NEW REAGENTS AND METHODS FOR MASS SPECTROMETRY-BASED PROTEOMICS INVESTIGATIONS

Thesis by

Chang Ho Sohn

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2011

(Defended May 9, 2011)

Acknowledgements

I dedicate this thesis to my parents. Like many other Korean parents they have fully supported my higher education endeavors, even despite their own limited educational opportunities. Their love and endless sacrifice has made this miracle possible.

I wholeheartedly thank my adviser and mentor Jack Beauchamp. His endless passion and curiosity for science is an inspiration to my own intellectual motivation. I especially appreciate his laissez-faire approach that offered me the time and latitude to individually develop my ideas during graduate school.[1] He has taught me (indirectly) that, "Always, inspiration and imagination come from the slack, not from the diligence." I would also like to thank my committee members Geoffrey Blake, James Heath, and Douglas Rees, along with former committee member Linda Hsieh-Wilson, for all of their support and encouragement.

I was fortunate to have many fantastic collaborators on this project. I was greatly blessed by Caltech's interdisciplinary research environment and am especially indebted to three brilliant Caltech scientists. Cheol K. Chung who was a former postdoc in the Grubbs lab, was my best organic chemistry teacher ever. Before he trained me, I had been a typical chemical physics graduate student who mostly tunes knobs for high-voltage power supply and laser beams. Synthetic experiments with him have changed my career in chemistry a lot. As an insightful biochemist/biologist, J. Eugene Lee inspired me to develop two novel chemical reagents: isobaric tags and cross-linkers that are both involved with click reactions. Heather Agnew, an expert in click chemistry, helped me to optimize the downstream sample treatment for cross-linked and

[1] Scientific Autobiography of Jack Beauchamp. *J. Phys. Chem. A* **2002,** *106* (42), 9625-9628.

clicked peptides. I hope that our work will become the cornerstone for future applications of click reactions in mass spectrometry-based proteomics.

The PEL fellows, Mike, Bobby, Geoff, Natasa, and Sonja, have been highly supportive in our collaborations. Many technical challenges during the LC-MS sample analyses have been solved by these wonderful scientists.

Current and past JLB members, Daniel, Evan, Hugh, and Tae-Young, were a great team to share and bounce off ideas. Truly, each of you was most helpful to me.

Joe Loo at UCLA has been such a great colleague and friend by sharing his novel ideas and state-of-the-art mass spectrometers for our collaborative work. Of course, our work would not have been made possible without the contributions of his mentees, Sheng, Prasanna, Ivory, and Gregg.

To my Korean folks, Christopher, Sanghyun, Sukwon, Hyungjun, Melanie, Narae, Young-In, Boram, Youngshik, Kiwook, Sinchul, Haekong, Andrew and Oh-Hoon at Caltech and Sangwoo, Dann, Sunghwan and Sangjin at Harvard—I will always cherish the fond memories you have given me. Thanks guys!

To my loving girlfriend and future wife, Lupita Hyung Ryeol Kim, whom I deeply thank for her patience and support. What a lucky guy I am!

Finally, I want to acknowledge a generous fellowship from Kwanjeong Educational Foundation for my Ph.D study abroad. I sincerely appreciate their supports.

Abstract

New chemical reagents and methods have been employed for mass spectrometry (MS)-based proteomics investigations. Many chemical reagents are synthesized to be covalently attached to biomolecules, especially peptides and proteins. The properties of the resulting peptide conjugates are characterized by various tandem mass spectrometric techniques (*e.g.*, collision-induced dissociation (CID), electron capture dissociation (ECD), electron transfer dissociation (ETD), infrared multiphoton dissociation (IRMPD), and free radical initiated peptide sequencing (FRIPS)). In Chapter 2, the effect of high electron affinity tags in ECD and ETD is investigated using their peptide conjugates. The initial intramolecular electron transfer from the high-lying Rydberg states to the covalently attached high electron affinity tag occurs in competition with the Coulomb stabilized π^* orbitals of the amide bonds in the model peptides. This leads to the inhibition of the normal sequence of ECD and ETD processes, yielding no backbone fragmentations. In Chapter 3, selective disulfide bond cleavages are observed by the FRIPS method. A newly prepared TEMPO-based FRIPS reagent is labeled to model peptides containing disulfide bonds and subject to CID to monitor free radical induced cleavages. Highly selective C–S and S–S bond cleavages are observed and their reaction mechanisms are proposed. In Chapter 4, novel Caltech isobaric tags (CITs) for protein quantification are developed and validated using various model samples. A newly discovered low-energy gas-phase fragmentation pathway, a nucleophilic substitution of the N3 in the 1,2,3-triazole ring generated by copper-catalyzed azide-alkyne cycloaddition (CuAAC) inspired us to create CITs. This selective cleavage is applied to the formation of the reporter ions to quantify protein expression level in cells. Chapter 5 describes clickable cross-linkers (CXLs) developed for elucidation of three-dimensional protein structures

and protein-protein interactions (PPIs). In CXLs, cross-linking reactions are separated from the conjugation of affinity tags, avoiding steric hindrance. Cross-linked peptides are enriched from the complex mixture of yeast lysate and cross-linked ubiquitin digests using avidin affinity chromatography, showing high sensitivity of the CXL-based analysis. The low-energy pathway used for CIT reagents is also adopted to produce the reporter ion, filtering MS/MS scans of cross-linked peptides from those of unmodified peptides.

Table of Contents

Table of Figures

Table of Schemes

List of Tables

Table of Equations

1. Introduction

1.1. Background

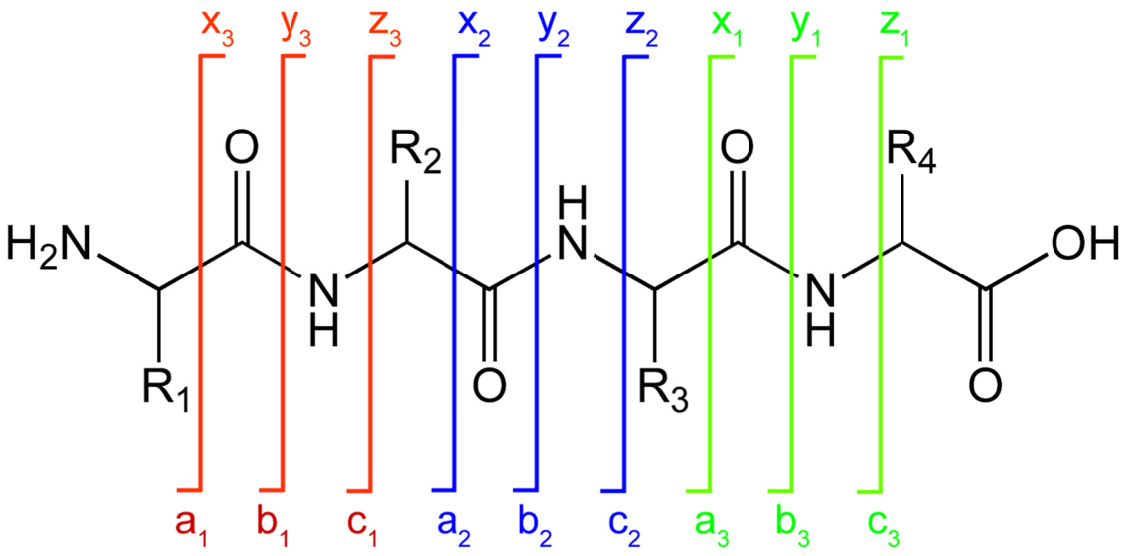
Development of two soft ionization methods, matrix-assisted laser desorption/ionization (MALDI)[1,2] and electrospray ionization (ESI)[3] has opened up the era of mass spectrometry (MS)-based proteomics. Using these soft ionization techniques, molecular weights of biopolymers are easily determined and the structures and reaction dynamics of biomolecular ions in the gas phase can be investigated by mass spectrometers.

After the ionization of analytes, tandem mass spectrometry (MS/MS) is used for isolation and fragmentation of precursor ions of interest.[4] By examining fragment ions, the structure and reactivity of the precursor ion can be studied. If energy deposition from the step of ionization is enough to proceed for fragmentation, the precursor ion undergoes metastable ion decomposition (MID).[5] For efficient fragmentation, the internal energy of the precursor ion needs to be increased. In most unimolecular decomposition reactions, the acquired internal energy is statistically distributed to vibrational modes of the precursor ion via internal conversion. When the threshold energy for a certain reaction is reached, bond cleavages occur. Based on the Rice-Ramsberger-Kassel-Marcus (RRKM)[6]–Quasi Equilibrium Treatment,[7] the kinetic constant is a function of the internal energy. Therefore, it is critical to define the internal energy of the precursor ion for prediction of the results of unimolecular decomposition reactions.

In the early stage, due to the lack of proper ion activation techniques, MS/MS experiments were performed using MID that requires no special setups. For peptide sequencing, post-source

decay (PSD) of peptide ions generated by MALDI is conducted using reflectron time-of-flight (TOF) mass spectrometers.

After technical improvements, collision-induced dissociation (CID) or collisionally activated dissociation (CAD) was introduced and is still implemented in contemporary mass spectrometers. By inelastic collision with buffer gases (N_2, Ar or He), the translational energy of buffer gas molecules are transferred to the precursor ions, augmenting their internal energy. In the low energy CID that is widely used in MS-based proteomics platforms, ion activation occurs via multiple collisions resulting in slow heating of the precursor ion. Upon low energy collisional activation, peptide ions undergo mainly amide backbone cleavages, yielding b- and y-type ions (**Scheme 1.1**).[8] The mobile proton theory was proposed to explain peptide fragmentation by collisional activation.[9]



Scheme 1.1

In CID, peptide fragmentation patterns highly depend on the sequence and the charge state of the peptide ion. Often, selective bond cleavages are observed (*e.g.*, the C-terminal sides of aspartic/glutamic acids[10,11], and the N-terminal side of proline[12]), which is especially useful for *de novo* sequencing. Yet these features can yield poor sequencing results due to the suppression of other competing reaction channels. Especially, chemical bonds in post-translational modifications (PTMs) are preferentially fragmented prior to those of backbones, leading to the loss of their connectivity.

To address the problems described above, alternative ion activation method, electron capture dissociation (ECD)[13] and its variation, electron transfer dissociation (ETD)[14] were developed. Multiply charged peptide or protein ions generated by ESI are reacted with an electron or anionic reagent ions, resulting in charge-reduced precursor ions. The recombination energy gained by electron capture or transfer is redistributed to vibrational modes via internal conversion. Ultimately, backbone amide bonds in the charge-reduced ion are cleaved, leading to the formation of c- and z-type ions. Unlike CID, ECD and ETD less suffer from sequence discrimination and, most importantly, preserve labile bonds, enabling successful sequencing of PTMs. In spite of their usefulness, however, the underlying mechanisms have been debated since their inventions. In Chapter 2, we investigate the effect of high electron affinity tags in ECD and ETD of peptides. By labeling tags with various electron affinities, their effects on the initial electron capture kinetics and subsequent intramolecular electron transfer, followed by proton transfer, are studied.

In response to the inventions of electron-based dissociation techniques, Hodyss *et al.* reported an alternative method for free radical initiated peptide sequencing (FRIPS).[15] The water soluble, commercially available free radical initiator, Vazo 68, is conjugated to peptides and its

bioconjugates are subject to CID. The selective gas-phase fragmentations are observed via abstraction of the alpha- or beta-hydrogen, followed by β-cleavage. The resulting fragment ions are mainly a- and x-type ions along with neutral losses, and some c- and z-type ions that are similar to those observed in ECD and ETD. FRIPS does not require multiply charged precursor ions and specialized instruments for ion activation. Also, phosphorylated residues are preserved, holding a promise for its applicability in PTM analyses. In Chapter 3, we extend our FRIPS method to the analysis of disulfide bonds in peptides. The model peptides are labeled by a newly prepared second generation 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO)-based FRIPS reagent. The peptide conjugates are analyzed by collisional activations. After loss of TEMPO, the regioselective acetyl radical is formed, followed by highly selective C–S and S–S cleavages. The mechanisms of those processes are investigated and further application of FRIPS for disulfide analysis in proteins is discussed.

Recently, quantitative MS-based proteomics has been utilized to monitor the relative and absolute protein expression in cells.[16,17] By comparing the level of protein expression in various cell conditions, one can study the functions of individual proteins and their interactions with other proteins. As a popular chemical labeling approach, tandem mass spectrometry–based isobaric tags have been employed for protein quantification.[18,19] Yet, the high cost of the commercially available isobaric reagents hinders their wide usages. In Chapter 4, novel isobaric tags that are easy and cheap to synthesize are described. These new reagents were inspired by the observation of the low energy fragmentation pathway triggered by a nucleophilic substitution of N3 nitrogen in the 1,2,3-triazole ring.

Chemical cross-linkers coupled with MS have been applied to elucidation of three-dimensional protein structure and protein-protein interactions (PPIs).[20-24] The detection of cross-

linked peptides from enzymatic digestion of cross-linked protein samples is, however, still a challenging task due to their substoichiometric quantities in mixtures of abundant unmodified peptides. Numerous functionalized cross-linkers have been developed to facilitate selective and sensitive identification of cross-linked peptides. In Chapter 5, we report clickable chemical cross-linkers (CCXLs) that afford significant advantages over other reagents. Ubiquitin, a small model protein, is used to validate the applicability of CCXLs in studies of protein structures and PPIs.

1.2.Contents of Thesis

1.2.1. Investigation of Ion Activation Methods

1.2.1.1. The Mechanisms of Electron Capture Dissociation and Electron Transfer Dissociation

In Chapter 2, electron capture dissociation (ECD) and electron transfer dissociation (ETD) of doubly protonated electron affinity (EA)-tuned peptides were studied to further illuminate the mechanism of these processes. For this purpose, a series of electron affinity tuning tags were synthesized. The initially captured electron to high-*n* Rydberg states of the doubly charged peptide ion undergoes through-space or through-bond electron transfer to the EA-tuning tags or low-*n* Rydberg states via potential curve crossing in competition with transfer to the amide π^* orbital. This interrupts the normal sequence of events in ECD or ETD leading to backbone fragmentation by forming a stable radical intermediate. The implications which these results have for previously proposed ECD and ETD mechanisms are discussed.

1.2.1.2. The Mechanism of Disulfide Cleavage by Acetyl Radical

Chapter 3 describes the mechanism of disulfide bond cleavage in gaseous peptide ions triggered by a regiospecific covalently attached acetyl radical. We describe a second generation 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO)-based free radical initiated peptide sequencing (FRIPS) reagent that yields acetyl radical peptide ions in a single step of collisional activation. Highly selective C−S and S−S bond cleavages are observed from collisional activation of FRIPS conjugates containing disulfide bonds. To probe their mechanisms, model peptides having a deuterated β-carbon at the disulfide bond are employed. It is suggested that the major pathway of S−S bond cleavage is triggered by hydrogen abstraction from the α-carbon, followed by radical substitution to the S−S bond, yielding thiirane and thiyl radical products. A minor contribution by direct radical substitution to the disulfide bond is also considered. Density functional theory calculations are performed to explore energetics of the proposed mechanisms for disulfide bond cleavage.

1.2.2. Protein Quantification and Structural Studies

1.2.2.1. Caltech Isobaric Tags for Protein Quantification

In Chapter 4, versatile and modular isobaric tags for protein quantification, referred to as Caltech Isobaric Tags (CITs) are reported. CIT is based on a newly discovered fragmentation pathway, a nucleophilic substitution of N3 nitrogen in the 1,2,3-triazole ring. Unlike commercially available isobaric tags (*e.g.*, iTRAQ and TMT), the number of isobaric combinations of CIT reagents, in theory, is unlimited. The sizes of reporter ions can be easily tuned to avoid the low mass–cut off problem in ion trap mass spectrometers. CITs can also be prepared with relatively low expense and less effort. CIT is applied to model systems and heavy

to light ratios of the CIT reporter ions show excellent linear responses within a two-orders-of-magnitude dynamic range, agreeing with initial mixing ratios.

1.2.2.2. Clickable Cross-Linkers for Elucidation of Protein Structures and Protein-Protein Interactions

In the last chapter, we present novel clickable cross-linkers (CXLs) for sensitive and selective detection of cross-linked peptides from complex mixtures. CXLs are homobifunctional amine-reactive cross-linkers composed of a central tertiary amine connected to a terminal alkyne. CXLs hold several distinct advantages over other currently available amine-reactive cross-linkers. Ubiquitin, a small and lysine-abundant protein, is used here as a model system for protein structural study using CXLs. The observed cross-linked lysine residues and their connectivities are consistent with the high resolution X-ray crystal structure. The application of CXLs in more complex systems (*e.g.*, *in vivo* cross-linking) is also tested using HEK293 cells, showing good cell-permeability and water solubility.

1.3. Conclusion

New chemical reagents and methods have been employed for mass spectrometry (MS)-based proteomics investigations. Both liquid- and gas-phase properties of peptides of interest are modified by incorporation of various chemical reagents and their changes are investigated by mass spectrometry. Selective and orthogonal labeling reactions allow us to perform sensitive detection of target molecules, expediting mass spectrometry–based analyses of proteomes. Especially, the application of copper-catalyzed azide-alkyne cycloaddition (“click” reaction) to mass spectrometry-based proteomics would be invaluable.

2. Probing the Mechanism of Electron Capture and Electron Transfer Dissociation Using Tags with Variable Electron Affinity

(Reproduced in part with permission from Sohn, C.H.; Sohn, C. H.; Chung, C. K.; Yin, S.; Ramachandran, P.; Loo, J. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2009**, *131*, 5444.

Abstract Electron capture dissociation (ECD) and electron transfer dissociation (ETD) of doubly protonated electron affinity (EA)-tuned peptides were studied to further illuminate the mechanism of these processes. The model peptide FQpSEEQQQTEDELQDK, containing a phosphoserine residue, was converted to EA-tuned peptides via β-elimination and Michael addition of various thiol compounds. These include propanyl, benzyl, 4-cyanobenzyl, perfluorobenzyl, 3,5-dicyanobenzyl, 3-nitrobenzyl and 3,5-dinitrobenzyl structural moieties, having a range of EAs from –1.15 to 1.65 eV, excluding the propanyl group. Typical ECD or ETD backbone fragmentations are completely inhibited in peptides with substituent tags having EA over 1.00 eV, which are referred to as electron predators in this work. Nearly identical rates of electron capture by the dications substituted by the benzyl (EA = –1.15 eV) and 3-nitrobenzyl (EA = 1.00 eV) moieties are observed, which indicates the similarity of electron capture cross sections for the two derivatized peptides. This observation leads to the inference that electron capture kinetics are governed by the long range electron-dication interaction and are not affected by side-chain derivatives with positive EA. Once an electron is captured to high-*n* Rydberg states, however, through-space or through-bond electron transfer to the EA-tuning tags or low-*n* Rydberg states via potential curve crossing occurs in competition with transfer to the amide π^* orbital. The energetics of these processes are evaluated using time-dependent density functional theory with a

series of reduced model systems. The intramolecular electron transfer process is modulated by structure-dependent hydrogen bonds and is heavily affected by the presence and type of electron withdrawing groups in the EA-tuning tag. The anion radicals formed by electron predators have high proton affinities (approximately 1400 kJ/mol for the 3-nitrobenzyl anion radical) in comparison to other basic sites in the model peptide dication, facilitating exothermic proton transfer from one of the two sites of protonation. This interrupts the normal sequence of events in ECD or ETD leading to backbone fragmentation by forming a stable radical intermediate. The implications which these results have for previously proposed ECD and ETD mechanisms are discussed.

2.1. Introduction

Following the development of electron capture dissociation (ECD) of multiply protonated peptide or protein ions,[13] numerous studies have been carried out to investigate the mechanism of this process and to explore its broad applicability to mass spectrometry (MS)-based structural studies of peptides and proteins.[25-31] Unlike collision-induced dissociation (CID)[32-34] or infrared multiphoton dissociation (IRMPD),[35,36] ECD and its analogue, electron transfer dissociation (ETD),[14] generate abundant sequence ions and the sites of peptide backbone cleavage are relatively less discriminated by the side-chains of nearby amino acids. These methods also preserve labile side-chains with post-translational modifications (PTMs), allowing easier identification and localization of PTMs compared with CID or IRMPD.[37,38] While ECD and ETD preferentially cleave a disulfide bond, thermal activation methods (CID and IRMPD) do not generate abundant C−S or S−S bond cleavage fragments unless peptides are cationized by metal ions.[39,40] This makes ECD and ETD methods of choice for characterizing phosphorylation,[41-44]

glycosylation,[45-51] methylation[52] and disulfide linkage[53] of proteins to elucidate important biological processes such as cell signaling and cell differentiation and proliferation. Owing to recent instrumental developments, ECD and ETD have been successfully implemented to various mass analyzers such as the linear ion trap,[14] hybrid quadrupole-Time-Of-Flight (QqTOF),[54,55] Fourier transform ion cyclotron resonance (FTICR)[56] and, most recently, orbitrap[57-59] instruments. These developments satisfy the varying requirements of a wide range of applications where resolution, sensitivity, dynamic range and compatibility with various chromatographic methodologies are important parameters to consider for the mass spectrometric analyses of biological samples of ever increasing complexity.

