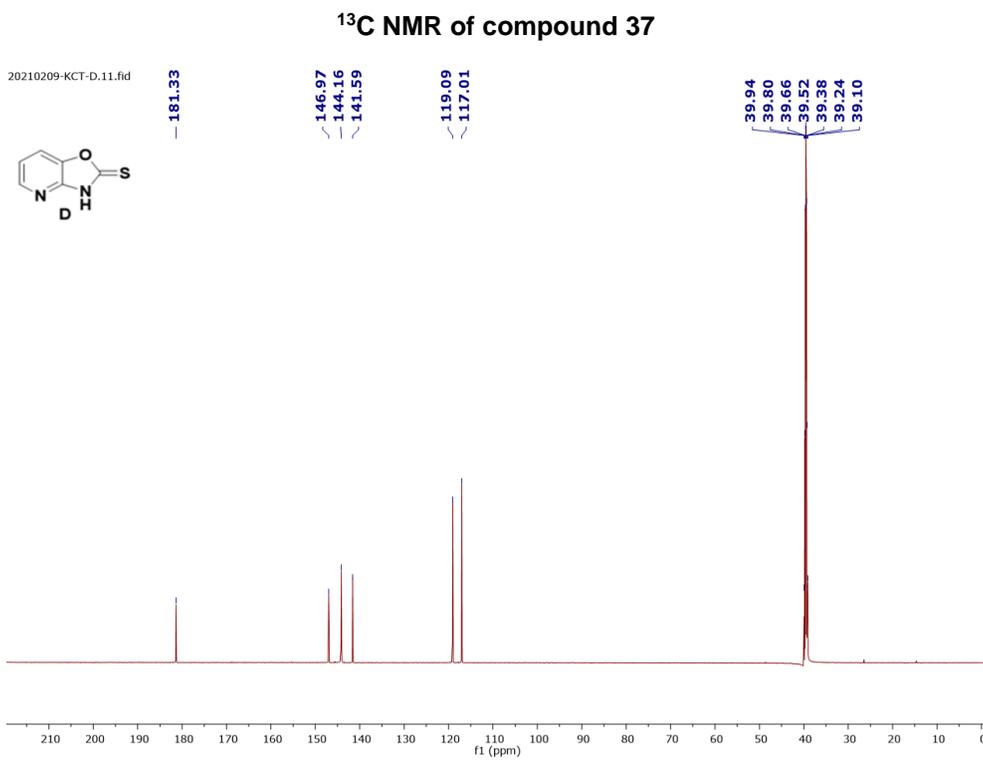
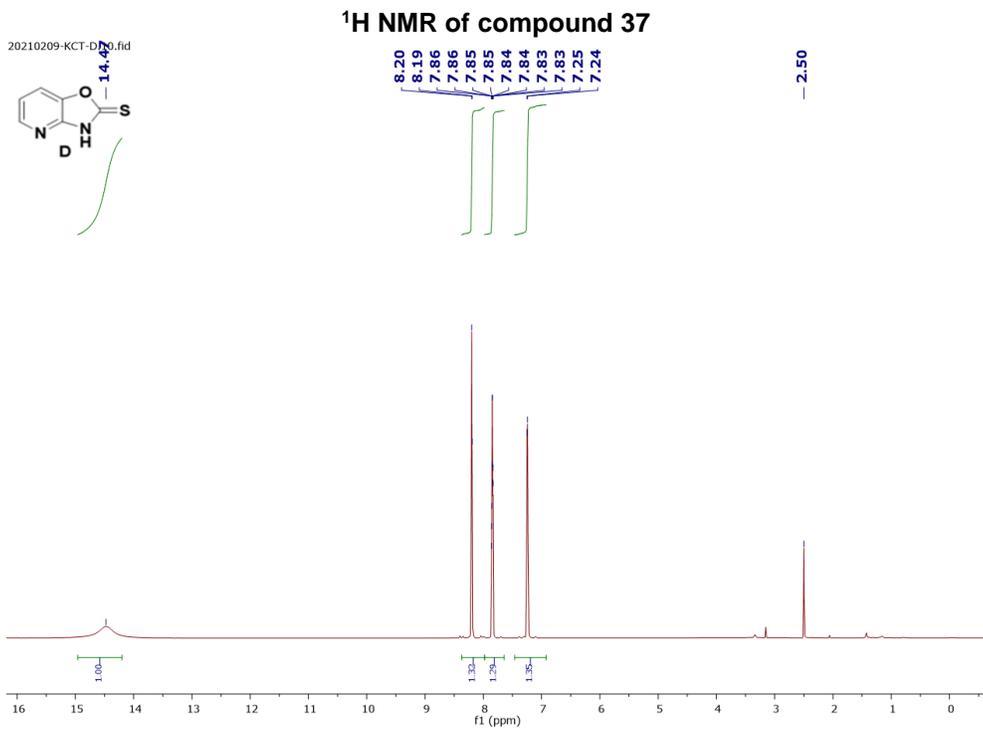
2-Amino-3-hydroxypyridine **36** (1.1 g, 10 mmol) was dissolved in methanol / water (10:1, 81 mL) and KOH (1.38 g, 10 mmol) and CS₂ (1.8 mL, 30 mmol) were added sequentially. The reaction mixture was refluxed for 6 h. The solution was then cooled to RT and treated with glacial acetic acid (5 mL) and diluted with water (150 mL). The resulting precipitate was filtered off and washed with hexane for three times to obtain compound **37** as off-white powder (82 %, 1.24 g). ¹H NMR (600 MHz, DMSO-*d*₆) δ 14.47 (s, 1H), 8.20 (d, *J* = 5.0 Hz, 1H), 7.86 – 7.83 (m, 1H), 7.25 (d, *J* = 5.3 Hz, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 181.33, 146.97, 144.16, 141.59, 119.09, 117.01.