Since its conception, however, ECD has elicited lively discussions in the mass spectrometry community with regard to its mechanism. Initial electron capture to high-*n* Rydberg states was first proposed by McLafferty and coworkers.[13,25,53,60] In this model, the protonation sites (*i.e.*, protonated amine, guanidine or imidazole residues) of a peptide ion are believed to be internally solvated by amide oxygens via one or more hydrogen bonds. Electron localization occurs to one of the positively charged sites, which subsequently forms a hypervalent radical in the ground electronic state via internal conversion, with the energy released in this process contributing to the overall vibrational excitation of the ion. Subsequent transfer of a hydrogen atom to an amide oxygen facilitates β-cleavage of the adjacent N−C_α bond through an aminoketyl radical intermediate. The resulting fragments are the residues of the peptide N-terminus and C-terminus, denoted as *c* and $z^\bullet$ ions, respectively. This process, referred to as the Cornell mechanism,[61] was initially suggested to be a non-ergodic reaction.[26] The preservation of non-covalent interactions along with backbone cleavages was demonstrated as a proof of non-ergodicity in ECD.[62,63] Supportive theoretical and experimental observations for the Cornell mechanism were subsequently reported elsewhere.[64-66]

Even though the Cornell mechanism provided a reasonable picture for ECD, some backbone fragmentations were not easily explained.[67] The characteristic ECD fragmentation processes are still observed in some peptide cations where electron capture does not yield a mobile hydrogen atom. These include peptides cationized by metal ion attachment[68] or fixed charge derivatives (*i.e.*, quaternary ammonium or phosphonium groups).[69,70] In addition, the guanidinium groups in peptides are poorly solvated by amide oxygens and hydrogen atom transfer from an arginine radical to an amide carbonyl is endothermic.[61,71] With either of these circumstances, *c*- or *z*-type ions are still prominent in ECD spectra.[61]

The Utah-Washington mechanism[61] (UW mechanism), recently proposed independently by Simons and coworkers[72-80] and Turecek and coworkers,[61,69,71,81-90] provides an alternative view of the mechanism explaining the relatively indiscriminate distribution of N−C_α bond cleavage processes observed in ECD and ETD. Coulomb stabilization by positively charged groups allows the amide π^* orbital to possess a positive electron affinity (EA).[91] Electron attachment to Coulomb stabilized amide π^* orbitals makes the amide group an exceptionally strong base with a proton affinity (PA) in the range 1100-1400 kJ/mol.[84] The amide anion radical is able to abstract a proton in an energetically favorable process via conformational changes, even from relatively distant proton donors. The resulting intermediate is identical to the aminoketyl cation radical proposed in the Cornell mechanism and can undergo the same N−C_α bond cleavage. This process does not require invoking either the mobile "hot" hydrogen atom hypothesis or non-ergodicity of dissociation. ECD of multiply cationized ions where the charge carriers are metal ions or fixed charge derivatives can also be explained by ion-dipole interactions and the intramolecular electron transfer between the charge-stabilized amide π^* orbital and the N−C_α σ^* orbital, followed by N−C_α bond cleavage. The UW mechanism is supported by recent theoretical and experimental investigations.[82,83,92,93]

Despite many efforts of the past decade, there is still much to be learned about the mechanistic details of ECD and ETD. The sizes of peptides or proteins are too large to accurately quantify the energetics of these processes based on high level *ab initio* or density functional calculations. Recently, Williams and coworkers quantified the energetics of the ECD process involving a hydrated gaseous peptide dication by examining the extent of water evaporation resulting from electron capture.[94] The conformational dynamics of multiply protonated peptides and proteins also contributes to uncertainties in identification of a particular charged site associated with the capture dynamics of an electron in high-*n* Rydberg states and the specification of the eventual site of electron localization in the cation radical. To circumvent these problems, relatively simple model systems have been investigated with high level quantum mechanical calculations.[89,95-97] The amide-I vibration (C=O stretching mode) dynamics was also examined as a simple model of the vibrational energy propagation in α-helix fragmentation upon ECD and ETD.[98]

To constrain the charged or radical site, recent studies have shown the effect of incorporation of permanent charged tags in peptides on backbone[69,70,99] and disulfide cleavage.[100] Improved sequence coverage of glycosylated and phosphorylated peptides has also been demonstrated using permanently charged tags.[101] Tags comprising strongly basic sites of proton localization as well as radical traps have been incorporated to study their effect on typical ECD fragmentations.[86,102] However, electron traps with a range of EAs have not been considered.

Turecek and coworkers used 2-(4′-carboxypyrid-2′-yl)-4-carboxamide (pepy) group[86] which has much higher gas-phase basicity (923 kJ/mol) compared to other basic groups in the peptide with the expectation that it is always protonated in the peptide dication. Thus it actually functions in the same manner as permanently charged tags such as quaternary ammonium or phosphonium groups by trapping an electron at the site of protonation because of its higher recombination energy. The resulting radical is also stable and does not contribute a labile hydrogen atom that

might be transferred to an amide carbonyl and lead to backbone cleavage. As a result, they observed the termination of N-C_α backbone cleavage in analogy with many other permanent tag experiments.

O'Connor and coworkers used the coumarin tag[102] which has a relatively low electron affinity (<0.6 eV), and hence, based on the experiments described in this work, cannot terminate peptide backbone cleavage solely by operation as an electron trap. Instead, the courmarin group acts as a free radical (hydrogen atom) scavenger to terminate the ECD process. In the experiments of O'Connor and coworkers, it is likely that initial electron capture and subsequent relaxation of the charge-reduced cation radical initially forms the aminoketyl intermediate, which in turn transfers the labile hydrogen atom to the coumarin substituent before cleavage of the peptide backbone can occur.

In the present work, we synthesized a series of EA-tuned peptides, which were generated from phosphopeptides, by attaching thiol groups having EAs ranging from −1.15 eV to 1.652 eV in their precursor forms. The model peptide, FQpSEEQQQTEDELQDK, was chosen because it has a C-terminal lysine residue, thus simulating a typical tryptic peptide, and also has a phosphoserine residue for inserting the EA-tuning tags between the N-terminal amine and the C-terminal lysine. For the synthesis of the EA-tuned peptides, a dehydroalanine residue is prepared by eliminating a phosphate group under basic conditions, followed by Michael addition of thiols to generate various benzylic cysteine residues. The derivatized peptide dications generated by electrospray are analyzed by ECD and ETD to investigate the effect of the EA-tuning tags. We observe that, with sufficiently high EA, the tag leads to inhibition of the backbone dissociation process normally observed in ECD and ETD experiments. We propose that this results from relaxation processes involving through-space or through-bond electron transfer from an initially formed high-*n* Rydberg state to the tag, followed by proton transfer to the resulting radical anion moiety.

The implications of these results for previously proposed mechanisms of electron capture and electron transfer dissociation are discussed. In addition, the present experiments allow for interpretation of matrix-assisted laser desorption/ionization (MALDI) in-source decay processes[103] resulting from MALDI plume chemistry involving electrons and multiply protonated ions and have important implications for the study of peptides possessing nitrated tyrosine as a PTM.[104,105]

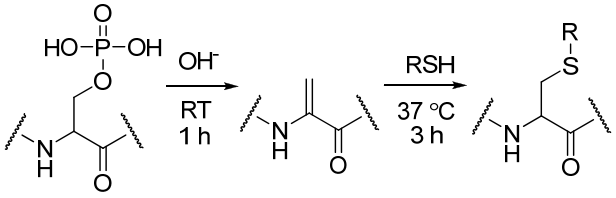
2.2. Experimental Section

2.2.1. Materials

Monophosphopeptide from β-casein (FQpSEEQQQTEDELQDK) was obtained from Anaspec (San Jose, CA). Thioacetic acid (HSAc), 0.3 N saturated barium hydroxide ($Ba(OH)_2$) solution, propanethiol (PT), benzyl bromide, 4-cyanobenzyl bromide, perfluorobenzyl bromide, 2-nitrobenzyl bromide, 4-nitrobenzyl bromide, 3,5-dinitrobenzyl chloride, 3-nitrobenzylthiol (3NBT), 1,3-dibromobenzaldehyde, sodium borohydride, mesyl chloride and α-cyano-4-hydroxycinnamic acid (CHCA) were acquired from Sigma-Aldrich (St. Louis, MO). Hydrochloric acid in methanol (~1.25 M) and 1-fluoro-3,5-dinitrobenzene were purchased from Fluka (Buchs, Switzerland). Methanol (MeOH), ethanol (EtOH), anhydrous N,N-dimethylformamide (DMF), anhydrous dichloromethane (DCM), dimethylether, acetonitrile (ACN), tetrahydrofuran (THF), ethyl acetate (EtOAc), anhydrous potassium carbonate (K_2CO_3) and OmniSolv™ high purity water were provided by EMD (Darmstadt, Germany). Dimethylsulfoxide (DMSO), formic acid (FA), and trifluoroacetic acid (TFA) were supplied by Mallinckrodt Inc. (Phillpsburg, NJ). All chemicals mentioned above were used as received without further purification. For desalting, OMIX™-100 μL size C-18 tips were purchased from Varian Inc. (Palo Alto, CA).

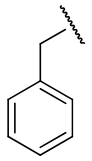
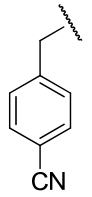
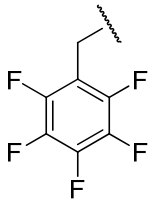
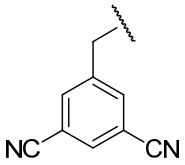
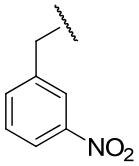
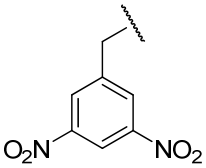
Synthesis of Benzylic Thiol:

Synthesis of the EA-tuned Peptide:



Scheme 2.1

Table 2.1 Electron Affinities of Thiol Precursors

Name of Benzyl Group	Benzyl-	4-Cyanobenzyl-	Perfluorobenzyl
Structure			
Electron Affinity (eV)	-1.15 ± 0.05[a]	0.258 ± 0.018[b] or 0.26 ± 0.1[c]	0.434 ± 0.081[e] or 0.730 ± 0.080[f]
Name of Benzyl Group	3,5-Dicyanobenzyl-	3-Nitrobenzyl-	3,5-Dinitrobenzyl-
Structure			
Electron Affinity (eV)	0.91 ± 0.1[g]	1.00 ± 0.010[h]	1.652 ± 0.048[j]

[a-j] Electron affinities are quoted from Ref. 111-118, respectively. For **Figure 2.7**, c and e were chosen for each compound due to the consistency of experimental methods.

2.2.2. Synthesis of the EA-Tuning Tags and Derivatized Peptides

The EA-tuning tags (benzyl thiols) were prepared from the corresponding benzyl halides. The literature procedure was followed with minor modification for better yield.[106,107] To synthesize thioesters, each benzyl halide (5 mmol) was dissolved in 15 mL of THF with 6 mmol of HSAc and 6 mmol of anhydrous K_2CO_3 in an air-free flask. The mixture was stirred at room temperature under a steady stream of N_2. The reaction time for each precursor varied from 1 to 1.5 h and the completion of reactions was monitored by thin-layer chromatography (TLC). The crude thioacetate obtained after standard aqueous work-up was sufficiently pure to use directly in the next step. The deacetylation reaction was carried out by adding 3mL of hydrochloric acid in methanol to a solution of the crude thioacetate in methanol and stirring at ~55-60 °C for 15~18 h. The thiol products were purified by flash chromatography on silica (1:20 EtOAc/hexane eluent) and identified by ^{1}H NMR (Supporting Information). Solid products such as 2-nitrobenzyl thiol, 4-nitrobenzyl thiol and 3,5-dinitrobenzyl thiol were dissolved in DMF at ~3-4 M concentration. All products were stored in sealed vials at 4 °C up to 6 months without any noticeable degradation.

Reactions involving formation of a dehydroalanine by β-elimination followed by Michael addition were used to attach the EA-tuning tags to our model phosphopeptide. A 20 μg portion of monophosphopeptide (FQpSEEQQQTEDELQDK) was dissolved in 40 μL of 4:3:1 mixture of H_2O/DMSO/EtOH (Solvent A) or 40 μL of 20% ACN (Solvent B), which proved optimal after extensive screening of solvent systems. In particular, these solvent systems provide enhanced solubility of thiols as described elsewhere.[108-110] Whereas solvent A generally worked well with all of the thiol compounds, solvent B proved better suited for perfluorobenzyl thiol. However, solvent B gave poor product recovery for nitrobenzyl thiols. An aliquot of 10 μL of 0.3 N

(saturated) $Ba(OH)_2$ solution was added and allowed to react at room temperature for 1 h. One μL of each thiol either in its liquid form or DMF solution was then added to the peptide solution, and the mixture was allowed to react at 37 °C for 3 h. The extended reaction time (~4-6 h) is required for less nucleophilic thiols such as 3,5-dinitrobenzyl thiols to improve the yield. Heating the mixture over 6 h at higher temperature results in poorer product recovery. The reaction was terminated by adding 1 μL of FA. The product mixture was vortexed and spun down by centrifugation. Supernatant was subjected to desalting using an OMIX^{TM}-100 μL size C-18 tip following the standard procedure. Identities of final products, eluted in 0.1% TFA, 50% ACN, 50% H_2O for MALDI or 0.1% FA, 50% MeOH, 50% H_2O for electrospray ionization (ESI), were confirmed by MS and directly used for ECD and ETD experiments. MALDI-MS spectra of the derivatized peptides were further investigated to seek the presence of prompt in-source decay backbone fragments (*i.e.*, *c* and *z* ions). The synthetic procedures above and the EAs of precursors[111-118] are summarized in **Scheme 2.1** and **Table 2.1**, respectively. The details for synthesis of 3,5-dicyanobenzyl thiol are available in Supporting Information. 1-Fluoro-3,5-dinitrobenzene (Sanger's reagent)[119] was conjugated to the N-terminal amine to be compared with 3,5-dinitrobenzylcysteine containing peptides synthesized by β-elimination and Michael addition reaction. The procedure described in the literature[120] with reaction conditions optimized for the selective N-terminal amine derivatization was used without any modification.

2.2.3. Mass Spectrometry

All ECD and IRMPD spectra were recorded using a 7-Tesla linear ion trap–Fourier transform (LTQ-FT) mass spectrometer (Thermo Scientific, San Jose, CA) with a nanoelectrospray ion source.[121] The flow rate was ~50 nL/min and spray voltage was varied from 1.0 to 1.5 kV by monitoring ion signals. Other critical parameters were capillary temperature 200 °C, capillary

voltage 30 V, and tube lens offset 200 V for maximal ion intensity. Other instrumental parameters were varied to optimize the intensities of the target ions in the linear ion trap prior to injection into the ICR cell. In ECD experiments electron irradiation occurred for 100 ms at ~5-7% of full energy scale, approximately corresponding to electron energy less than 1 eV and ~30 milliamp. Supplemental activation was accomplished by multiphoton excitation using a continuous 20 W CO_2 infrared laser for 100 ms at ~45-90% of full energy scale, approximately corresponding to 5 J/cm^2. The resolving power of FT MS scans was selected at 100,000 FWHM. For both ECD and IRMPD/ECD experiments, 100 scans were recorded.

ETD experiments were performed on a Thermo LTQ XL linear ion trap mass spectrometer (Thermo Scientific) modified for ETD. The eluted sample from the desalting step was directly infused into the microspray source at a flow rate of 2.0 μL/min. Spray conditions for maximizing ion counts included spray voltage 5.0 kV, capillary temperature 275 °C, capillary voltage 36 V and tube lens offset 70 V. The electron transfer reagent generated from the chemical ionization (CI) source was introduced to the linear ion trap from the rear of the instrument and allowed to react with isolated ions. Fluoranthene (EA ~0.7 eV)[122] was used for the CI reagent. The pressure of fluoranthene was 1×10^{-5} torr with a maximum injection time of 50 ms. Alternatively, isolated cations were collisionally activated for 200 ms prior to ETD in order to compare with IRMPD/ECD spectra.[123] ETD spectra were accumulated for ~1 min (ca. 50 scans) to accumulate a reasonable signal-to-noise ratio.

MALDI TOF spectra were acquired using a Voyager DE-PRO mass spectrometer (Applied Biosystems, Foster City, CA) equipped with a 20 Hz nitrogen laser (337 nm). All spectra were recorded in reflectron mode with 20 kV acceleration voltage, 150 ns delay extraction time and 75% grid voltage. 0.3 μL of the derivatized peptide solution was mixed with 0.3 μL of 10 mg/mL of CHCA matrix solution in 0.1% TFA, 50% ACN, 50% H_2O and spotted on a stainless steel

MALDI sample plate. Well-crystallized spots by the standard dried droplet method[124] were introduced into the mass spectrometer for analysis. Usually 100 laser shots were averaged. Recorded spectra were analyzed using Xcalibur (Thermo Electron, San Jose, CA) for ECD and ETD and Data Explorer (Applied Biosystems, Foster City, CA) for MALDI. Fragment ion masses were calculated using MS-Product of Protein Prospector.[125]

2.2.4. Quantum Mechanical Calculation

The PC GAMESS[126] (version 7.10) under Windows XP environment was used for the energetics of dicyanobenzene. To compare with previous work done by Polasek and Turecek, we used the same level of calculation and basis sets reported elsewhere.[127] The geometries were optimized using Becke's general gradient exchange functional[128] with Lee, Yang and Parr's correlation functional[129] (B3LYP) with the 6-31+G(d,p) basis set for dicyanobenzene (DCB), protonated dicyanobenzene ($DCBH^+$), dicyanobenzene anion radical ($DCB^{-\bullet}$) and hydrogen attached dicyanobenzene radical ($DCBH^\bullet$). For all open-shell systems, the spin-unrestricted method (UB3LYP) was used. Observed spin contamination in UB3LYP was small enough to be ignored ($<S^2>$ expectation values were 0.75-0.77). Optimized structures were further characterized by calculating vibrational frequencies and thermodynamic values using the same level of theory at 298.15K and 1.0 atm. To further refine the electronic energy of the system, electronic energies from UB3LYP/6-311+G(2df,p) and spin-restricted MP2 (ROMP2) with the same basis set were averaged (B3-ROMP2 energy).[127] Spin contamination in spin-unrestricted MP2 (UMP2) for open-shell systems was significant with an $<S^2>$ expectation value ~1.6. Therefore, the UMP2 method was not used for this work.

All other quantum mechanical calculations, including time-dependent density functional excited states analyses, were performed by GAMESS-US[130] (version April 11, 2008 R1) under

linux environment. The geometries of the model systems (**Figure 2.8**) were optimized at the B3LYP/6-31++G(d,p) level. All vertical electron affinities and recombination energies of the model systems were calculated without geometry relaxation. Further energy refinement was performed at the same level of theories described above for dicyanobenzene with the 6-311++G(2df,p) basis set. The M06 density functional[131] with the same basis sets was also used to estimate the energetics of the electron capture process. Calculations of the energetics of vertical electron capture with excited states were performed using time-dependent density functional theory (TDDFT) at the UB3LYP/6-31++G(d,p) and 6-311++G(2df,p) level as implemented in GAMESS for open-shell systems. Molecular orbitals (MOs) of excited states were prepared by linear combination of virtual orbitals with given coefficients from TDDFT calculations. Generated MOs were plotted using MacMolPlt[132].

All geometries of optimized structures from quantum mechanical calculations with electronic, zero-point energy, enthalpy corrections and excited state energies are available in Supporting Information.

2.3.Results

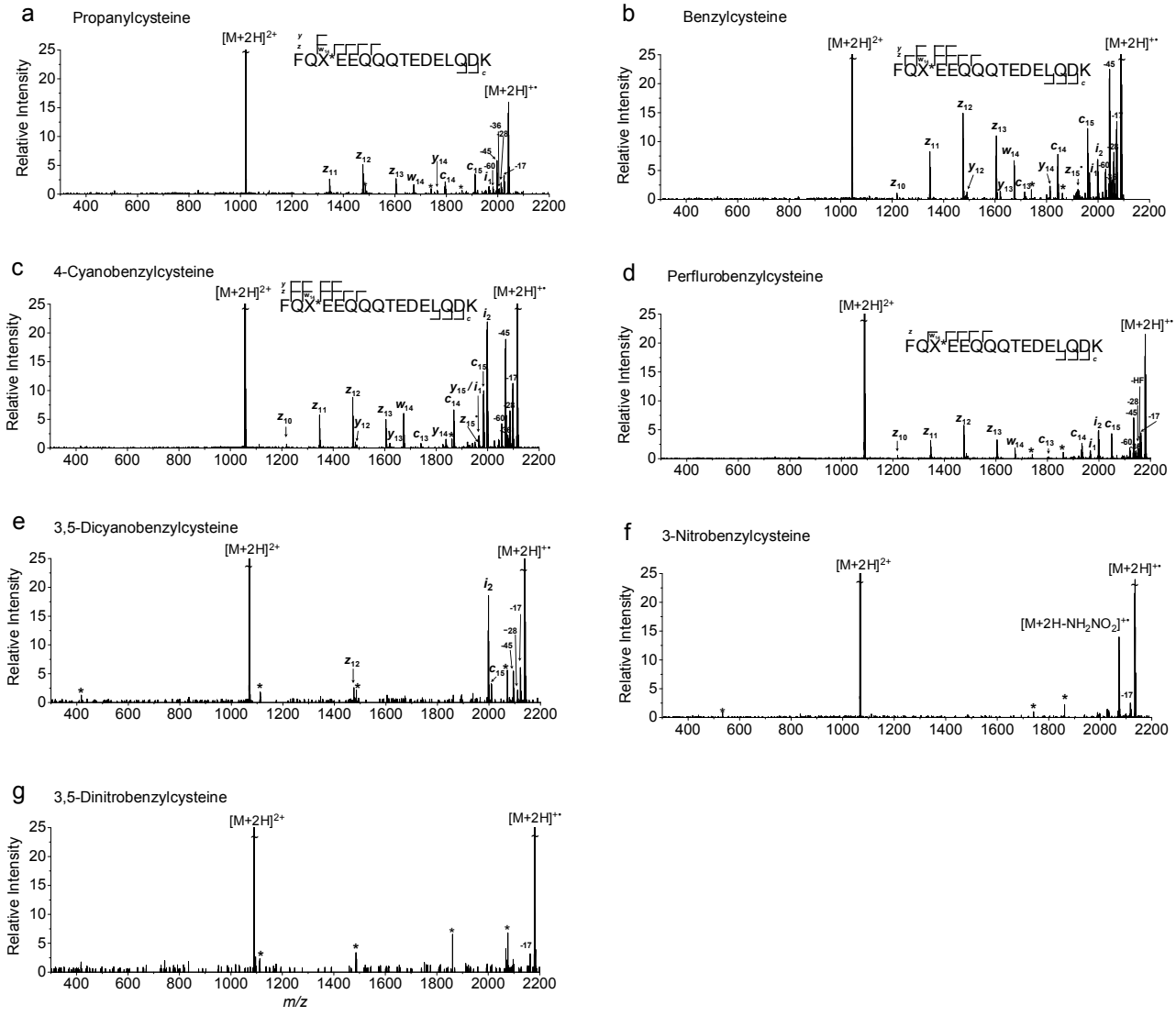
2.3.1. ECD of the EA-tuned Peptides

Each derivatized peptide was confirmed by electrospray ionization to form mainly doubly protonated ions. These ions are respectively denoted as $[P+2H]^{2+}$, $[B+2H]^{2+}$, $[4CB+2H]^{2+}$, $[PFB+2H]^{2+}$, $[35DCB+2H]^{2+}$, $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ for the model peptides FQX*EEQQQTEDELQDK, where X* is propanylcysteine, benzylcysteine, 4-cyanobenzylcysteine, perfluorobenzylcysteine, 3,5-dicyanobenzylcysteine, 3-nitrobenzylcysteine and 3,5-dinitrobenzylcysteine. To examine the effect of substitution position in the nitrobenzyl moiety, 2-nitrobenzyl and 4-nitrobenzylcysteine containing peptides were studied. The peptides

derivatized with 2NBT and 4NBT gave ECD and ETD spectra essentially identical to those of 3NBT (Supporting Information). Therefore, only spectra of $[3NB+2H]^{2+}$ are discussed in this paper. To investigate the effect of the location of the EA-tuning tags in the peptide, the 3,5-dinitrophenyl group was attached to the N-terminal amine of the model peptide using 1-fluoro-3,5-dinitrobenzene and the resulting peptide was subject to ECD experiments. 3,5-Dicyanobenzyl thiol (35DCBT) derivatized peptides were studied to compare different types of functional groups for tags having EA near 1.00 eV. The spectra acquired from the 2NBT, 4NBT and N_α-3,5-dinitrophenyl derivatized peptides are available in Supporting Information.

Figure 2.1 depicts ECD spectra of the derivatized peptides. Except **Figure 2.1**a, the spectra are presented in order of increasing EA of the benzyl substituents. The fragment ions induced by subsequent β-fission of a $z_n^\bullet$ ion and side-chain losses (−R• or −RS•; R is a substituent side-chain) of $[M+2H]^{+\bullet}$ are denoted as w_n, i_1 and i_2, respectively. The C-terminal ions (z_{10} to z_{15}) and the N-terminal ions (c_{13} to c_{15}) were detected in most of the spectra. Some of the C−N amide bond cleavages (*y* ions) were also observed in ECD of $[B+2H]^{2+}$ and $[4CB+2H]^{2+}$ (**Figures 2.1b-c**).

The most prominent peak among ECD type ions is z_{12} as discussed by Savitski *et al.*[133] Note that −1 or +1 Da shift from *c* or $z^\bullet$ ions by the abstraction of a C_α hydrogen were observed as reported by O'Connor *et al.*[30] and Savitski *et al.*[134] We label these as $c^\bullet$ and *z* ions which match with *c*−1 and $z^\bullet$+1 ions. In some cases, both $c^\bullet/c$ ions and $z^\bullet/z$ ions are identified simultaneously. Predominant *z* ions from $z^\bullet$ ions are believed to be formed by the abstraction of the C_α hydrogen in the derivatized cysteine residues which contain a methionine-like thioether bond and a strong electron withdrawing group at benzylic side-chains, resulting in a more reactive C_α−H bond. The z_{14} ion was not observed in any ECD spectrum. Related to this, the presence of w_{14} indicates the facile side-chain loss reaction pathway for the EA-tuning tags compared to side-chain losses from the remaining amino acids in the model peptides (**Figure 2.1**).[135]



ECD of doubly protonated model peptides. a) propanylcysteine, b) benzylcysteine, c) 4-cyanobenzylcysteine, d) perfluorobenzylcysteine, e) 3,5-dicyanobenzylcysteine, f) 3-nitrobenzylcysteine and g) 3,5-dinitrobenzylcysteine containing peptides, respectively. An asterisk indicates instrumental noise.

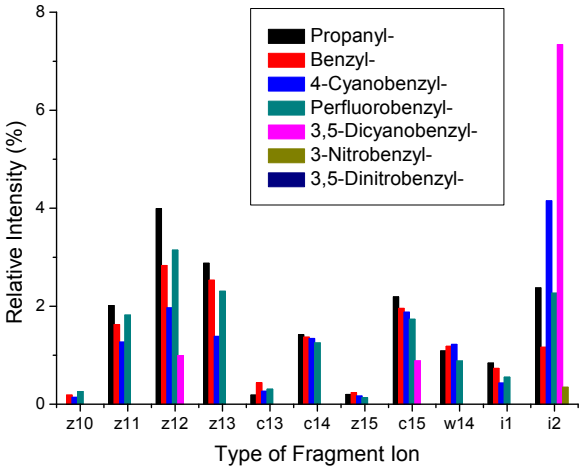
Figure 2.1 ECD of doubly protonated model peptides

The ECD spectrum of $[P+2H]^{2+}$ shown in **Figure 2.1a** exhibits a pattern of ECD backbone fragmentation typical of that observed in tryptic peptide dications.[133] Our model peptides have flexible gas-phase structures, allowing frequent interactions between protonated sites and backbone amide carbonyls. Considering the effects of Coulomb stabilization and hydrogen bonded carbonyls, both the Cornell mechanism and the UW mechanism are expected to be operational in this case and it remains unknown which one is more dominant for elucidating ECD spectrum of $[P+2H]^{2+}$.

As the EA of the tag is increased, the relative abundance of ECD-type ions diminishes (**Figure 2.1**). Relative yields of typical ECD backbone fragment ions, which all ECD peaks in each spectrum are normalized for comparison, are summarized in **Figure 2.2**. The peaks from ECD of $[PFB+2H]^{2+}$ deviate from the observed trend in that the peak abundance is decreasing as EA of the tag is increased. The unusually high abundance of i_2 in ECD of $[4CB+2H]^{2+}$ and $[35DCB+2H]^{2+}$ (**Figures 2.1c and e**) can be attributed to the stability of the RS• radical.

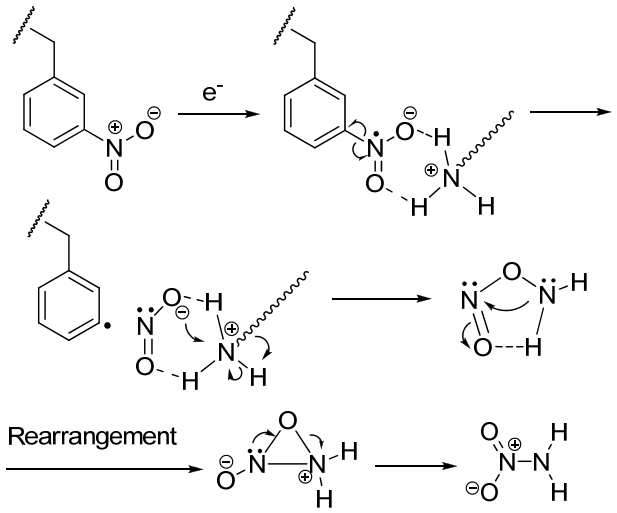
Remarkably, ECD spectra of $[35DCB+2H]^{2+}$, $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ exhibit essentially very small or no backbone fragmentation (**Figures 2.1e-g**). The loss of 17 Da from $[3NB+2H]^{+\bullet}$ and $[35DNB+2H]^{+\bullet}$ at m/z 2116.875 and 2161.862 turns out to be hydroxyl radical rather than ammonia by the comparison of measured and calibrated exact masses (**Figures 2.1f-g**). The mass deviation from the loss of hydroxyl radical in ECD of $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ is 0.73 and 0.06 ppm while that of ammonia is 10.51 and 11.07 ppm, respectively. Polasek and Turecek previously reported loss of hydroxyl radical from the phenylnitronic radical and characterized the energetics of this process.[127] More details about hydroxyl radical loss will be discussed in the following section. The loss of 17 Da from the remaining cation radicals is determined as ammonia.

Figure 2.2 Relative intensities of ECD fragment ions.



Relative intensities of ECD fragment ions. Intensities are taken from ECD spectra and reported as a total percent of the sum of the intensities of backbone fragments, side-chain losses and the charge reduced cation radical. The intensities attributed by −1 or +1 Da shift from $\boldsymbol{c}$ or $\boldsymbol{z}^{\bullet}$ ions by the abstraction of a C_{α} hydrogen are summed up to those of $\boldsymbol{c}$ or $\boldsymbol{z}^{\bullet}$ ions.

In addition, ECD of $[PFB+2H]^{2+}$ contains a product involving HF neutral loss (−20 Da) at m/z 2158.843, indicating possible formation of the perfluorobenzyl anion radical group, followed by proton transfer and loss of HF (**Figure 2.1d**). A similar process has been reported for gaseous perfluorobenzylalkylammonium ions forming the zwitterionic neutral radical by electron transfer and subsequent intramolecular proton transfer.[136]



Scheme 2.2

A neutral loss of 62 Da from $[3NB+2H]^{2+}$ was observed at m/z 2071.882 as a main fragment (**Figure 2.1f**). Considering the specific Coulomb interaction between positively charged groups (*i.e.*, the N-terminal amine and the ε-amine of lysine) and the nitrobenzylic cysteine anion radical formed by electron attachment, NH_2NO_2 is proposed as a reasonable candidate for this loss. However, it is not straightforward to propose a mechanism for NH_2NO_2 neutral loss. We tentatively suggest the process for NH_2NO_2 loss shown in **Scheme 2.2**. In the ECD spectrum of $[35DNB+2H]^{2+}$, a cation radical, $[35DNB+2H]^{+\bullet}$, is the most abundant product ion (**Figure 2.1g**). After NH_2NO_2 loss from $[35DNB+2H]^{+\bullet}$, the resulting product is less stable in comparison to that

of $[3NB+2H]^{+\bullet}$. Therefore, the product involving 62 Da loss in the ECD spectrum of $[35DNB+2H]^{2+}$ is not significant.

The ECD spectrum of doubly protonated N_{α}-3,5-dinitrophenyl derivatized peptide was also investigated to demonstrate the effect of the position of 3,5-dinitrophenyl group and its connectivity (from thioether to secondary amine) in the model peptide. No ECD-type backbone fragmentation is observed while most of the prominent side-chain losses remain as unknown peaks (Supporting Information). This observation is consistent with ECD of $[35DNB+2H]^{2+}$. It also clearly demonstrates that the presence of the 3,5-dinitrophenyl group in the model peptide is responsible for inhibition of ECD and ETD backbone cleavage processes rather than its location or chemical connectivity.

2.3.2. IRMPD/ECD of the EA-tuned Peptides

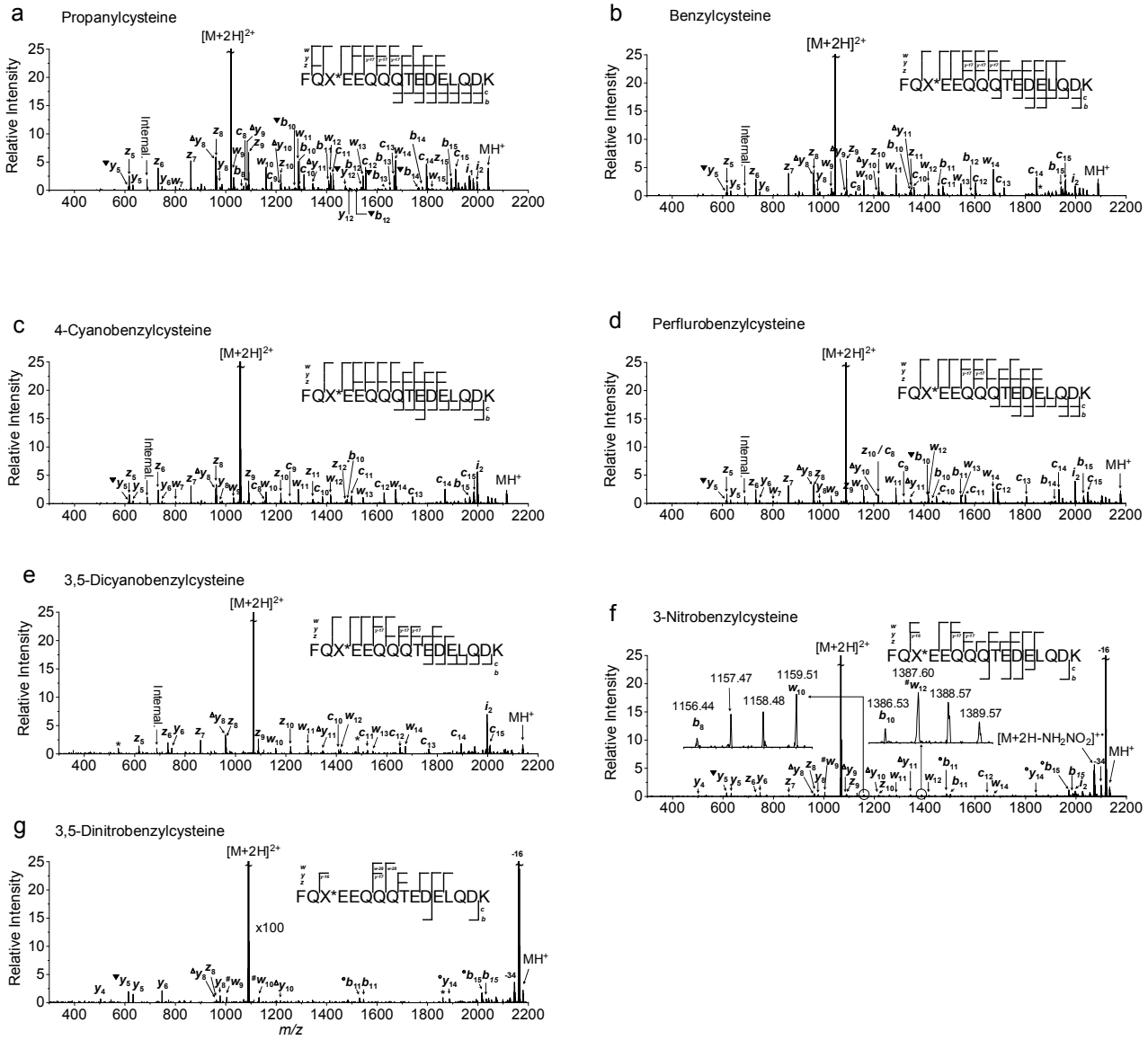
To further examine the stability of cation radicals considered in this study, IRMPD with ECD (IRMPD/ECD) was performed. Precursor ions were heated by infrared photons to just below the onset of backbone cleavage. Electrons were simultaneously injected into the ICR cell without isolation of heated precursor ions. It is reasonable to assume that the IRMPD/ECD spectra contain not only ECD fragments of heated precursor ions but some direct IRMPD fragments. Delayed electron injection (100 ms) into the ICR cell for reaction with ions preheated by infrared photons did not generate spectra significantly different from those obtained with simultaneous activation. Therefore, only simultaneous excitation by infrared photons and electrons (IRMPD/ECD) is discussed in this work.

The IRMPD/ECD spectra are shown in **Figure 2.3**. In comparison with ECD-only spectra, many of the C−N bond cleavages (*b*, *y* ions) from cation radicals were detected.[137] Hydrogen atom loss from the charge-reduced cation radical, $[M+2H]^{+\bullet}$, is predominant in every

IRMPD/ECD spectra, yielding $[M+H]^+$. The loss of 17 and 18 Da from *b* and *y* ions in IRMPD/ECD spectra are assigned as ammonia and water, respectively. It is worth noting that abundant ECD type fragments (*c*, *z* and *w* ions) are observed in IRMPD/ECD spectra of $[P+2H]^{2+}$, $[B+2H]^{2+}$, $[4CB+2H]^{2+}$ and $[PFB+2H]^{2+}$ while those of $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ exhibit a lower yield of these fragments (**Figure 2.3**). The IRMPD/ECD of $[35DCB+2H]^{2+}$ presents slightly reduced but still prominent peak intensities (**Figure 2.3e**). The existence of abundant *w* ions is attributed to the higher level of vibrational excitation provided by infrared photons. Unusual $w-C_2H_4$ ions are observed in **Figures 2.3f-g**, which are also believed to be induced by additional vibrational excitation.

The isotope distributions of *b* ions in the IRMPD/ECD spectra were investigated for the presence of $[b+1]^{+\bullet}$ ions formed by addition of a hydrogen atom to a typical *b* ion (**Figure 2.3**). The b_8 and b_{10} ions have abundant peaks 1 Da higher than their calculated monoisotopic masses. Mass deviations from the theoretical masses were, however, large enough not to assign those peaks as $[b+1]^{+\bullet}$ ions unlike a previous report.[86] The most dominant *b* ions (b_{11} and b_{15}) are observed at the C-terminus of aspartic acid residues and likely result from a salt-bridge mechanism (**Figure 2.3**).[10,11,138] However, no significant yield of $[b+1]^{+\bullet}$ ions from b_{11} and b_{15} ions was found, suggesting that the origin of b_{11} and b_{15} ions is the consequence of the direct IRMPD (data not shown). IRMPD/ECD spectra of the model peptides (**Figure 2.3**) were carefully examined for the presence of $[y+1]^{+\bullet}$ ions but none were detected.

Figure 2.3 IRMPD/ECD of doubly protonated model peptides.



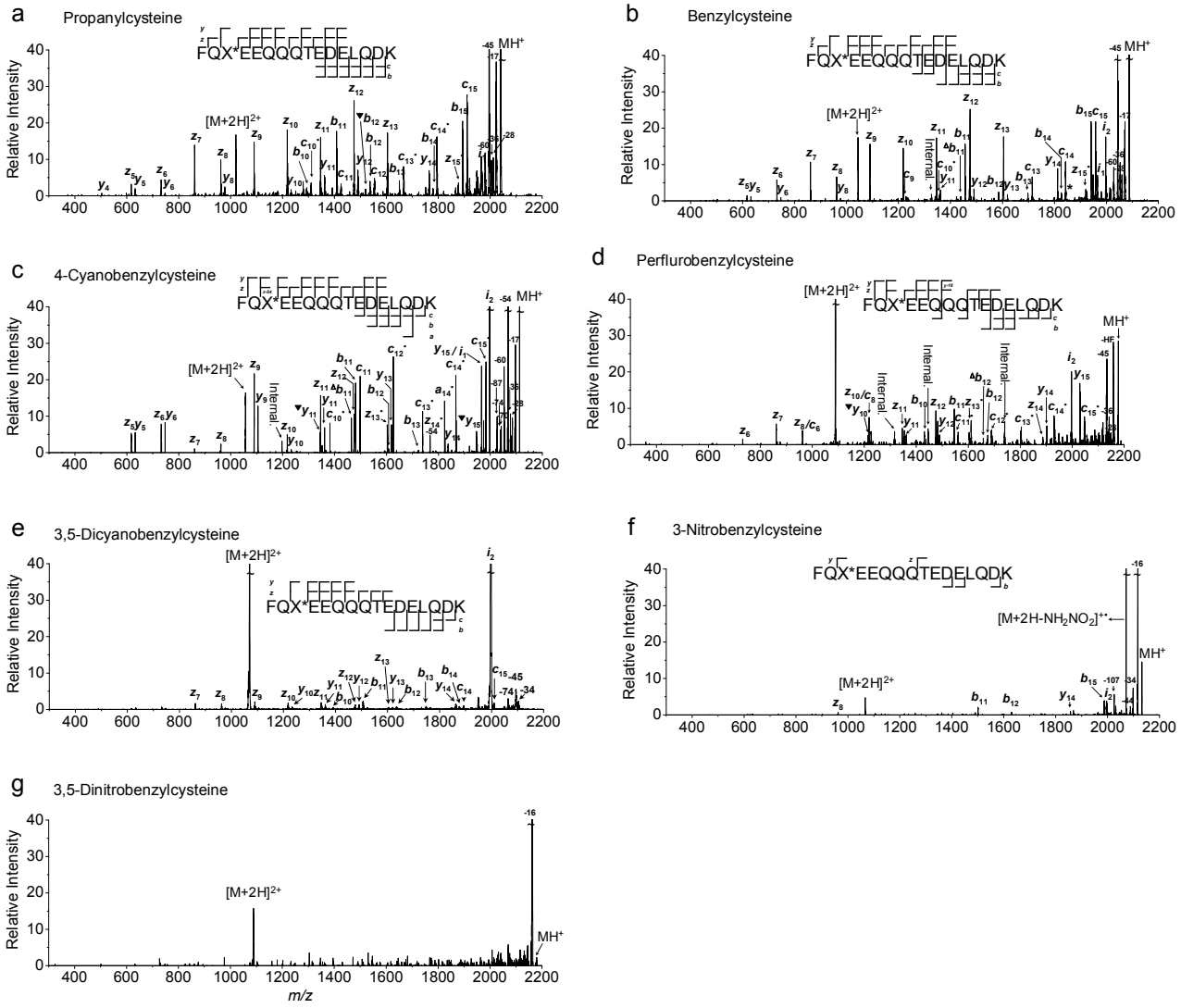
IRMPD/ECD of doubly protonated model peptides. a) propanylcysteine, b) benzylcysteine, c) 4-cyanobenzylcysteine, d) perfluorobenzylcysteine, e) 3,5-dicyanobenzylcysteine, f) 3-nitrobenzylcysteine and g) 3,5-dinitrobenzylcysteine containing peptides, respectively. Precursor ions were heated by infrared photons to just below the onset of backbone cleavage. Electron irradiation was applied simultaneously with infrared excitation without isolation of heated precursor ions. Symbolic superscript appendixes °, Δ, ▼ and # indicates loss of hydroxyl radical from $[\boldsymbol{b}+1]^{+\bullet}$ and $[\boldsymbol{y}+1]^{+\bullet}$ ions, and ammonia, water and ethylene from either ***b*** and ***y*** or ***w*** ions, respectively. An asterisk indicates instrumental noise.

2.3.3. ETD of the EA-tuned Peptides

In a separate set of experiments, ETD spectra of the derivatized peptides were obtained to investigate possible differences between ECD and ETD. Without supplemental activation by collision prior to the electron transfer reaction, significant yields of *c* or *z* fragment ions were not observed in any ETD spectra. Hence, only spectra from the ETD of collisionaly activated ions (ETcaD)[123] are discussed in this work.