Synthesis of compound **38**

To solution of oxazolo[4,5-b]pyridine-2(3H)-thione **37** (500 mg, 3.29 mmol) in dry DMF (12 mL) was added K₂CO₃ (454 mg, 3.29 mmol), MeI (206 μ L, 3.29 mmol) sequentially at 0 °C, then the reaction was warmed to room temperature and stirred for 6 h. The reaction mixture was washed with ethyl acetate and brine for three times. The organic layer was collected and dried over anhydrous MgSO₄, filtered, and concentrated. The crude was purified by flash chromatography (hexane: ethyl acetate 5:1) to afford compound **38** as white powder (79 %, 431 mg). ¹H NMR (600 MHz, CDCl₃) δ) δ 8.46 – 8.44 (m, 1H), 7.70 – 7.69 (m, 1H), 7.18 (dd, *J* = 8.0, 5.0 Hz, 1H), 2.81 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.35, 156.22, 145.89, 144.25, 118.86, 117.08, 14.82.

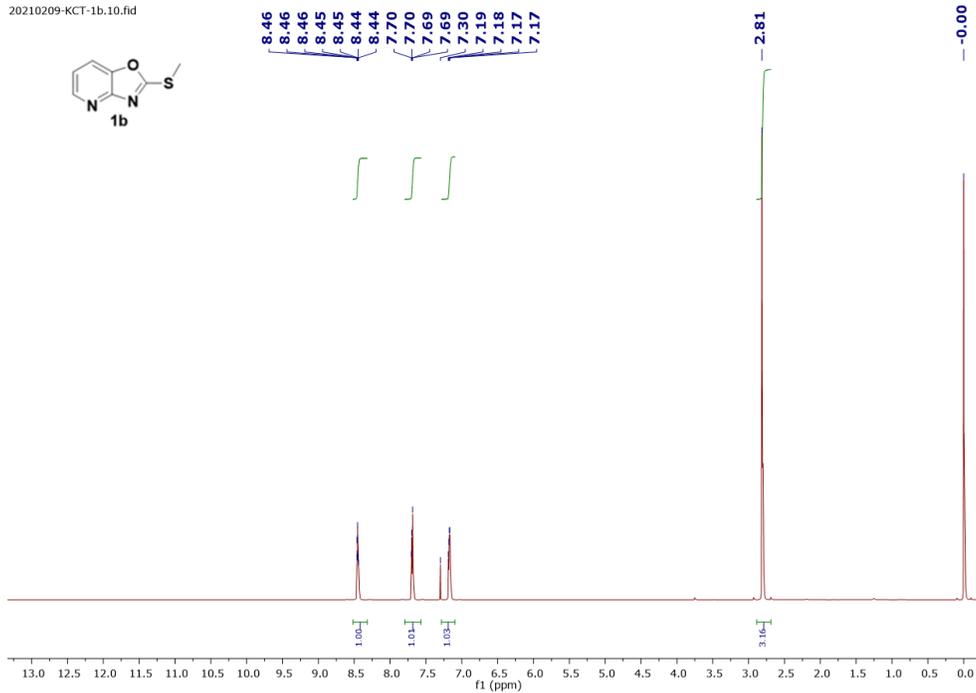
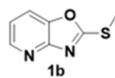
Synthesis of compound **1c-yne**

2-(Methylthio)oxazolo[4,5-b]pyridine **38** (83 mg, 0.5 mmol) was dissolved in dry ACN (3 mL), and treated with propargyl bromide (600 μ L, 5 mmol). The reaction mixture was refluxed for 24 h. The mixture was cooled to room temperature and ACN was removed by rotary evaporation. The crude was purified by recrystallization with ether to afford compound **1c-yne** as brown powder (80%, 246 mg). ¹H NMR (600 MHz, DMSO-*d*₆) δ 9.01 – 8.99 (m, 1H), 8.89 (dd, *J* = 8.8, 4.5 Hz, 1H), 8.00 – 7.97 (m, 1H), 5.71 (s, 2H), 3.99 (s, 1H), 2.95 (s, 3H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 172.02, 150.31, 147.93, 137.91, 126.18, 121.09, 80.54, 75.00, 44.04, 15.02.