ETcaD spectra of the derivatized peptides are shown in **Figure 2.4**. While peptide dications are the most abundant peaks in ECD and IRMPD/ECD spectra, hydrogen atom loss (**Figures 2.4a-e**) or hydroxyl radical loss (**Figures 2.4f-g**) from $[M+2H]^{+\bullet}$ is dominant in the ETcaD spectra. The relative intensities of precursor peptide dications and charge-reduced cation radicals observed in ECD and ETcaD spectra indicate that ETcaD has a higher dissociation product yield than ECD (**Figures 2.1** and **2.4**). ECD-like side-chain losses such as −17, −28, −36, −45 and −60 Da for ETcaD of $[P+2H]^{2+}$, $[B+2H]^{2+}$, $[4CB+2H]^{2+}$ and $[PFB+2H]^{2+}$ (**Figures 2.4a-d**) were identified. Loss of hydroxyl radical and NH_2NO_2 from $[3NB+2H]^{+\bullet}$ and hydroxyl radical from $[35DNB+2H]^{+\bullet}$ (**Figures 2.4f-g**) were observed. With ETD the coverage of sequence ions is generally better than that observed in ECD spectra. The ETcaD spectrum of $[P+2H]^{2+}$ (**Figure 2.4a**) includes 6 out of 15 possible *c* ions (c_8 to c_{15}) and 10 out of 15 possible *z* ions (z_5 to z_{13} and z_{15}) while that of the ECD spectrum spans 2 out of 15 possible *c* ions (c_{14} and c_{15}) and 4 out of 15 possible *z* ions (z_{10} to z_{13}) (**Figure 2.1a**). The pattern of hydrogen abstraction forming $c^{\bullet}/c$ ions and $z^{\bullet}/z$ ions becomes more complex in comparison to the ECD data (**Figures 2.1** and **2.4**). No evidence was found for the presence of $[b+1]^{+\bullet}$ and $[y+1]^{+\bullet}$ ions in an examination of the isotope distributions of *b* and *y* ions. Therefore, all *b* and *y* ions are believed to be induced by direct action of vibrational excitation prior to ion/ion reaction.

Figure 2.4 ETD of doubly protonated model peptides.



ETD of doubly protonated model peptides. a) propanylcysteine, b) benzylcysteine, c) 4-cyanobenzylcysteine, d) perfluorobenzylcysteine, e) 3,5-dicyanobenzylcysteine, f) 3-nitrobenzylcysteine and g) 3,5-dinitrobenzylcysteine containing peptides, respectively. Supplemental activation was performed prior to reaction with fluoranthene anion. Symbolic superscript appendixes °, Δ and ▼ indicates loss of hydroxyl radical from $[\boldsymbol{b}+1]^{+\bullet}$ and $[\boldsymbol{y}+1]^{+\bullet}$ ions, and ammonia, water from ***b*** and ***y*** ions, respectively.

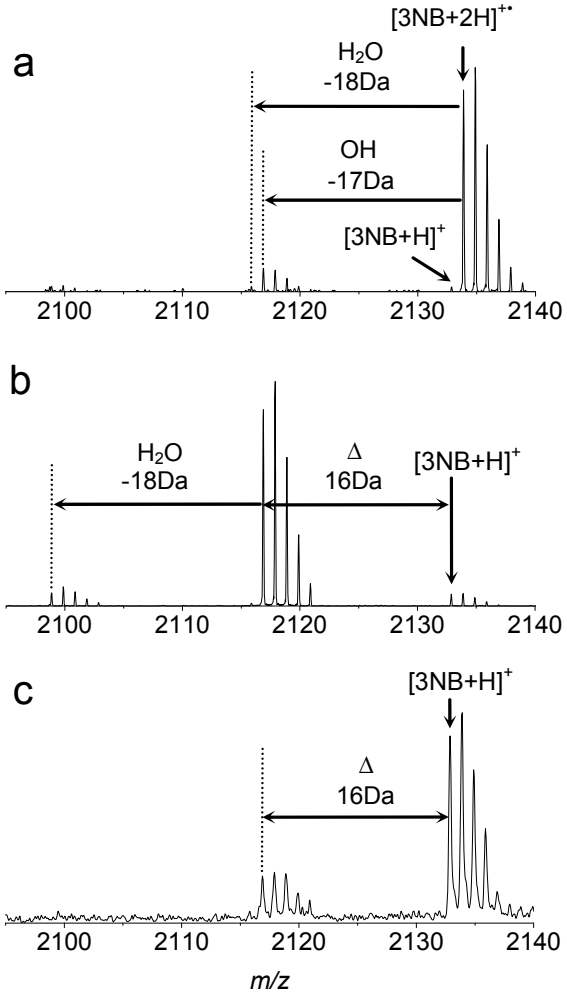
Despite differences between ECD and ETD (*i.e.*, electron capture/transfer cross section, exothermicity from electron transfer reaction depending on the electron affinity of the electron carrier reagent, and time scale of reaction or ion detection), typical backbone fragmentation is almost completely inhibited in ETcaD spectra of $[35DCB+2H]^{2+}$, $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ (**Figures 4e-g**). This observation reinforces the validity of the electron predator model for both ETD and ECD.

The presence of cleaved but hydrogen-bonded *c*, *z* fragment complexes were hypothesized in a previous study.[30] This possibility can be explored using a high level of vibrational excitation in the peptide cation radicals. As seen in **Figures 2.3f-g** and **2.4f-g**, this fails to yield significant abundances of ECD or ETD type backbone fragments. This supports the conjecture that stable peptide cation radicals are formed rather than hydrogen bonded *c* and *z* fragment complexes. However, the IRMPD/ECD of $[35DCB+2H]^{2+}$ exhibits slightly more abundant fragment yields compared to the corresponding ECD and ETcaD spectra (**Figures 2.1e, 2.3e** and **2.4e**). This also indicates that the nascent $[35DCB+2H]^{+\bullet}$ cation radical is less stable compared to $[3NB+2H]^{+\bullet}$ and $[35DNB+2H]^{+\bullet}$ under the higher level of vibrational excitation.

2.3.4. Hydroxyl Radical Loss and Ion Formation Mechanism in MALDI Plumes

As seen in **Figures 2.1f-g**, hydroxyl radical loss occurs from $[3NB+2H]^{+\bullet}$ and $[35DNB+2H]^{+\bullet}$. In IRMPD/ECD, several peaks are observed 16 Da less than some *b* and *y* ions, indicating loss of hydroxyl radical from intermediately formed $[b+1]^{+\bullet}$ and $[y+1]^{+\bullet}$ ions (**Figures 2.3f-g**). Relevant to hydroxyl radical and related losses, formation of the phenylnitronic radical and its dissociation energetics were investigated in detail by Polasek and Turecek.[127] The phenylnitronic radical is quite stable on the microsecond life time[127] and does not appear to initiate significant backbone cleavages or other side chain losses in ECD of $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$. However, the

Figure 2.5 Hydroxyl loss from the charge-reduced cation radical



Hydroxyl loss from the charge-reduced cation radical of 3-nitrobenzylcysteine containing peptide. a) ECD, b) IRMPD/ECD and c) MALDI-TOF MS of 3-nitrobenzylcysteine containing peptide. The peak 16 Da less than [**3NB**+H]$^+$ is loss of hydroxyl radical from the charge reduced cation radical, [**3NB**+2H]$^{+\bullet}$ based on our discussion. a) and b) are magnified **Figures 2.1f** and **2.3f**, respectively, in m/z region between 2050 and 2140 around [**3NB**+2H]$^{+\bullet}$ ion. c) was recorded using a time of flight mass spectrometer equipped with a 337nm N_2 laser in the reflector mode. 100 shots were averaged. 10 mg/ml CHCA was used for matrix.

phenylnitronic radical group easily undergoes a direct homolytic cleavage leading to hydroxyl radical loss and this process, which has an extremely low reverse reaction barrier (*ca.* ~0 kJ/mol),[127] is especially prominent with higher levels of vibrational excitation (**Figures 2.3f-g** and **2.4f-g**). The loss of HONO is calculated to be less energetically favorable,[127] consistent with our observation that this is a less prominent dissociation pathway (**Figures 2.1f-g**, **2.3f-g** and **2.4f-g**). These theoretical calculations and experimental observations clearly support the formation of nitrobenzyl anion radical group and intramolecular proton transfer to it in ECD, IRMPD/ECD and ETcaD spectra of the nitrobenzylcysteine containing peptides.

Hydroxyl radical loss also provides an explanation for the product appearing 16 Da less than $[3NB+H]^+$ in the MALDI MS (**Figure 2.5c**). A similar loss from the 3-nitrotyrosine residue in MALDI MS of peptides has been reported previously.[139-141] In the MALDI plume, a number of free electrons exist and may react with desorbed primary ions and neutrals.[142] Protons can also be provided by numerous matrix molecules. From these observations, we suggest that ion yields in MALDI may in part result from charge neutralization process by electron capture of multiply protonated ions. This has also been discussed in several papers.[143-145] However, prompt in-source decay backbone fragments (*i.e.*, *c* and *z* ions) from the derivatized peptides were not observed in this work (data not shown).

2.3.5. Kinetics of Electron Capture

At the inception of this study, we speculated that the tags having positive electron affinities might increase the overall efficiency of electron capture. This would be the case if, following the initial electron capture event, electron autodetachment competes with further relaxation of the nascent radical cation to yield ECD products. To investigate this possibility, ECD spectra of simultaneously isolated $[B+2H]^{2+}$ and $[3NB+2H]^{2+}$ ions were recorded. Similar initial ion signal

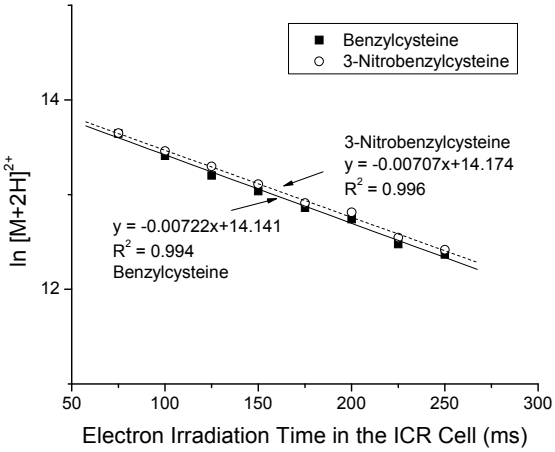
intensities of peptide dications in the FT MS spectrum ($[B+2H]^{2+}/[3NB+2H]^{2+}$ = ~0.95) were established, and electron irradiation time was sequentially increased from 75 to 250 ms in order to monitor the relative electron capture kinetics. Assuming a constant electron flux during the irradiation period, the rate of electron capture can be expressed as in **Equation 2.1**,

$$-\frac{d[(\mathrm{M}+2\mathrm{H})^{2+}]}{dt} = k_{\mathrm{obs}}[(\mathrm{M}+2\mathrm{H})^{2+}][e^{-}]_{\mathrm{s}}$$

Equation 2.1

where $[(M+2H)^{2+}]$ and $[e^{-}]_s$ are the number of the precursor ions and electrons, and k_{obs} is the observed rate constant of the electron capture process. **Equation 2.1** yields first order kinetics for the doubly charged ions, demonstrated by the data in **Figure 2.6**, where the logarithm of the $[B+2H]^{2+}$ and $[3NB+2H]^{2+}$ ion intensities versus electron irradiation time in the ICR cell are plotted. The nearly identical slopes indicates similar electron capture rates for $[B+2H]^{2+}$ and $[3NB+2H]^{2+}$. No change is observed that can be attributed to the higher EA tag. This is consistent with earlier studies which conclude that electron capture rates into high-*n* diffuse Rydberg states possess probabilities that vary as the square of the total charge of the ion.[25,26] The eventual site at which the electron becomes localized is determined by through-space and through-bond electron transfer processes subsequent to the initial capture.[79]

Figure 2.6 Electron capture kinetics



Variation in the natural logarithm of $[\mathbf{B}+2H]^{2+}$ and $[\mathbf{3NB}+2H]^{2+}$ with electron irradiation time in the ICR cell. Both precursor ions were simultaneously isolated for ECD with similar ion intensities. Slopes indicate that the electron predator has no effect on the rate of electron capture.

2.4.Discussion

2.4.1. Effect of EA-tuning Tags on Nascent Cation Radicals

The percent yield of each ECD fragmentation channel is depicted as a function of EA of tags in **Figure 2.7**. **Equations 2.2-5** are used to calculate relative yield of different ECD processes, where a = charge-reduced radical cations ($[M+2H]^{+\bullet}$), b = Σ [c_i + z_i + w_i ions], c = Σ [side-chain loss] and d = Σ [other backbone fragments (*b* and *y* ions) and subsequent loss of H_2O or NH_3]. For each term, background noise was subtracted and isotopic contributions of each ion were summed up.

Total ECD yield = (a + b + c) / (a + b + c + d) × 100 **Equation 2.2**

Total EC, no D yield = a / (a + b + c + d) × 100 **Equation 2.3**

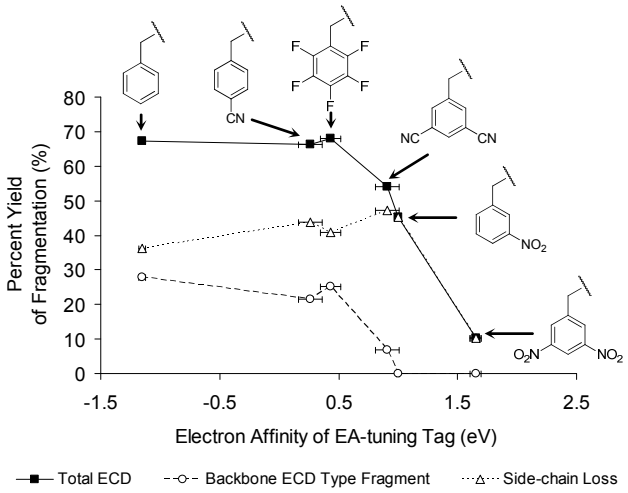
Backbone ECD-type fragment yield = b / (a + b + c + d) × 100 **Equation 2.4**

Side-chain loss yield = c / (a + b + c + d) × 100 **Equation 2.5**

As seen in **Figure 2.7**, yield of *c*- and *z*- type backbone fragmentation generally diminishes with increasing EA of tags in the model peptides. Typical ECD-type backbone fragments start to disappear when EA of the tag exceeds ~1.0 eV, independent of the functionality of the tag. It should be also noted that abundant side-chain losses in ECD of $[35DCB+2H]^{2+}$, $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$ are mostly contributed by tag-related peaks such as RS• (i_2), •OH and NH_2NO_2 losses, and not by other amino acids in the peptides.

Different electron relaxation processes have different exothermicities, but they also lead to final states with dissociation pathways having very disparate activation energies. Therefore, it is important to consider the factors related to the stability of nascent peptide cation radicals formed

Figure 2.7 Relationship between the electron affinities and yields of ECD



Relationship between the electron affinities of the tags and percent yields of various ECD fragmentation channels, including total ECD (solid line), backbone ECD-type fragment (dash line) and side-chain loss yield (dotted line), respectively. The horizontal error bars are taken from references for the electron affinities of tags in Ref. 111-118. Each isotope distribution of ions is summed and normalized followed by **Equation 2.2-5**.

in the electron capture and relaxation process. **Figure 2.7** clearly demonstrates that EA of the tag is the most important parameter relating to stability of the cation radicals. A secondary factor appears to be the PAs of different intermediate anion radicals. Namely, if two tags have similar positive EA with different PAs of the corresponding anion radicals, ECD-type backbone fragmentation of the tag with lower PA is more prominent. This idea is supported by calculated energetics of dicyanobenzene and nitrobenzene (**Table 2.2**) and by observed ECD spectra (**Figures 2.1e-f**). It is obvious that the most stable cation radical is $[35DNB+2H]^{+\bullet}$ which exists mostly as a nascent cation radical with minimal fragmentation. To summarize, exceptional stability of nascent cation radicals is conferred by the generation of a stable radical center by electron capture followed by intramolecular proton transfer.

The present investigation also leads to the conclusion that ECD and ETD may not generate abundant backbone cleavages in characterization of tyrosine nitration, which is widely observed in proteins as a post-translational modification.[104,105]

2.4.2. Quantum Mechanical Calculations

To further investigate the energetics and mechanism of electron capture in the presence of our tags, we performed several quantum mechanical calculations using a series of model compounds. First, the energetics of adding an electron, proton and hydrogen atom to the electron predators were evaluated to illuminate the stability and reactivity of model nascent cation radicals. Dicyanobenzene and nitrobenzene were chosen as model compounds to represent electron predators. The energetics of each process for nitrobenzene are derived from a previous study[127] and are used here. Second, time-dependent density functional calculations of a series of reduced model peptide systems (**Figure 2.8**) were performed to estimate the relative energies among the

Table 2.2 Enthalpies from Quantum Mechanical Calculations on 1,3-dicyanobenzene.

1,3-Dicyanobenzene

	Level of Quantum Mechanical Calculation	Proton Affinity[a]	Electron Affinity[b]	Proton Affinity of Anion[a]	Hydrogen Affinity[a]
Theoretical	B3LYP/6-31+G(d,p)	787.5	1.073	1312.4	108.7
	B3LYP/6-311+G(2df,p)	787.7	1.096	1307.0	100.5
	ROMP2/6-311+G(2df,p)	756.1	0.780	1269.9	39.1
	B3-ROMP2[c]	771.9	0.938	1288.5	68.9
Experimental		779.3	0.91	N/A	N/A

Nitrobenzene

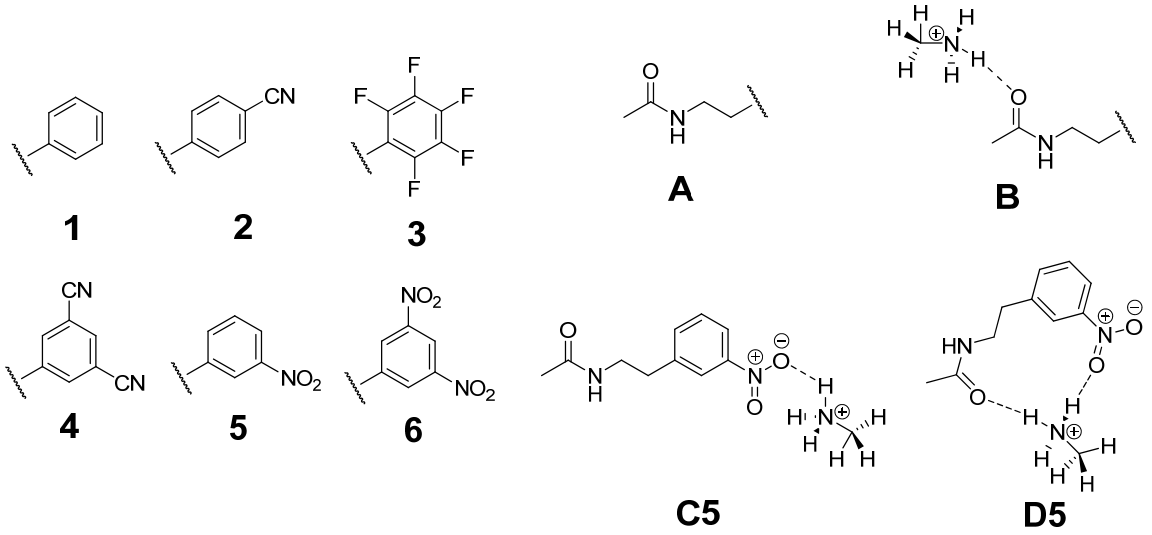
	Level of Quantum Mechanical Calculation	Proton Affinity[a]	Electron Affinity[b]	Proton Affinity of Anion[a]	Hydrogen Affinity[a]
Theoretical	B3LYP/6-31+G(d,p)	806.5	1.288	1384.8	201.8
	B3LYP/6-311+G(2df,p)	806.6	1.252	1386.9	195.5
	ROMP2/6-311+G(2df,p)	776.5	0.718	1385.5	148.7
	B3-ROMP2[c]	791.6	0.985	1386.2	172.1
Experimental		800.3	1.00	N/A	N/A

[a] unit of kJ/mol, [b] unit of eV, [c] 1/2 (B3LYP + ROMP2).

Enthalpies from quantum mechanical calculations on 1,3-dicyanobenzene. Enthalpies of protonated, electron and hydrogen attached species of 1,3-dicyanobenzene were calculated and compared with those of nitrobenzene from Polasek and Turecek.

excited states of cation radicals. These model systems comprise a series of N-(substituted-phenyl)acetamides with (B1-B6) or without (A1-A6) methyl ammonium, which forms a strong hydrogen bond to the amide carbonyl. For N-(3-nitrophenyl)acetamide and N-(3,5-dicyanophenyl)acetamide, the structures having strong hydrogen bonds to the substituted moieties such as the nitro or cyano groups are considered (C4 and C5). In particular, for *N*-(3-nitrophenyl)acetamide, the very stable structure formed with strong hydrogen bonds to both amide carbonyl and nitro oxygen (D5) is investigated. The vertical electron affinities and recombination energies were also calculated to provide vertical electronic energies of the lowest electronic states of each model species. This facilitates evaluation of the relative exothermicities of different electron relaxation processes to specific orbitals related to different reaction pathways (*i.e.*, forming a stable radical intermediate or forming precursors that can lead to typical ECD backbone fragmentation processes). Before discussing the electron capture process, it is appropriate to consider the sites of protonation in our model peptide cations. Unlike the 2-(4′-carboxypyrid-2′-yl)-4-carboxamide group studied by the Turecek group[86] as a radical trap, our electron predators, a term used to describe the superior electron trapping abilities of 3,5-dicyanobenzyl, 3-nitrobenzyl and 3,5-dinitrobenzyl groups, are not stronger gas-phase bases (PA[1,3-Dicyanobenzene] = 779.3 kJ/mol, PA[Nitrobenzene] = 800.3 kJ/mol)[146] than other possible protonation sites such as the N-terminal amine (PA[Glycine] = 866.5 kJ/mol)[146] or the ε-amine of lysine (PA[Lysine] = 966.0 kJ/mol).[146] Therefore, peptide dications are not likely to be protonated at the site of the EA-tuning tags. The probable sites of protonation in the model peptide chosen for this study are the N-terminal amine and lysine amine.

Figure 2.8 Structures of the model compounds for quantum mechanical calculations.



Structures of the model compounds for quantum mechanical calculations. These are prepared by a combination of each aromatic functional group (**1**-**6**) with either acetamide (**A**) or methyl ammonium acetamide complex by a hydrogen bond to the amide carbonyl (**B**). Some methyl ammonium complexes having 3,5-dicyanophenyl (**4**) and 3-nitrophenyl ring (**5**) form hydrogen bonds with the cyano and the nitro group (**C4** and **C5**) and both the amide carbonyl and nitro group, simultaneously (**D5**).

Table 2.2 summarizes all calculated energies related to dicyanobenzene and nitrobenzene. The protonation sites of 1,3-dicyanobenzene (DCB) and nitrobenzene are the nitrogen of one of the cyano groups and the oxygen of the nitro group, respectively.[127,147] The full sets of optimized structures and electronic energies, zero-point energy corrections and enthalpies of DCB, $DCBH^+$, $DCB^{-\bullet}$ and $DCBH^\bullet$ are available in Supplemental Information. The enthalpy of each species is compared with that of Polasek and Turecek's report for nitrobenzene.[127] The adiabatic electron affinity of 1,3-dicyanobenzene calculated at the B3-ROMP2/6-311+G(2df,p)//B3LYP/6-31+G(d,p) level in this work is 0.937 eV, in good agreement with the experimental value of 0.91 eV.

An important observation from these calculations is the difference of hydrogen affinity of 1,3-dicyanobenzene (69.8 kJ/mol) and nitrobenzene (172.1 kJ/mol), which contrasts with their similar EAs (EA[1,3-dicyanobenzene] = 0.91 eV, EA[nitrobenzene] = 1.00 eV). The ~2.5 times higher hydrogen affinity of nitrobenzene compared to that of 1,3-dicyanobenzene may in part be responsible for the absence of any significant ECD type backbone fragment from the 3-nitrobenzyl derivatized peptide (**Figure 2.1f**) while the 3,5-dicyanobenzyl derivatized peptide exhibits small yields of *c* and *z* ions (**Figure 2.1e**). It is also noteworthy that both tags have higher hydrogen affinity than the amide carbonyls (21-41 kJ/mol).[81]

To estimate the overall energy released by the electron capture process, we calculated the vertical electron affinity of the neutrals and the vertical recombination energy of the cation-neutral complexes by adding an electron to each system without geometry optimization (**Table 2.3**). The general trend observed in **Table 2.3** is reasonable in comparison with the electron affinities of the tags listed in **Table 2.2**, regardless of the presence of Coulomb stabilization conferred by the methyl ammonium ion. Notably, electron affinities of A3 and B3 were estimated as slightly negative values regardless of the calculation methods, in contrast to the experimentally

Table 2.3 The Vertical Electron Affinities and Vertical Recombination Energies of the Model Compounds

Species	UB3LYP	UM06	UB3LYP	UM06	ROMP2	B3-ROMP2[a]
	6-31++G(d,p)		6-311++G(2df,p)			
A1	-0.424	-0.605	-0.394	-0.519	-0.589	-0.491
A2	0.074	0.171	0.136	0.275	-0.390	-0.127
A3	-0.025	-0.194	-0.028	-0.141	-0.418	-0.223
A4	0.801	0.947	0.875	1.040	-0.262	0.306
A5	0.819	0.962	0.852	1.019	-0.393	0.230
A6	1.584	1.725	1.602	1.753	N/A[b]	N/A[b]
B1	2.950	2.698	2.946	2.739	2.740	2.843
B2	3.184	2.969	3.196	3.033	2.872	3.034
B3	3.120	2.890	3.112	2.930	2.858	2.985
B4	3.494	3.348	3.523	3.425	2.947	3.235
B5	3.521	3.403	3.530	3.455	2.834	3.182
C4	3.569	3.567	3.612	3.657	2.919	3.266
C5	4.668	4.798	4.700	4.853	4.305	4.502
D5	4.060	4.209	4.095	4.264	3.599	3.847
B6	3.966	4.000	3.974	4.031	2.991	3.482

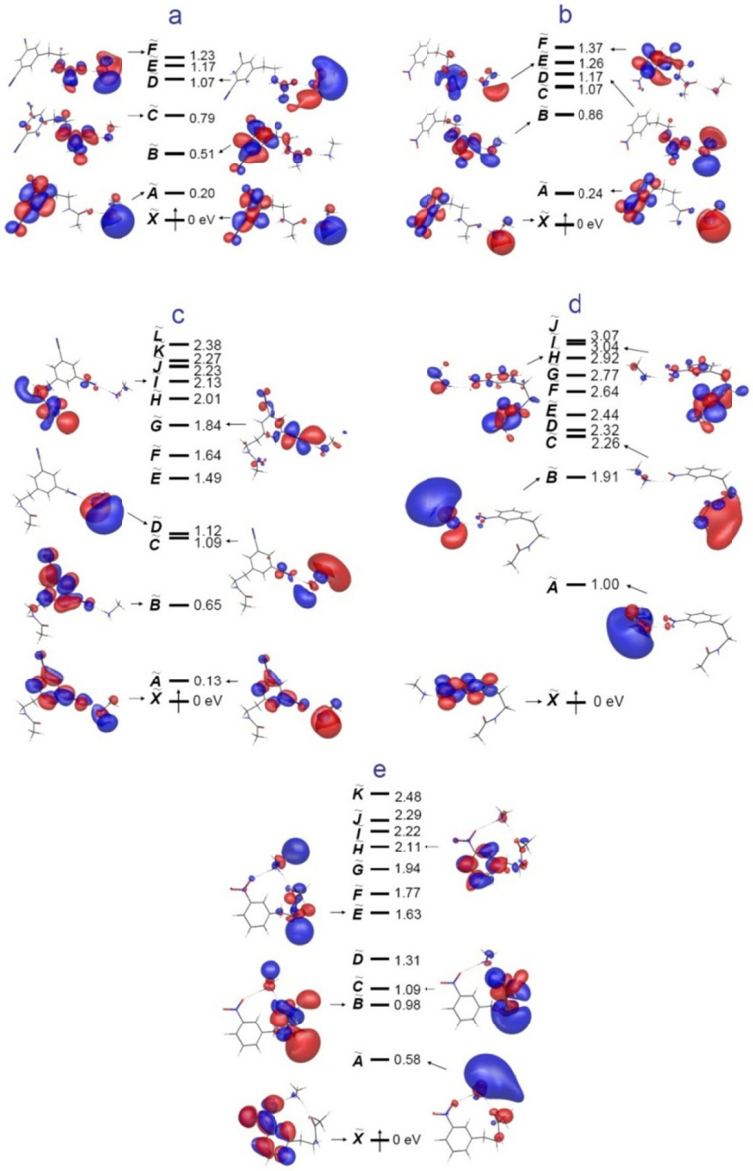
[a] 1/2 (UB3LYP + ROMP2) without zero point energy correction. [b] Unrestricted open-shell SCF was not converge.

The vertical electron affinities and vertical recombination energies of the model compounds described in **Figure 2.8** at various levels of theories. All energies are in units of electron volt.

reported values in **Table 2.1**. However, Frazier *et al.* reported negative electron affinities of the π^* orbitals of perfluorobenzene,[148] which lends support to the validity of the calculated negative vertical electron affinities. ROMP2 vertical electron affinities for A2 through A5 seem to be erroneous showing all negative values. This manifest error may be caused by the limitation of the restricted spin calculation. It should be stressed that recombination energies of methyl ammonium complexes are highly dependent on their particular hydrogen bond acceptors. Also, although B6 has two nitro groups on the phenyl ring, C5 undergoes the most exothermic recombination process.

To further investigate the relative energetics of excited states during the relaxation of a captured electron, we performed time-dependent density functional calculations on the model systems shown in **Figure 2.8**. Excited state orbitals of charge-neutralized B4, B5, C4, C5 and D5 radicals generated by TDDFT calculations are depicted in **Figure 2.9**.[149] These excited MOs clearly reveal the effects of different hydrogen bonding partners. As seen in **Figures 2.9a** and **2.9b**, a hydrogen bond to the amide carbonyl lowers the energy of the amide π^* orbital, while the nitrophenyl π^* orbital mixed with the ground Rydberg orbital of the methyl ammonium ion give rise to nearly degenerate lowest states (X and A states). The relative energy gaps among orbitals in which we are interested are quite similar in both B4 and B5 (**Figures 2.9a-b**). If the methyl ammonium ion directly interacts with an oxygen of the nitro group as in C5, it significantly stabilizes the nitrophenyl π^* orbital, pushing the ground Rydberg orbital (A state) and the amide π^* orbital (H and I states) to higher levels (**Figure 2.9d**). This effect is diminished by having another hydrogen bond with the amide carbonyl simultaneously with the nitro group (**Figure 2.9e**). However, this reordering of orbitals is not observed in the case of C4 despite the presence of the similar hydrogen bond with the cyano group (**Figure 2.9c**). As seen in excited state MOs of

Figure 2.9 Excited state molecular orbitals obtained from time-dependent density functional calculations



Excited state molecular orbitals obtained from time-dependent density functional calculations of a) **B4** b) **B5** c) **C4** d) **C5** and e) **D5** at the UB3LYP/6-311++G(2df,p) level. See **Figure 2.8** for the structure of each species.

B4 and B5, the first two excited states of C4 are constituted from the dicyanophenyl π^* orbitals mixed with the ground Rydberg orbitals of the methyl ammonium ion, being nearly degenerate.

In summary, these theoretical calculations and experimental observations lead to two conclusions. First, the inhibition of typical ECD backbone fragmentation requires a certain level of intrinsic positive electron affinity of the tag. The efficiency of the electron trap is further augmented by structure-dependent hydrogen bonds to the derivatized functional groups. In particular, the higher proton affinity of the nitro group compared to the cyano group (**Table 2.2**) facilitates more stable hydrogen bond formation with the N-terminal amine or lysine ε-amine. This results in higher populations of structural conformations which stabilize the nitrophenyl π^* orbital and push other orbitals to higher levels. It is thus a reasonable prediction that the nascent $[35DCB+2H]^{+\bullet}$ cation radical would be less stable than $[3NB+2H]^{+\bullet}$ and $[35DNB+2H]^{+\bullet}$. This prediction is consistent with our observations of small fractions of typical ECD backbone fragmentation in ECD, IRMPD/ECD and ETcaD of $[35DCB+2H]^{2+}$ (**Figures 2.1e, 2.3e** and **2.4e**). Therefore, we conclude that the electron relaxation process after the initial electron capture to high-lying Rydberg states is modulated by the presence of tags with positive EAs and their structure-dependent hydrogen bonds.

Second, the formation of a stable and regiospecific radical center[15,137] on the nitrophenyl tags raises a question regarding the operation of the UW mechanism for ECD-type backbone fragmentation in the EA-tuned peptides. This mechanism invokes the engagement of Coulomb stabilized amide π^* orbitals in the electron relaxation and subsequent backbone cleavage processes. Although this process is energetically exothermic and has a lower barrier than the Cornell mechanism,[61,69,75,84] backbone fragmentation was not observed in the presence of electron predators. In addition, the proton affinity of the amide carbonyl group ($PA[CH_3CONHCH_3]$ = 888.5 kJ/mol, the protonation site being the carbonyl oxygen)[146,150] is higher than those of the

cyanophenyl and nitrophenyl groups (**Table 2.2**). This suggests that the amide carbonyl groups would more frequently participate in strong hydrogen bond formation than either the cyanophenyl or nitrophenyl group. Thus, more populated conformations that could induce the formation of the aminoketyl intermediate should contribute to the probability leading to typical ECD cleavage processes. However, backbone fragmentation is inhibited in the presence of the electron predator. This contradiction leads to the implication that, even with the assistance of Coulomb stabilization, the amide π^* orbital cannot capture an electron to form a stable bound state that in turn would be expected to result in backbone fragmentation processes. However, it is possible that the presence of the electron predator could modulate the probability of intramolecular electron transfer from a hign-*n* Rydberg orbital to the amide π^* orbital by intercepting and trapping the electron. This may prevail even when transient conformations of the peptide render electron capture by the amide π^* orbital energetically more favorable.

2.4.3. Comparison of ECD, ETD and the Effect of Augmented Vibrational Excitation

The ECD and ETD experimental methodologies have several different aspects. The electron capture/transfer cross sections are different due to different electron transfer media (*i.e.*, free electron for ECD and anion radical for ETD). Both methods also have dissimilar recombination energies, modified by the EA of the electron transfer reagent. In addition, the time scales associated with different instruments or instrumental parameters during the electron capture/transfer process, followed by dissociation, are different.

Inelastic scattering as well as electron transfer during energetic collisions between electron transfer reagent anions and peptide dications could result in higher internal energies of the resulting peptide cation radicals. Similarly, in the case of ECD, recombination involving

energetic electrons as well as inelastic electron-peptide cation collisions may yield peptide cation radicals with excess internal energy. As a result, it is difficult to assess the internal energy distribution of peptide cation radicals formed by electron capture or transfer reactions. Therefore, we only discuss the recombination energy gained by the electron capture and transfer processes.

In the present work, we used fluoranthene with EA ~0.7 eV for the electron transfer reagent. Therefore, the overall recombination energy of ETD is smaller than that of ECD by ~0.7 eV, and fragmentation yields may be reduced in ETD relative to ECD. As noted above, supplemental activation by collision is required to acquire abundant backbone fragments. However, as seen in **Figures 2.1, 2.3** and **2.4**, the general dissociation patterns in ECD, IRMPD/ECD and ETcaD spectra are not significantly different, including the absence of ECD or ETD type fragmentation of $[3NB+2H]^{2+}$ and $[35DNB+2H]^{2+}$. This similarity leads to the conclusion that the overall recombination energy gained by either electron capture or transfer does not affect subsequent fragmentation processes. Excess vibrational excitation, provided either by IR photon absorption or by collisions with an inert gas, also does not produce any significant difference, which also indicates that the levels of vibrational excitation for dissociating ion populations in each case are similar.

2.5. Conclusion

We have elucidated some key aspects of the mechanism of electron capture dissociation and electron transfer dissociation of doubly protonated peptides. The 20 common amino acids, in the absence of post-translational modifications, do not have positive electron affinities. Using the model peptide FQpSEEQQQTEDELQDK, we have modified the phosphoserine residue to incorporate a range of functional groups of widely varying electron affinity, include propanyl, benzyl, 4-cyanobenzyl, perfluorobenzyl, 3,5-dicyanobenzyl, 3-nitrobenzyl and 3,5-dinitrobenzyl

structural moieties, having a range of EA from –1.15 to 1.65 eV, excluding the propanyl group. Typical ECD or ETD backbone fragmentations are completely inhibited in peptides with substituent tags having EA over 1.00 eV, which we refer to as electron predators. The kinetics of the initial electron capture are not modified by the presence of the electron predators, consistent with the expectation that electron capture kinetics are governed by the long range electron-dication interaction. Once an electron is captured to high-*n* Rydberg states, however, we propose that through-space or through-bond electron transfer to the EA-tuning tags or low-*n* Rydberg states via potential curve crossing occurs in competition with transfer to the amide π^* orbital. This conjecture is supported by time-dependent density functional theory applied to a series of reduced model systems. The intramolecular electron transfer process is modulated by structure-dependent hydrogen bonds and is heavily affected by the presence and type of electron withdrawing groups in the EA-tuning tag. The anion radicals formed by electron predators have high proton affinities (approximately 1400 kJ/mol for the 3-nitrobenzyl anion radical) in comparison to other basic sites in the model peptide dication, facilitating exothermic proton transfer from one of the two sites of protonation. This forms a stable radical intermediate and interrupts the normal sequence of events in ECD or ETD leading to backbone fragmentation through the intermediacy of an aminoketyl radical which fragments by β-cleavage of the adjacent $N-C_\alpha$ bond. Even in the presence of Coulomb stabilization from nearby charges it does not appear that one can infer that the amide π^* orbital can compete with the electron predators, with electron affinities in excess of 1.0 eV, as the eventual site of localization of the captured electron.

The phenynitronic group formed by sequential electron and proton transfer to a nitrophenyl group in a peptide undergoes a facile hydroxyl loss. This process provides an explanation for the unusual peak observed in MALDI MS of peptides containing a nitrophenyl group, 16 Da less than $[M+H]^+$. It indicates the role of electrons in charge reduction processes converting multiply

charged peptides and proteins to the more usual singly charged ions observed in MALDI MS. Nitration of tyrosine is an important post-translational modification associated with cell signaling pathways and oxidative inflammatory responses.[104,105] Interestingly, this process introduces an electron predator that exhibits behavior similar to what we observe with our derivatized peptides.[151] We are exploring the possibility that this can be exploited to facilitate the detection of trace peptides where this PTM is present.

2.6. Acknowledgement

This work was supported by the National Science Foundation through grant CHE-0416381 and the Beckman Institute at California Institute of Technology. The computational resource was kindly provided by the Materials and Process Simulation Center at California Institute of Technology. C. H. S. acknowledges a fellowship from the Kwanjeong Educational Foundation. P.R. acknowledges support from the NIH/NIDCR UCLA Research Training Program (T32 DE007296). The NIH/NCRR High-End Instrumentation Program supported the acquisition of the LTQ-FT mass spectrometer (grant S10 RR023045 to J.A.L.). The authors thank Professor Jack Simons for discussions regarding the electron capture process, Professor Woon-Seok Yeo for help with synthesis, Professor Francis Turecek, Dr. Yousung Jung and Dr. Jiyoung Heo for assistance with the computational analysis, and Dr. Hugh I. Kim for discussions of reaction mechanisms.

Supporting Information Available:

^{1}H-NMR peaks of thiol compounds, synthesis of 3,5-dicyanobenzyl thiol, ECD and IRMPD/ECD spectra of 2-nitrobenzyl, 4-nitrobenzyl and N_{α}-3,5-dinitrophenyl derivatized peptides, geometries, energetics and molecular orbitals of the model species from quantum mechanical calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

3. Investigation of the Mechanisms of Inter- and Intramolecular Disulfide Bond Cleavages in Model Peptides by Covalently Attached Acetyl Radical

Abstract We investigate the mechanism of disulfide bond cleavage in gaseous peptide ions triggered by a regiospecific covalently attached acetyl radical. We describe a second generation 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO)-based free radical initiated peptide sequencing (FRIPS) reagent (**10**) that yields acetyl radical peptide ions in a single step of collisional activation. Derivatized model peptide ions containing inter- or intramolecular disulfide bonds are collisionally activated to yield an N-terminal acetyl radical group via loss of TEMPO radical. Further collisional activation of the acetyl radical peptide results in highly selective C−S and S−S bond cleavages. To probe their mechanisms, model peptides having a deuterated β-carbon at the disulfide bond are employed. It is suggested that the major pathway of S−S bond cleavage is triggered by hydrogen abstraction from the α-carbon, followed by radical substitution to the S−S bond, yielding thiirane and thiyl radical products. A minor contribution by direct radical substitution to the disulfide bond is also considered. Density functional theory calculations are performed to explore energetics of the proposed mechanisms for disulfide bond cleavage.

3.1. Introduction

As an important post-translational modification, identification and characterization of disulfide bonds in proteins are critical for their three-dimensional structure determination.[152] The disulfide bond, a strong covalent linking of two cysteine residues between protein subunits, contributes significantly to stabilization of tertiary protein structures[153,154] and helps to maintain protein activity in cellular environments.[155-158] Despite their important role, analysis of disulfide bonds in

proteins remains a challenging task exacerbated by their fragility toward redox stress. The native disulfide connectivity can be easily lost by reduction and re-oxidation of disulfides, which may occur randomly during sample preparation for analysis. To avoid this problem, pretreatment of disulfides by reduction and alkylation is often performed, with concomitant loss of important structural information.[159]

Not surprisingly, the rapid expansion of experimental methodology employing high performance mass spectrometry (MS) in protein identification and structural determination has included the development of new approaches for disulfide bond characterization.[160] Low-energy collisional activation of protonated peptides containing disulfide bonds usually leads to a mixture of amide backbone and disulfide C−S bond cleavage, with essentially no S−S bond rupture due to the higher activation energy required for this process.[161] Protein digests generated by proteases that retain intact disulfide bonds typically contain both inter- and intramolecular disulfide linkages. Cleavage of intermolecular disulfide bonds leads to separated peptide fragments and further collisional activation can yield fragments revealing the point of connection. Intramolcular disulfide bonds require multiple steps of activation to locate the linkage sites. Therefore, only limited structural information can be acquired by conventional low-energy collision induced dissociation (CID) of protonated peptides containing intramolecular disulfide bonds.[162] Some of the low-energy CID approaches with certain limited conditions generate more information-rich fragments. Gaseous peptide ions lacking mobile protons typically exhibit highly selective C−S bond cleavages by low-energy CID.[163] This effect is especially prominent in singly protonated disulfide containing peptide ions produced by matrix-assisted laser desorption ionization (MALDI).[164] CID of anionic disulfide-bridged peptides also generates cleavage products from C−S bond fragmentation but their intensities are usually weak and the pattern of fragmentation is complex.[165-167]

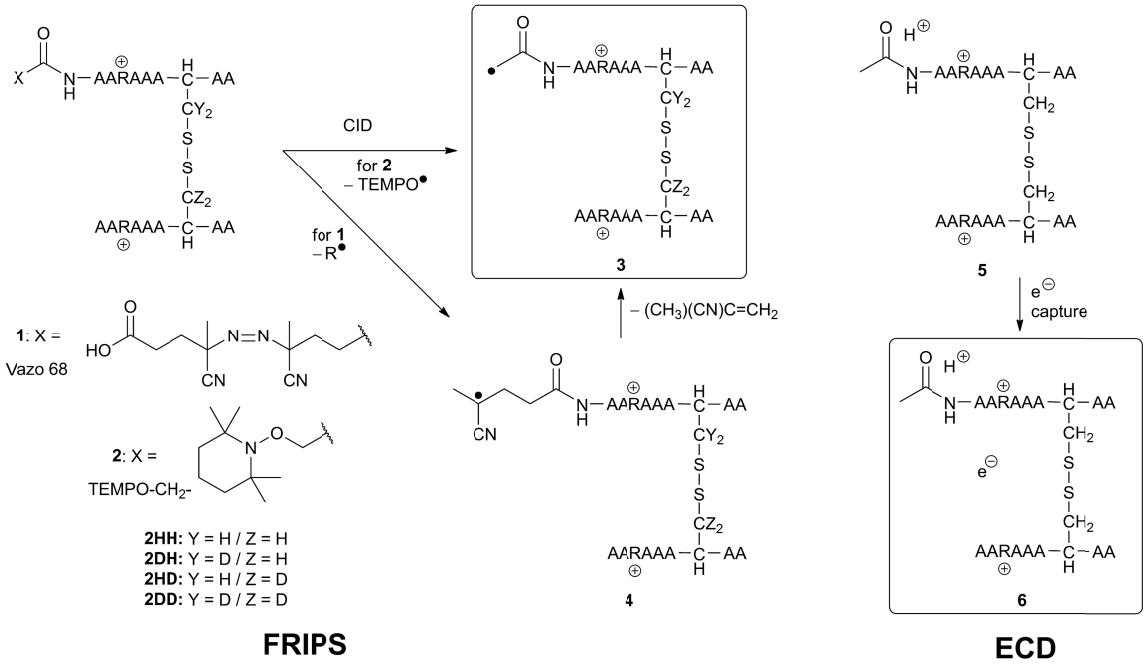
Metal cationized disulfide containing peptides have also been thoroughly investigated by MS. The patterns of disulfide fragmentation with various metal complexes under low-energy CID are diverse.[168] For example, disulfide bond containing peptides cationized by a gold cation undergo efficient S−S bond cleavages by low-energy CID.[169] Alkali or alkaline earth metal–peptide complexes cleave C−S bonds, yielding highly selective H_2S_2 loss.[39,40] This signature neutral loss can be used for fast screening of disulfide containing peptides in complex mixtures, including those resulting from peptic digests.[39,40] The observed processes are triggered by anionic enolation of cysteine residues at backbone C_α positions by metal cations, followed by sequential cleavage of the C−S bonds.