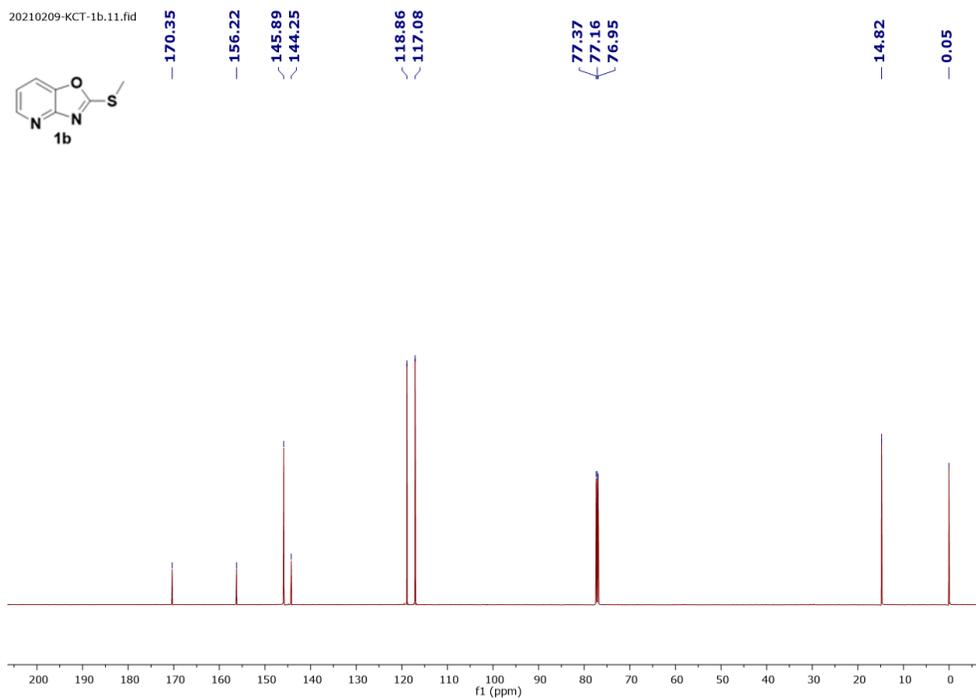
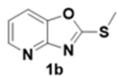
¹H NMR of compound 38

20210209-KCT-1b.10.fid

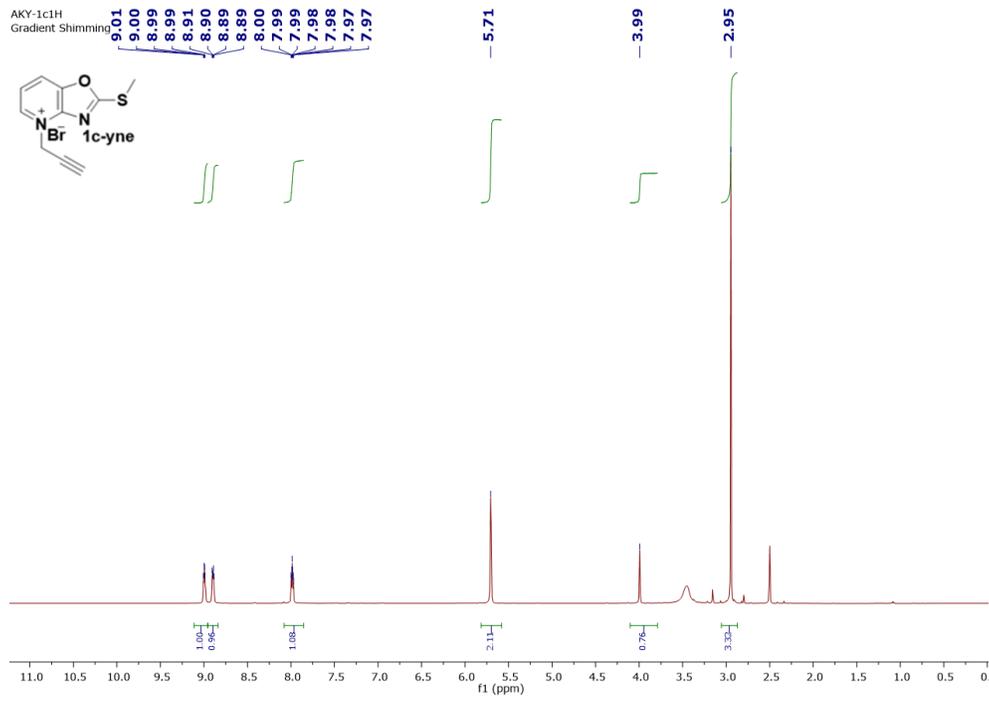


¹³C NMR of compound 38

20210209-KCT-1b.11.fid



¹H NMR of compound 1c-yne



¹³C NMR of compound 1c-yne