Electron capture dissociation (ECD)[13] and its variation, electron transfer dissociation (ETD)[14], have been very attractive methods for analysis of disulfide linkages, deriving advantage from selective cleavage of S−S bonds in peptides and proteins.[53] In ECD and ETD, the ion activation occurs via electron capture or transfer to multiply charged peptide cations, followed by electronic state relaxation via internal conversion.[170] Initially, dissociative addition of a hydrogen atom to the disulfide bond was suggested to explain the selective S−S bond fragmentation (the Cornell mechanism).[53,171] Direct electron capture or intramolecular electron transfer from high-lying orbitals to the Coulomb stablized σ* orbital of the S−S bond were alternatively proposed.[72,74,76,77,79,80,172,173] Even when all charged sites in the peptide cations are fixed-charge groups such as quaternary ammoniums, S−S bond cleavage is still observed in the absence of a mobile hydrogen atom.[100] This observation supports an alternative mechanistic viewpoint, in which the Coulomb stabilization between positively charged sites and disulfide bonds promotes exothermic intramolecular electron transfer to the σ* orbital of the S−S bond (mainly the Utah part in the Utah-Washington mechanism). The detailed processes for initial electron capture and

subsequent S−S bond cleavage in various disulfide bond containing peptides and proteins remain an active subject for further experimental and theoretical investigations.[174]

We have previously described a method for free radical initiated peptide sequencing (FRIPS) via multistep collisional activation of peptides conjugated with a reagent (Vazo 68, **Scheme 3.1**) that introduces a regiospecific free radical center.[15] FRIPS of peptide conjugates can effect backbone fragmentation, producing mainly a- or x- and z-type ions whose charges are located at the N- or C-terminal sides of peptides, respectively. These fragment ions are not usually observed in low-energy CID spectra of protonated peptide cations. The radical center produced by FRIPS typically reacts by H-atom abstraction either from α- or β-carbons, followed by β-cleavage. This leads primarily to side-chain loss and backbone fragmentation, respectively.

In this paper, we apply our FRIPS methodology to characterize disulfide bond connectivity in peptides and their mechanisms of disulfide bond cleavages are investigated. In addition, we compare and contrast the distinctly different processes associated with FRIPS and ECD that trigger radical-driven cleavages in the gas phase. Our earlier study employed Vazo 68, a commercially available water soluble free-radical initiator, conjugated to the peptide N-terminus to induce hydrogen deficient radical-driven fragmentations (**Scheme 3.1**, **1**).[15]



Scheme 3.1

Upon sequential collisional activation, a regiospecific acetyl radical group (**3**) is produced, which abstracts C_α and C_β hydrogens with concomitant β-cleavage to yield mainly side-chain losses and backbone fragmentations, respectively. In the present study, we employ a second generation FRIPS reagent (**Scheme 3.2**, **10**) inspired by Lee *et al.*[175] The peptide sequencing performance of this reagent is validated with model peptides including the tryptic peptide, HSDAVFTDNYTR, the intramolecular disulfide bond containing peptides, Arg8-Vasopressin, and Arg8-Conopressin G, and the intermolecular disulfide bond containing peptide from a tryptic digest of Arg8-Conopressin G (**Scheme 3.4**). To probe the mechanisms of disulfide bond cleavages, the model peptides having one or two β-deuterium(s) at the disulfide bond are employed (**Scheme 3.1**, **2** with no deuterium (**2HH**), β-deuteriums at the α-chain (**2DH**), at the β-chain (**2HD**), and at both chains (**2DD**), respectively). To further study the reactivity of gas-phase peptide radical ions (**3**

versus **6**) with different hydrogen environments (*i.e.*, whether the system is "hydrogen-rich" or "hydrogen-deficient"), we prepared the model peptide systems depicted in **Scheme 3.1**. Acetylation of the model peptide AARAAACAA, followed by disulfide bond formation with the parent peptide, provides a structural mimic (**5**) of the regiospecific acetyl radical peptide cation (**3**) generated by FRIPS reagent conjugates (**1** or **2**). Electron capture by the triply protonated mimic (**5**) produces the charge-reduced doubly protonated ion (**6**). The two peptide cations **3** and **6** that are produced, respectively, by collsional activation of FRIPS conjugates and ECD, are differentiated by a localized radical center in the "hydrogen-deficient" species (**3**) vs. the presence of a labile hydrogen atom in the "hydrogen-rich" species (**6**). We show that activation of both species leads predominantly to cleavage of the disulfide linkage. Reaction mechanisms are proposed for each, showing that distinct reactivity in each model system results from the nature of the radical centers and structural features which constrain reaction energetics. Quantum chemical calculations using third generation meta-hybrid density functionals (BMK[176], M05-2X[177] and M06-2X[131], chosen for their better performance in organic radical reactions) along with the conventional B3LYP[128,129] functional were performed to quantify energetics of observed reaction processes.

3.2.Experimental

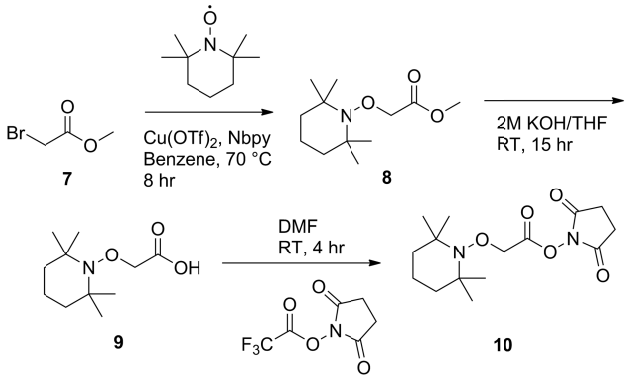
3.2.1. Peptide Fragmentation Nomenclature

Peptide fragmentation nomenclature used in this study for c- and z-type ions produced by ECD and ETD versus a- and x-type ions produced by FRIPS is explained in details in Supporting Information. Briefly, we follow the ECD type nomenclature described in the first ECD paper[13] for both FRIPS and ECD fragment ions.

3.2.2. Materials

The model peptides, AARAAACAA, in 95% peptide purity, and AARAAACAA with a deuterated β-carbon at the cysteine residue, in 98% isotope purity, and 75% peptide purity, were purchased from Biomer Technology LLC (Pleasanton, CA, USA). Vasoactive intestinal peptide (VIP) from human, residue 1-12 (HSDAVFTDNYTR) that simulates tryptic peptides, and Arg8-Vasopressin (CYFQNCPRG-NH_2, Cys1 and Cys6 are connected via a disulfide bond) were purchased from AnaSpec (San Jose, CA, USA). Arg8-Conopressin G (CFIRNCPRG-NH_2, Cys1 and Cys6 are connected via a disulfide bond) was purchased from Bachem (Torrance, CA, USA). A soluble free-radical initiator, Vazo 68, was purchased from DuPont (Wilmington, DE, USA). Acetic anhydride was purchased from Mallinckrodt Inc. (Phillpsburg, NJ, USA). All solvents are HPLC grades and were purchased from EMD Merck (Gibbstown, NJ, USA). Sequencing grade TPCK treated trypin was purchased from Promega (Madison, WI, USA). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise specified. For desalting, OMIX 100 μL size C-18 tips were purchased from Varian Inc. (Palo Alto, CA, USA).

3.2.3. Synthesis of TEMPO-based FRIPS Reagent



Scheme 3.2

Synthesis strategy for the second generation TEMPO-based FRIPS reagent (**10**) is summarized in **Scheme 3.2**. The previous development by Lee *et al.*[175] was modified by replacing 2-(bromomethyl)benzoic acid methyl ester with methyl 2-bromoacetate for the current study. Briefly, TEMPO is coupled to the acetyl methyl ester group, followed by deprotection and activation of the carboxylic acid group. Ten mg/mL of the final product (**10**) in acetonitrile (ACN) was reacted with 50 μg of model peptides in 100 mM phosphate buffer (pH 8.5) for 2 hr. After desalting, the resulting peptide conjugates were analyzed by electrospray ionization (ESI) ion trap mass spectrometers. See Supporting Information for details on the synthesis of the TEMPO-based FRIPS reagent and peptide conjugation.

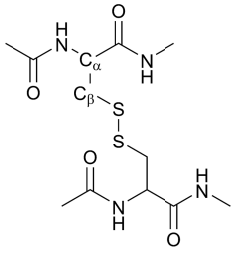
3.2.4. Mass Spectrometry

FRIPS experiments were performed using a LCQ Deca XP ion trap mass spectrometer (Thermo, San Jose, CA, USA). Prepared peptide solutions were directly infused to the mass spectrometer at 3 μL/min by a syringe pump with an electrospray ionization source. Critical parameters of the mass spectrometer include spraying voltage 3.5 kV, capillary voltage 41~42 V, capillary temperature 275 °C and tube lens voltage −50~60 V. Other ion optic parameters were optimized by the auto-tune function in the LCQ tune program for maximizing the signal intensity. The precursor isolation window for MS^n experiments was set to 3.5 m/z and normalized collisional energy in the LCQ tune program was varied from 23% to 28% based on residual precursor ion intensities. For FRIPS and CID spectra, 100 scans were recorded.

ECD was performed in the Proteome Exploration Laboratory of the Beckman Institute at Caltech using a 7-tesla linear ion trap Fourier transform ion cyclotron resonance (LTQ-FTICR) mass spectrometer (Thermo, San Jose, CA, USA) equipped with the Nanomate (Advion BioSciences Inc., Ithaca, NY, USA) nanospray unit. The spraying voltage was 1.4 kV and the gas

pressure was 0.3 psi. Critical parameters of the mass spectrometer include capillary voltage 49 V, capillary temperature 200 °C and tube lens voltage 180 V. Other ion optic parameters were optimized by the auto-tune function in the LTQ tune program for maximizing the signal intensity. The target resolution at 400 m/z was set to 100,000. The precursor isolation window for ECD experiments was set to 5.0 m/z and electron irradiation occurred for 100ms at 5% of the full energy scale in the LTQ tune program, corresponding to electron energies less than approximately 1 eV. For ECD spectra, 100 scans were recorded. In particular, the TEMPO-based FRIPS reagent conjugates of intermolecular disulfide bond containing peptides (**2**), Arg8-Conopressin G, and trypsin digested Arg8-Conopressin G were analyzed by ion trap scans in a hybrid LTQ-FTICR mass spectrometer with the nanospray and ion optics conditions described above.

3.2.5. Quantum Chemical Calculation



N,N′-diacetyl-cystine-*N*-methylamide

Scheme 3.3

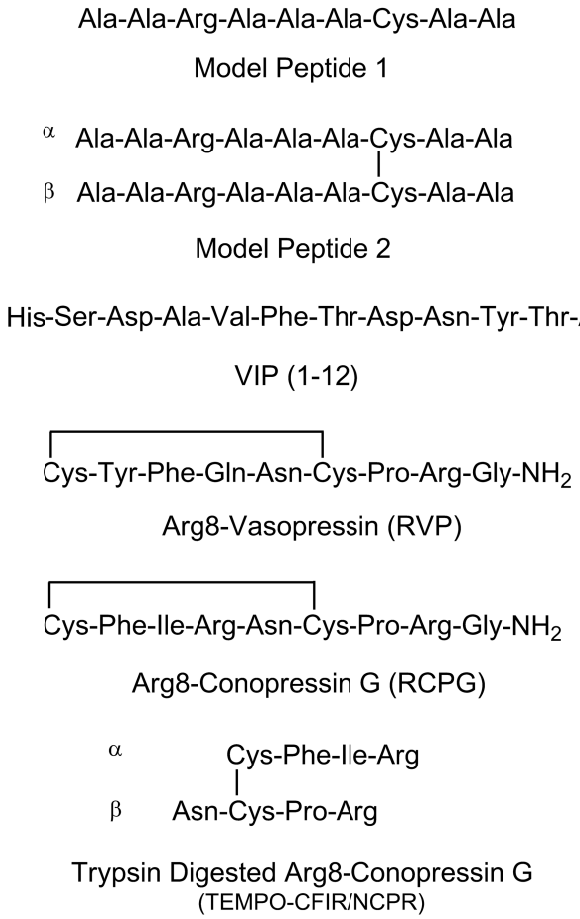
N,N′-diacetyl-cystine-*N*-methylamide (**Scheme 3.3**) was used for the simple model system to describe the intermolecular disulfide bond linked peptide used in this study, CH_3CONH-[AARAAACAA]-S-S-[AARAAACAA] (**5**). Initial geometries were generated by the MC/MM conformer search using Macromol 8.0 (Schrödinger Inc., Portland, OR, USA) as implemented in

Maestro 8.0 (Schrödinger Inc., Portland, OR, USA) under the Linux environment. The OPLS 2005 was used for the force field model. Within 5 kcal/mol energy, all low-energy conformers were initially recorded. After manual screening of obtained structures to avoid redundancy, low-energy conformers were selected for further structure optimization by density functional theory (DFT). For radical species, *ad hoc* assignment of the formal charge state was employed to implement MC/MM calculations for conformer searching. In particular, conformers of the C_β hydrogen abstracted *N,N′*-diacetyl-cystine-*N*-methylamide radical were searched by substitution of the β-carbon with boron to simulate the trigonal bonding environment. Each conformer was subject to a geometry optimization using Jaguar 7.5 (Schrödinger Inc., Portland, OR, USA) at the B3LYP/6-31G(d) level. By monitoring the occurrence of imaginary vibrational frequencies, only non–transition state structures (*i.e.*, no imaginary vibrational frequency) were further optimized using a higher basis set at the B3LYP/6-311++G(d,p) level. The transition state structures were searched using the LST or QST methods by interpolating initial guesses for reactants, products and transition states. Single point energy was refined using Q-Chem 3.1 (Q-Chem Inc., Pittsburg, PA, USA) by the BMK, M05-2X, M06-2X and B3LYP density functionals with the 6-311++G(3df,3pd) basis set. The three new generation meta-hybrid functionals other than B3LYP were chosen for their ability to more reliably predict the energetics of organic radical reactions.[131,176] For the open-shell systems, the spin-unrestricted method was used. All calculations were performed using computational resources kindly provided by the Material and Process simulation center at the Beckman Institute, Caltech.

3.3.Results and Discussion

All peptides used in this study are shown in **Scheme 3.4**. The model peptide 1 and 2 are the simple cysteine and intermolecular disulfide bond containing peptides. The N-terminal amine of

the α-chain in the model peptide 2 is acetylated for ECD or conjugated by FRIPS reagents. The β-deuteriums are inserted in each peptide chain of the model peptide 2 to probe the mechanism of disulfide bond cleavage (See also **Scheme 3.1**). The VIP (1-12) peptide is used for the validation of sequencing performance of the newly prepared FRIPS reagent (**10**). This model system simulates FRIPS spectra of tryptic peptides which are produced by trypsin, the most popular protease in the MS-based bottom-up proteomics analysis. Arg8-Vasopressin and Arg8-Conopressin G are chosen to investigate the reactivity of the acetyl radical with the intramolecular disulfide bond in various charge states (*i.e.*, +1 for Arg8-Vasopressin and +2 for Arg8-Conopressin G). These two peptides share many structural features such as an intermolecular disulfide bond between Cys1 and Cys6, and the amidated C-terminal, which form the vasopressin family peptides and show similar bioactivities.[178] An intermolecular disulfide bond containing model peptide is produced by trypsin digestion of the TEMPO-based FRIPS reagent conjugate of the Arg8-Conopressin G (TEMPO-CFIR/NCPR). This model peptide is introduced to investigate the reactivity of the acetyl radical with an intermolecular disulfide bond when other amino acids having more reactive hydrogens of the α- and β-carbons than those of alanine are present.



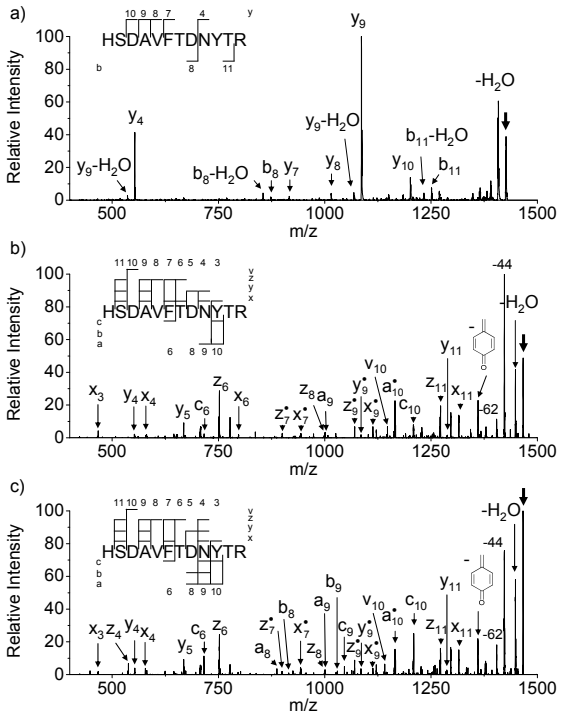
Scheme 3.4

3.3.1. VIP (1-12)

Evaluation of the second generation TEMPO-based FRIPS reagent is performed by singly protonated HSDAVFTDNYTR peptide cation (**Figure 3.1**). The CID spectrum of singly protonated VIP is dominated by y_4 and y_9 ions from the low-energy salt bridge cleavage pathway at the C-terminal sides of aspartic acids in the presence of the protonated arginine residue, yielding poor sequence coverage (**Figure 3.1a**, 6 out of 11, 55%).[10,11] The FRIPS spectra of singly protonated VIP from Vazo 68 and TEMPO-based FRIPS reagents exhibit extensive

backbone cleavage fragments which cover almost the full sequence of the peptide (**Figures 3.1b** and c, 10 out of 11, 91%). Almost all ion types of backbone fragments such as a, b, c, v, x, y and z ions appear in both **Figures 3.1b** and **c**. Not unexpected, the spectra from the two different FRIPS reagent conjugates (**Figures 3.1b** and **c**) share many common ions and their intensities are nearly identical. Observed backbone fragments are produced by either H-atom abstraction from the β-carbons, followed by β-cleavage (a and $x^{\bullet}$ ions, and $z^{\bullet}$ ions by CONH loss from $x^{\bullet}$ ions) or the charge-remote amide bond cleavage (b and y ions). H-atom transfer between a-type neutral fragment and $x^{\bullet}$ ion after β-cleavage may produce even electron x ions. This phenomenon is more dominant for smaller fragment ions such as x_3 and x_4. Interestingly, the formation of c_6/z_6 ions and $a_{10}^{\bullet}/c_{10}$ ions at the N-terminal of the threonine residues (Thr7 and Thr11) that cannot be simply explained by H-atom abstraction from the β-carbons, followed by β-cleavage is observed (**Figures 1b** and **c**). The same types of fragment ions nearby the threonine or serine residues are also found in the previous studies.[179,180] We are investigating the mechanism of this process and it will be independently reported. Initiation of side-chain loss usually occurs by H-atom abstraction from the α- or γ-carbons of backbone amino acids, followed by β-cleavage. Types of side-chain loss in hydrogen-deficient peptide radicals are thoroughly investigated by Sun *et al*.[179] Carbon dioxide loss (−44 Da) is induced by H-atom abstraction from the carboxyl groups at the aspartic/glutamic acid residues or the C-terminal carboxylic acid. Isobaric $\bullet CONH_2$ neutral loss produced by the H-atom abstraction from the α-carbon of asparagine, followed by β-cleavage may partially account for −44 Da loss in **Figures 1b** and **c**. Phenolic H-atom abstraction from tyrosine yields loss of para-quinone methide (−106 Da) via radical rearrangement of the phenoxyl radical group (**Figures 1b** and **c**).

Figure 3.1 CID and FRIPS of singly protonated HSDAVFTDNYTR



The CID spectrum (a) and FRIPS (Vazo 68 and the TEMPO-based reagent) spectra (b and c, respectively) of singly protonated HSDAVFTDNYTR. The CID spectrum (a) is dominated by y_4 and y_9 ions from the low-energy salt bridge cleavage pathway at the C-terminal sides of aspartic acids, effected by the protonated arginine residue. This results in poor sequence coverage (6 out of 11, 55%). b) and c) are nearly identical, yielding mainly backbone cleavage fragments (10 out of 11, 91% sequence coverage) and neutral losses. Bold arrows indicate the precursor ions.

3.3.2. Arg8-Vasopressin

Figures 3.2a-d depict FRIPS of Arg8-Vasopressin. The TEMPO-based FRIPS reagent (**10**) was conjugated to the N-terminal amine of Arg8-Vasopressin with a conversion yield of approximately 90% based on the relative signal intensities between FRIPS reagent conjugated and unmodified Arg8-Vasopressin peaks in **Figure 3.2a**. The singly protonated TEMPO-based FRIPS reagent conjugate of Arg8-Vasopressin (m/z 1281) is collisionally activated to generate the regiospecific acetyl radical cation (m/z 1125) by loss of TEMPO radical (**Figure 3.2b**). This process is energetically favored to produce the abundant acetyl radical cation. Thus it allows us to proceed to the following CID experiments up to MS4 for peptide sequencing by ensuring strong ion signals. It is less practical when Vazo 68 is used with the consequence that MS5 is required to characterize the intramolecular disulfide bond in Arg8-Vasopressin. Collisional activation of the acetyl radical cation (m/z 1125) induces mainly CH_2S loss (m/z 1079) by cleaving the S−S bond.

Figure 3.2 FRIPS and ECD of Arg8-Vasopressin

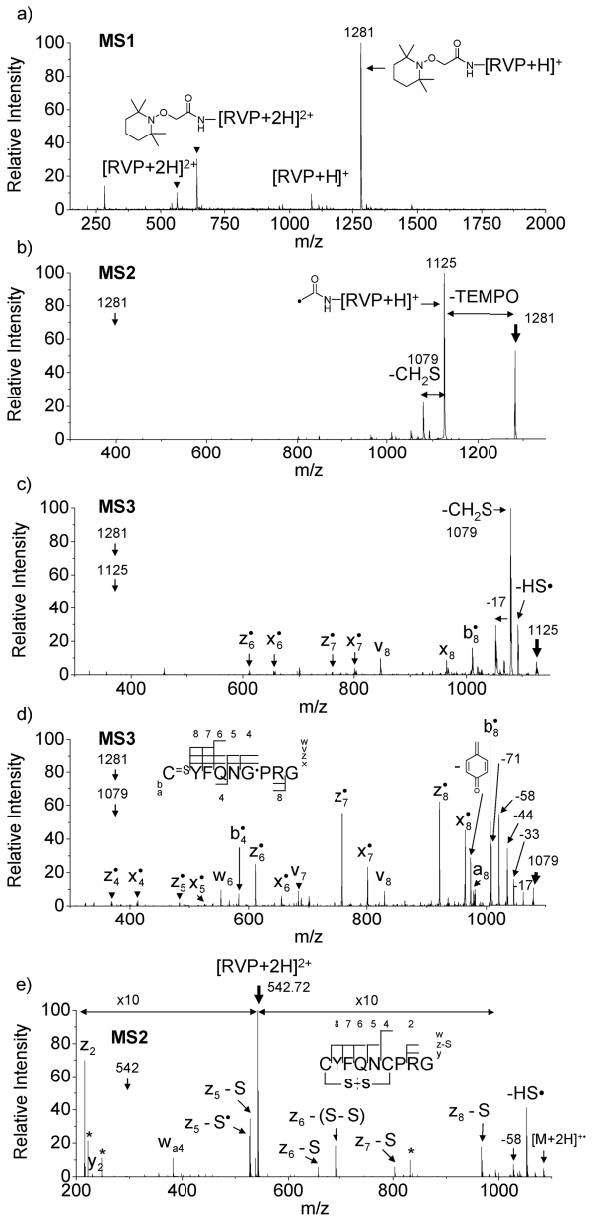
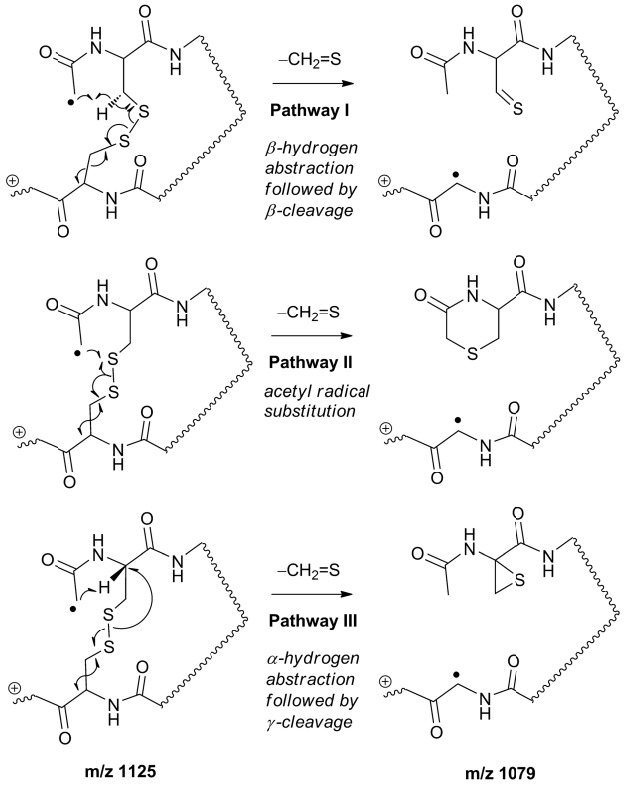


Figure 3.2. FRIPS and ECD of Arg8-Vasopressin. a) ESI-MS1 of the TEMPO-based FRIPS reagent conjugate of Arg8-Vasopressin. b) CID of the singly protonated TEMPO-based FRIPS reagent conjugate of Arg8-Vasopressin, m/z 1281 (MS2). c) CID of the acetyl radical cation, m/z 1125 (MS3). d) CID of the CH_2S loss product from the acetyl radical cation, m/z 1079 (MS3). RVP is Arg8-Vasopressin, $C^{=S}$ is thioaldehyde, thiomorpholin-3-one or thiirane products, and $G^{\bullet}$ is glycyl α-carbon radical. See **Scheme 3.5** for the proposed reaction mechanisms. e) ECD of the doubly protonated Arg8-Vasopressin. See **Scheme 3.6** for notations of z_n−S ions. * denotes electronic noise peaks. Bold arrows indicate the precursor ions.



Scheme 3.5

We propose that this process is initiated by H-atom abstraction at the β-carbon of Cys1, followed by β-cleavage (**Scheme 3.5**, pathway I). The resulting radical cation at m/z 1079 contains a modified residue whose side chain is thioaldehyde (−CH=S) at Cys1 position (the 2-amino-3-thioxopropanoic acid residue) and the glycyl α-carbon radical residue at Cys6 position. The possibility of H-atom abstraction at the β-carbon of Cys6 was considered but no correlated

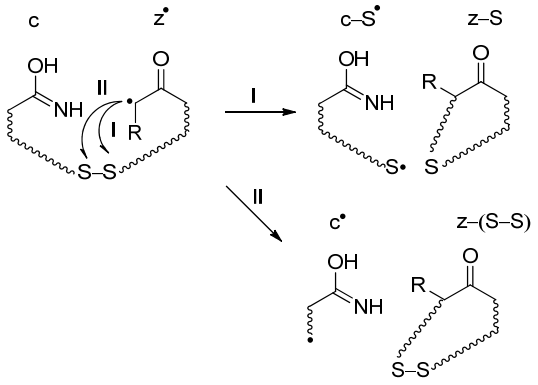
fragments were observed in CID of m/z 1079 (**Figure 3.2d**). Instead, the six-membered ring intermediate favors reaction at the β-carbon of Cys1. No direct β-cleavage from the glycyl α-carbon radical residue (*e.g.*, $b_6^{\bullet}/y_4$ and $b_7/y_3^{\bullet}$) is observed. The glycyl α-carbon radical cation at m/z 1079 is directly isolated from the MS2 stage and further collisionally activated in MS3 (**Figure 3.2d**). Following H-atom abstraction by the glycyl radical at other α- and β-carbon sites, followed by β-cleavage leads to side-chain losses (17, 33, 58, and 71 Da) and backbone fragmentation (b, x, z, v and w ions, **Scheme 3.5**). From these product ions, the peptide sequence and the position of the intramolecular disulfide bond are assigned (**Figure 3.2d**). Compared to what we observed in the previous study of alkali and alkaline earth metal complexes of disulfide bond containing peptides[39], the sequence coverage after CH_2S loss is more extensive, including six out of eight possible backbone fragments (**Figure 3.2d**).

An alternative mechanism for CH_2S loss via dissociative addition of the acetyl radical, so called radical substitution (S_H2) reaction, to the disulfide bond is described in **Scheme 3.5**, pathway II.[181] Radical substitution forms the stable six-membered thiomorpholin-3-one ring structure at the N-terminal, and releases the thiyl radical group by cleaving S−S bond at the counterpart cysteine residue. The residual internal energy after S−S bond cleavage leads to subsequent loss of CH_2S, yielding the glycyl α-carbon radical group at Cys6.

H-abstraction at the α-carbon of Cys1, followed by γ-cleavage is also considered (**Scheme 3.5**, pathway III). The first step of this pathway, H-abstraction reaction at the α-carbon, is energetically favored over abstraction at the β-carbon (tertiary versus secondary hydrogens). Also, the final thiirane product is more stable than thioaldehyde, yielding a thermodynamically favored process. However, in general 1,4-H transfer is rarely observed[182] and its steric hindrance over 1,5-H transfer makes this pathway kinetically less favored. Note that the overall fragmentation results in the pathway II and III after loss of CH_2S are indistinguishable by their mass to charge ratios to

those of the pathway I. In this regard, it is hard to discern the relative portions of each reaction pathway proposed in **Scheme 3.5**. We will later investigate this mechanism with intermolecular disulfide bond containing peptides that may experience less steric hindrance for H-abstraction at the α-carbon via the remote way, not by the direct 1,4 interaction.

ECD of the doubly protonated Arg8-Vasopressin is shown in **Figure 3.2e**. The most prominent process is loss of HS• that is involved with multiple bond cleavages in the disulfide bond. The same neutral loss was also observed in the previous study on the analogue model system, Lys8-Vasopressin.[53] Many N−C_α bonds are cleaved by ECD (6 out of 8, z_2, w_4, and $z_{5\text{-}8}$) and the observed sequence coverage is the same as FRIPS.



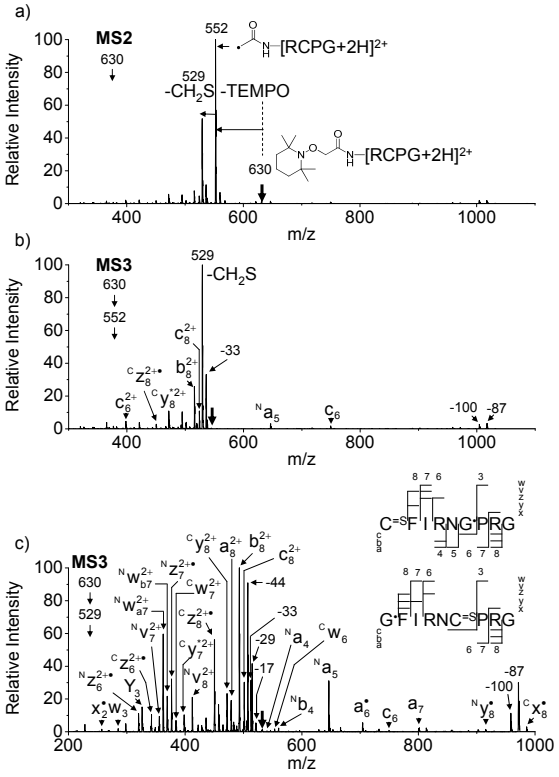
Scheme 3.6

One possible mechanism for intramolecular disulfide bond cleavage by ECD was previously proposed as follows.[53] The α-carbon radicals of the nascent $z^\bullet$ ions generated by N−C_α bond cleavage may attack one of the sulfur atoms in the disulfide bond, followed by disulfide bond cleavages (**Scheme 3.6**, pathway I for S−S bond cleavage and pathway II for C−S bond cleavage). As a result, two covalent bonds that form a cyclic peptide structure are cleaved, releasing z_n−S ions (**Figure 3.2e**). As an exception, the w_{a4} ion is formed via β-cleavage of the α-carbon radical

of the nascent $z^{\bullet}$ ion by cleaving a C−S bond (**Figure 3.2e**). This previously suggested mechanism explains both N−C_α bond and intramolecular disulfide bond cleavages. In addition, regardless of the actual mechanism for the formation of the nascent $z^{\bullet}$ ions (the Cornell versus the Utah-Washington), the final results after disulfide bond cleavage are identical.

H-atom transfer to one of the sulfur atoms, followed by disulfide bond cleavage by forming the thiyl radical (−S•) and thiol groups (−SH) is also considered.[53,171] The relatively high H-atom affinity of the disulfide bond makes this mechanism more plausible.[53,171] This process can also explain the observed dominant process for disulfide bond cleavage compared to N−C_α bond cleavages. For example, the most abundant product from HS• loss (**Figure 3.2e**) is associated with multiple bond cleavages, which is more easily explained by direct dissociative H-atom addition to the disulfide bond. The nascent thiyl radical can abstract H-atom from the α-carbon of the counterpart cysteine residue.[183,184] The resulting the α-carbon radical can undergo β-cleavage, yielding HS• loss. Direct electron capture or intramolecular electron transfer to the Coulomb stabilized σ* orbital of the S−S bond, followed by proton transfer yields the same thiyl radical and thiol products.[72]

Figure 3.3 FRIPS of doubly protonated Arg8-Conopressin G



FRIPS of doubly protonated Arg8-Conopressin G. a) CID of doubly protonated TEMPO-based FRIPS conjugate of Arg8-Conopressin G at m/z 630 (MS2). b) CID of the acetyl radical containing Arg8-Conopressin G dication at m/z 552 (MS3). c) CID of the CH_2S loss dication from the acetyl radical dication at m/z 529 (MS3). RCPG is Arg8-Conopressin G, $C^{=S}$ is thioaldehyde, thiomorpholin-3-one or thiirane products, and $G^{\bullet}$ is glycyl α-carbon radical. Superscripts ("N" and "C") on the left side of the fragment ions indicate the position of the $C^{=S}$ residue that is located at the N-terminal or the C-terminal of the peptide. * denotes ammonia loss. Bold arrows indicate the precursor ions.

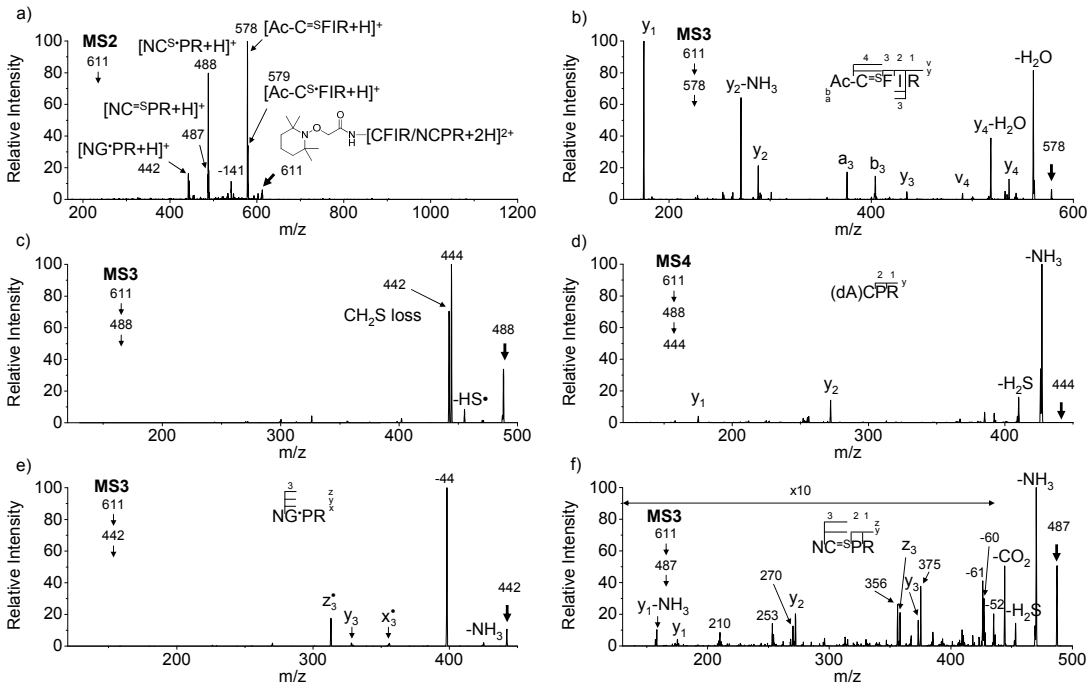
3.3.3. Arg8-Conopressin G

FRIPS spectra of doubly protonated Arg8-Conopressin G are shown in **Figure 3.3**. Compared to the FRIPS results for singly protonated Arg8-Vasopressin (**Figure 3.2**), we investigate the effect of the charge state of the precursor ion on the FRIPS fragmentation pathways (*i.e.*, disulfide bond, backbone and side-chain) with two protonated arginine residues. The regiospecific acetyl radical dication (m/z 552) is formed by collisional activation of the TEMPO-based FRIPS reagent conjugate of Arg8-Conopressin G dication (m/z 630, **Figure 3.3a**). Collisional activation of the acetyl radical dication (m/z 552) mainly produces the disulfide bond cleavage product (m/z 529) via CH_2S loss (**Scheme 3.5**). Concomitant losses of TEMPO radical and CH_2S occur regardless of the charge state (+1 or +2) of the precursor ions in both our model intramolecular disulfide bond containing peptides (**Figures 3.2** and 3.**3**). The prominent CH_2S loss is also observed in the MS2 stage (**Figure 3.3a**), which allows us to perform peptide sequencing in the MS3 stage with improved ion signals by reducing one step of collisional activation (**Figure 3.3c**). Interestingly, compared to the FRIPS spectrum of singly protonated Arg8-Vasopressin (**Figure 3.2d**), the FRIPS spectrum of doubly protonated Arg8-Conopressin G is the mixture of fragments of two possible glycyl radicals produced by loss of CH_2S at each cysteine residue. This indicates that the gas-phase conformation of doubly protonated Arg8-Conopressin is more flexible, allowing the acetyl radical group to react with both Cys1 and Cys6 residues. Although the peptide sequence of Arg8-Conopressin G is very similar to Arg8-Vasopressin, it is believed that the additional protonated arginine residue within the region bounded by the disulfide bond dramatically modifies its gas-phase structure, showing the distinct fragmentation pattern by forming two possible glycyl radical dications. As a result, the observed fragment ions from collisional activation of CH_2S loss dication (m/z 529) are attributed to two possible glycyl radicals at Cys1

and Cys6 positions, respectively (**Figure 3.3c**, *i.e.*, ${}^{N}z_{6}^{2+\bullet}$ versus ${}^{C}z_{6}^{2+\bullet}$ where the upper-left N or C indicate the positions of the resulting even electron species at the original cysteine residues after disulfide bond cleavage). The overall sequence coverage is extensive, revealing the full amino acid sequence of Arg8-Conopressin G (**Figure 3.3c**, 8 out of 8).

The reactivity of the intermolecular disulfide bond is investigated by collisional activation of doubly protonated TEMPO-CFIR/NCPR (a tryptic digest of the TEMPO-conjugated Arg8-Conopressin G, **Figure 3.4**). This model system simulates tryptic digests of disulfide bond containing proteins where cleavage fragments in part comprise two peptide chains derived from the original protein backbone, held by an intermolecular disulfide bond. Collisional activation of doubly protonated TEMPO-CFIR/NCPR mainly produces products from S−S bond cleavage. Interestingly, the product from TEMPO loss (−156 Da) is not observed (**Figure 3.4a**). Rather, loss of 141 Da (2,2,6,6-tetramethylpiperidine) is observed at m/z 540.8, indicating N−O bond cleavage (**Figure 3.4a**). This product may result from proton transfer from the protonated arginine residue to the TEMPO nitroxide tertiary amine residue and subsequent rearrangement for bond cleavage.

Figure 3.4 FRIPS of doubly protonated TEMPO-CFIR/NCPR



FRIPS of doubly protonated TEMPO-CFIR/NCPR. a) CID of doubly protonated TEMPO-CFIR/NCPR at m/z 611 (MS2). Highly selective S−S bond cleavage is observed, yielding both α- and β- chain cations. b) CID of the α-chain, Ac-$C^{=S}$FIR cation (even electron species) at m/z 578 (MS3), yielding the sequence of the α-chain. c) CID of the β-chain, $NC^{S\bullet}$PR cation (odd electron species) at m/z 488 (MS3). Losses of 44 (•$CONH_2$, m/z 444) and 46 (CH_2S, m/z 442) Da are prominent. d) CID of the β-chain cation at m/z 444 (even electron species, MS4). e) CID of the β-chain cation at m/z 442 (odd electron species, MS3). f) CID of the β-chain cation at m/z 487 (even electron species, MS3). The CID spectra of various β-chain peptide cation (**Figures 5d**, **e**, and **f**) allow sequencing of the β-chain. See Supporting Information for the precursor ion structure, TEMPO-CFIR/NCPR. $C^{=S}$ is thioaldehyde, thiomorpholin-3-one or thiirane products, $C^{S\bullet}$ is thiyl cysteine radical, $G^{\bullet}$ is glycyl α-carbon radical, and "Ac-" is the acetylated N-terminal group except for the N-terminal thiomorpholin-3-one product case. Bold arrows indicate the precursor ions.

The resulting products from S−S bond cleavage have the thiyl radical and the counterpart even electron species, thioaldehyde, thiomorpholin-3-one or thiirane products (**Scheme 3.5**), respectively (**Figure 3.4a**). Subsequent losses of CH_2S or $\bullet CONH_2$ from the thiyl radical at m/z 488 by CID yield the glycyl α-carbon radical (m/z 442) at Cys2 and dehydroalanine (m/z 444) at Asn1, respectively (**Figures 3.4a** and **c**). Collisional activation of the observed products in **Figure 3.4a** (m/z 578 of the α-chain, and m/z 442, 444, and 487 of the β-chain, respectively) provides *complete* sequence information of each peptide chain (**Figures 3.4b**, **d**, **e** and **f**). Even electron species such as cations at m/z 578 and 487 are mostly fragmented at amide bonds, yielding mainly b- and y-type ions with several neutral losses (**Figures 3.4b**, **d** and **f**). The glycyl α-carbon radical at m/z 442 abstracts H-atom from the β-carbon of the asparagine residue, which induces the C_α−C bond cleavage, releasing the $x_3^\bullet$ ion. Concomitant loss of isocyanic acid (O=C=NH) from the $x_3^\bullet$ ion yields the more stable $z_3^\bullet$ ion (**Figure 3.4e**).

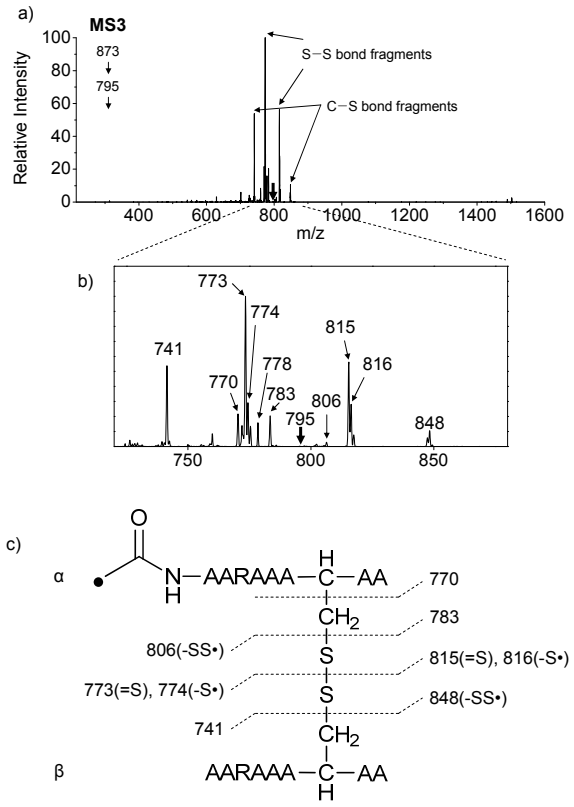
3.3.4. AARAAACAA Dimer

Figure 3.5 demonstrates disulfide bond cleavages effected by the acetyl radical in the model system, **3**. The regioselective acetyl radical group is generated by collisional activation of the doubly protonated TEMPO-based FRIPS reagent derivatized AARAAACAA peptide dimer (**2HH**, m/z 873). Further collisional activation of dication **3HH** severs the disulfide linkage, yielding various C−S (m/z 741, 783, 806, and 848) and S−S (m/z 773, 774, 815, and 816) bond cleavage fragments from each chain (**Figure 3.5b** and **c**). Compared to the FRIPS spectrum of doubly protonated TEMPO-CFIR/NCPR in **Figure 3.4**, abundant C−S bond cleavage fragments are observed in **Figure 3.5**. The difference should be noted that the acetyl radical dication at m/z 795 is subject to collisional activation in the MS3 stage (**Figure 3.5**), whereas the intact TEMPO conjugate is collisionally activated in the MS2 stage, yielding mainly the products from S−S bond

cleavage (**Figure 3.4**). Essentially no backbone fragmentation is observed due to higher bond dissociation energy of the C_β−H bond in alanine residues (**Figure 3.5a**).[179] In addition, S−S bond cleavage is more favored relative to C−S bond cleavage (**Table 3.1**). Relative fragmentation abundances are summarized in **Table 3.1**.

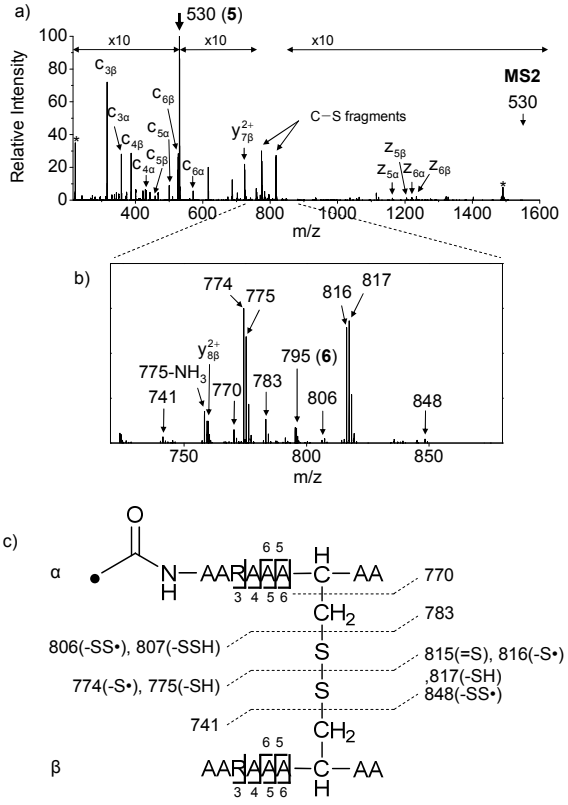
ECD of the triply charged intermolecular disulfide containing model peptide (**5**) is performed for comparison of the reactivity of the nascent charge-reduced radical dication to that of the regiospecific acetyl radical dication generated by FRIPS (**Figure 3.6**). The charge-reduced model peptide radical dication (**6**) produced by electron capture undergoes both backbone and disulfide fragmentations (**Figure 3.6a**). As noted in the introduction, disulfide bond cleavage is one of the most dominant reaction pathways in ECD and the process has been interpreted according to the viewpoints of both the Cornell[53] and Utah-Washington[72] mechanisms. More specifically, even compared to FRIPS of **2HH**, ECD is dominated by S−S bond cleavage, in preference to other C−S bonds, and backbone fragmentations leading to c- and z-type ions (**Figure 3.6**, **Table 3.1**).

Figure 3.5 FRIPS of the doubly protonated AARAAACAA disulfide-bridged dimer



a) FRIPS of the doubly protonated AARAAACAA disulfide-bridged dimer (**3HH**, m/z 795). •CH_2−CONH−[AARAAACAA]−S−S−[AARAAACAA] (**3HH**, m/z 795, bold arrow) is generated by collisional activation of TEMPO−CH_2−CONH−[AARAAACAA]−S−S−[AARAAACAA] (**2HH**, m/z 873) via loss of TEMPO radical. Essentially no backbone fragmentation is observed. b) Expansion of the m/z range in which disulfide cleavages occur. c) Scheme showing cleavage sites and fragment m/z values from each chain.

Figure 3.6 ECD of the triply charged alpha chain acetylated AARAAACAA disulfide-bridged dimer

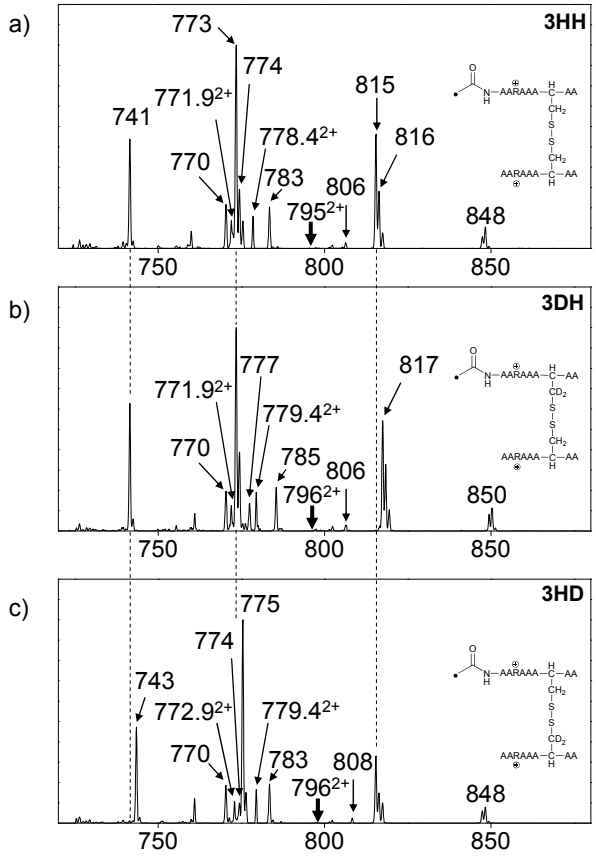


a) ECD of the triply charged alpha chain acetylated AARAAACAA disulfide-bridged dimer (**5**, m/z 530, bold arrow). The charge reduced doubly charged Ac−[AARAAACAA]−S−S−[AARAAACAA] (**6,** m/z 795) is generated by electron capture. Several ECD backbone fragments (c and z ions) are observed. b) Expansion of the m/z range in which disulfide cleavages occur. c) Scheme showing cleavage sites and fragment m/z values from each chain.

Table 3.1 Fragment Ions from FRIPS and ECD of AARAAACAA Disulfide-Bridged Dimer and Their Relative Yields.

FRIPS		ECD	
Fragment Type	Relative Yield (%)	Fragment Type	Relative Yield (%)
Backbone	0.5	Backbone	15.7
$a_{3\alpha}$, $a_{4\alpha}$, $a_{5\alpha}$, $a_{3\beta}$		$c_{3\alpha}$, $c_{3\beta}$, $c_{4\alpha}$, $c_{4\beta}$, $c_{5\alpha}$, $c_{5\beta}$, $c_{6\alpha}$, $c_{6\beta}$, $y_{6\beta}^{2+}$, $y_{7\beta}^{2+}$, $z_{5\alpha}$, $z_{5\beta}$, $z_{6\alpha}$, $z_{6\beta}$	
Side chain loss	1.2	Side chain loss	0
−101 and −89 from Arg			
Disulfide	98.3	Disulfide	84.3
C−S bond cleavage	27.1	C−S bond cleavage	8.3
m/z 741, 783, 806, 848		m/z 741, 783, 806, 807, 848	
S−S bond cleavage	72.9	S−S bond cleavage	91.7
m/z 770, 773, 774, 815, 816		m/z 758, 770, 774, 775, 815, 816, 817	

Figure 3.7 FRIPS of the deuterium-labeled doubly protonated AARAAACAA disulfide-bridged dimers



FRIPS of **3HH**, **3DH**, and **3HD**, (a, b, and c) respectively. Highly selective C−S (m/z 741/743, 783/785, 806/808, and 848/850) and S−S cleavage (m/z 773/775, 774/776, 815/817, 816/818) products are observed. See **Figure 3.5a** for the structures of the products. Almost no backbone fragmentation is observed, thus only the disulfide cleavage region is shown in here. Following CH_2S or CD_2S losses from the thiyl radicals at m/z 816 (a-3HH and c-3HD) or 818 (b-3DH) yield the glycyl α-carbon radical product at m/z 770. No significant change is observed in their relative abundances of the products by S−S bond cleavage ([m/z 817 in **3DH**] / [m/z 815 in **3HH**] = 0.995, [m/z 775 in **3HD**] / [m/z 773 in **3HH**] = ~1) among FRIPS of **3HH**, **3DH**, and **3HD**. The results support pathway III, **Scheme 3.5**.

3.3.5. Deuterium-Labeled AARAAACAA Dimer

To probe the mechanisms of disulfide bond cleavage in the model system, we introduced β-deuteriums at disulfide bonds of the α- and β-chains in the model peptides (See **Scheme 3.1**. **3HH**, **3DH**, **3HD**, and **3DD**, respectively) for FRIPS experiments. **Figure 3.7** shows the FRIPS spectra of **3HH**, **3DH** and **3HD**, respectively. For C−S bond cleavage, H-abstraction at the α-carbon, followed by β-cleavage may occur, yielding the products at m/z 741/743, 783/785, 806/808 and 848/850, respectively. Comparison of peak intensities indicates the difference of an isotopic effect on product distributions. For S−S bond cleavage, if the mechanism is involved with H-abstraction at the β-carbons, significant isotope effects on the fragmentation pattern should be observed from these experiments. Surprisingly, in FRIPS of **3DH**, no significant change is observed in the relative abundances of the products involving S−S bond cleavage ([m/z 817 in **3DH**] / [m/z 815 in **3HH**] = 0.995, **Figure 3.7**).[185] Moreover, in the FRIPS spectrum of **3HD**, corresponding peaks possibly from pathway I in **Scheme 3.5** (H-abstraction at the β-carbon in the β-chain, followed by β-cleavage) are missing (m/z 817) or very small (m/z 774). Rather, the products at m/z 775 and 816 are observed in FRIPS of **3HD**, and their relative intensities appear to be very similar to those in **3HH** ([m/z 775 in **3HD**] / [m/z 773 in **3HH**] = ~1). This major discrepancy arising from peak analysis based on the pathway I or II in **Scheme 3.5** can be resolved by the proposed pathway III in **Scheme 3.5**. First of all, the initial H-abstraction at the α-carbon is not affected by deuterium substitution at the β-carbons. Also, the final products, thiirane and thiyl radicals, can explain the observed peaks, m/z 775 and 816, in FRIPS of **3HD** by having both deuteriums in each peptide chain without abstraction. Therefore, it is proposed that the major process for S−S bond cleavage is initiated by H-abstraction at the α-carbon, followed by γ-cleavage, yielding thiirane and thiyl radical.

The doubly protonated products generated by neutral losses from the acetyl radical dication precursor provide clues for the action of pathway II (direct acetyl radical substitution, **Scheme 3.5**). The doubly protonated product via CH_2S loss is observed at m/z 778. If the acetyl radical is substituted to the sulfur atom in the β-chain side to cleave the S−S bond, the resulting product is an intact linear dication. Namely, a covalent bond is formed between the N-terminal acetyl carbon and the sulfur in the β-chain, and the thiyl radical in the α-chain. Subsequent loss of CH_2S yields the glycyl α-carbon radical as a doubly protonated species (for **3HH** at m/z 771.9 and for **3HD** at m/z 772.9). For FRIPS of **3DH**, loss of CD_2S is observed at m/z 771.9, confirming the proposed mechanism. The S−S bond cleavage product at m/z 815 in FRIPS of **3HD** via the acetyl radical substitution to the sulfur atom in the α-chain side, forming a cyclic structure between the N-terminal acetyl carbon and the sulfur in the α-chain, has the same mass to charge ratio as that from H-abstraction at the α-carbon, followed by γ-cleavage (pathway III in **Scheme 3.5**), which makes us hard to measure the contribution portion of the direct radical substitution mechanism. However, it is clear that the contribution of the direct radical substitution, which forms an intact dication when substituted to the sulfur atom in the β-chain, is not significant for S−S bond cleavage having no explanation for the base peak at m/z 775 in FRIPS of **3HD**. It can only be understood by pathway III in **Scheme 3.5**.

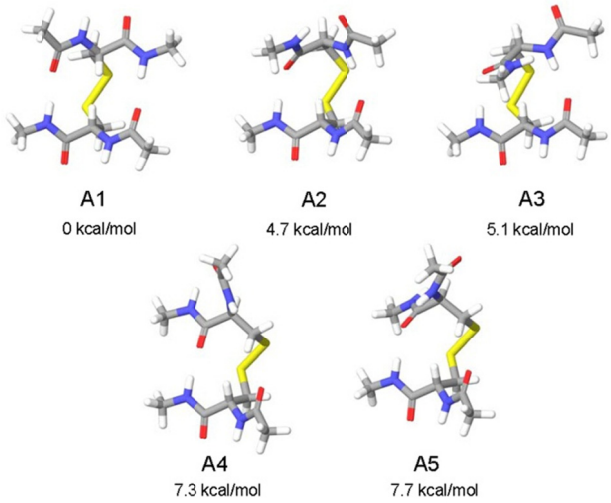
Based on the facts that the relative abundances of disulfide bond cleavage products show the similar distributions among FRIPS of **3HH**, **3DH**, and **3HD**, it is clear that pathway III (H-abstraction at the α-carbon, followed by γ-cleavage) is the dominant process for S−S bond cleavage, with a minor contribution from pathway II (direct radical substitution). The pathway I (H-abstraction at the β-carbon, followed by β-cleavage) may also play a role but only in a very minor way; the small product at m/z 774 in FRIPS of **3HD** is only explained by deuterium

abstraction (**Figure 3.7c**). For C−S bond cleavage, H-abstraction at the α-carbon, followed by β-cleavage may occur.

3.3.6. Quantum Chemical Computations

To investigate the energetics of the observed disulfide cleavage processes in collisionally activated acetyl radical cations, we use *N,N′*-diacetyl-cystine-*N*-methylamide (**Scheme 3.3**), and the untethered *N*-methylacetamide radical ($\bullet CH_2{-}CONH{-}CH_3$) as a model system. Several low-energy conformers of this model system are shown in **Figure 3.8**. The most stable conformer, **A1** is the all-trans form for amide bonds and hydrogen bonds are formed between amide oxygens and *N*-hydrogens in each chain. Other conformers have different dihedral angles from all-trans conformations between amide bonds, which distort hydrogen bonds between amide oxygens and *N*-hydrogens, yielding higher-energy conformers. Thus it is believed that the relative enthalpy difference in each conformer mainly results from disruption of interchain hydrogen bonds between amide *N*-hydrogens and carbonyl groups. Due to conformational diversity in the model system, we limit our consideration of reaction energetics to the lowest-energy structure in each reaction process without considering other possible conformers.

Figure 3.8 Low-energy conformers of model compounds



Low-energy conformers of *N,N'*-diacetyl-cystine-*N*-methylamide in relative enthalpy at 298.15K and 1 atm. Geometry optimization and thermochemical calculation were performed using B3LYP/6-311++G(d,p) level of theory and single point energy refinement was performed using B3LYP/6-311++G(3df,3pd) level of theory (in short, B3LYP/6-311++G(3df,3pd)// B3LYP/6-311++G(d,p)).

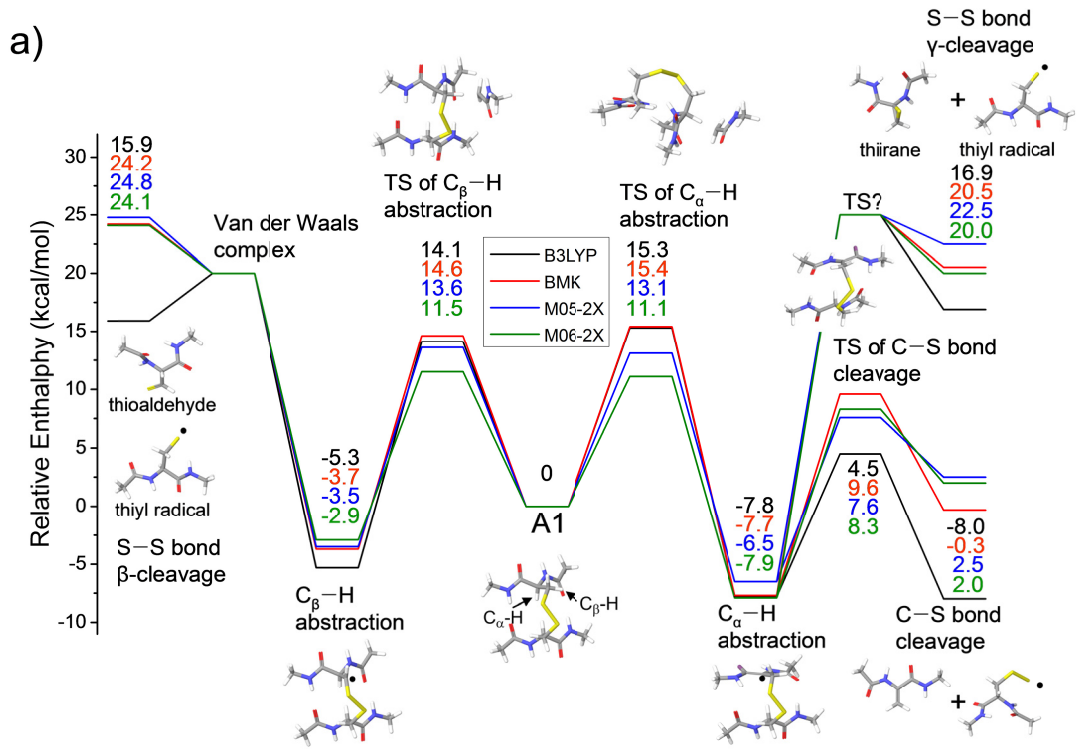
We first investigate C−S and S−S bond cleavages via abstraction of hydrogen atoms from α- and β-carbons, followed by β-[15,175] and γ-cleavages, respectively. Relative enthalpy changes associated with each reaction channel are shown in **Figure 3.9**.

For both C−S and S−S bond cleavage reactions, the enthalpy changes predicted by B3LYP systematically deviate from the results estimated by other functionals by ~8-10 kcal/mol (**Figure 3.9a**). This systematic deviation by B3LYP in the energetics of organic radical reactions has been reported previously.[176,186,187] The better performances of BMK and M05/06-2X functionals have been demonstrated in comparison with G3(MP2)-RAD results.[176,186,187] Therefore, we will discuss the energetics from the other three functionals, which are all in reasonable agreement.

As seen in **Figure 3.9**, H-abstraction at the β-carbon is exothermic but is a slightly less favored reaction (~4 kcal/mol) than H-abstraction at the α-carbon. Barriers for H-abstraction at each carbon are quite similar (~11-15 kcal/mol). Following β-cleavage reaction for C−S bond cleavage is ~7-10 kcal/mol endothermic, yielding acetyl-*N*-methyl dehydroalanine and acetyl-*N*-methyl cysteinyl radical and its barrier is ~14-17 kcal/mol. The overall enthalpy change for C−S bond cleavage is only ~0-2 kcal/mol endothermic. H-abstraction at the β-carbon, followed by S−S bond cleavage is endothermic by ~22 kcal/mol than that of H-abstraction at the α-carbon and subsequent C−S bond cleavage. In the S−S bond cleavage pathway via H-abstraction at the α-carbon, followed by γ-cleavage, the overall enthalpy change is ~2-4 kcal/mol favored over H-abstraction at the β-carbon, followed by S−S bond cleavage. In this regard, it is shown that the energetic of thiirane formation is more favored than that of thioaldehyde. During the step of S−S bond cleavage via H-abstraction at the β-carbon, followed by β-cleavage, no barrier was found, forming a Van der Waals complex between thioaldehyde and thyil radical. For the dissociation of a Van der Waals complex, a certain barrier needs to be overcome by breaking two hydrogen bonds between amide bonds.

Figure 3.9 Reaction energetics of disulfide bond cleavages

a) Reaction energetics for S−S bond cleavage (left side) and C−S bond cleavage (right side) of *N,N′*-diacetyl-cystine-*N*-methylamide via hydrogen abstraction from α- and β-carbons, followed by β- and γ-cleavages showing relative enthalpies in kcal/mol. Geometry optimization and thermochemical calculation (298.15K and 1 atm) were performed using B3LYP/6-311++G(d,p) level of theory and single point energy refinement was performed using B3LYP (black), BMK (red), M05-2X (blue), and M06-2X (green) density functionals with the 6-311++G(3df,3pd) basis set, respectively. Some barrier heights are not known. *N*-methylacetamide radical ($\bullet CH_2{-}CONH{-}CH_3$) and *N*-methylacetamide are omitted in molecular structure drawings except for transition states of the α- and β-hydrogen abstraction but their enthalpies are included in the relative enthalpy diagram. b) Schematic drawing of reaction mechanisms for S−S bond cleavage (down side) and C−S bond cleavage (upper side) of *N,N′*-diacetyl-cystine-*N*-methylamide via hydrogen abstraction from α- and β-carbons, followed by β- and γ-cleavages.


a)
Relative Enthalpy (kcal/mol)
30
25
20
15
10
5
0
-5
-10
15.9
24.2
24.8
24.1
Van der Waals complex
-5.3
-3.7
-3.5
-2.9
TS of Cβ−H abstraction
14.1
14.6
13.6
11.5
B3LYP
BMK
M05-2X
M06-2X
0
A1
TS of Cα−H abstraction
15.3
15.4
13.1
11.1
-7.8
-7.7
-6.5
-7.9
TS?
16.9
20.5
22.5
20.0
TS of C−S bond cleavage
4.5
9.6
7.6
8.3
-8.0
-0.3
2.5
2.0
S−S bond γ-cleavage
thiirane
thiyl radical
thioaldehyde
thiyl radical
S−S bond β-cleavage
Cβ−H abstraction
Cα-H
Cβ-H
Cα−H abstraction
C−S bond cleavage

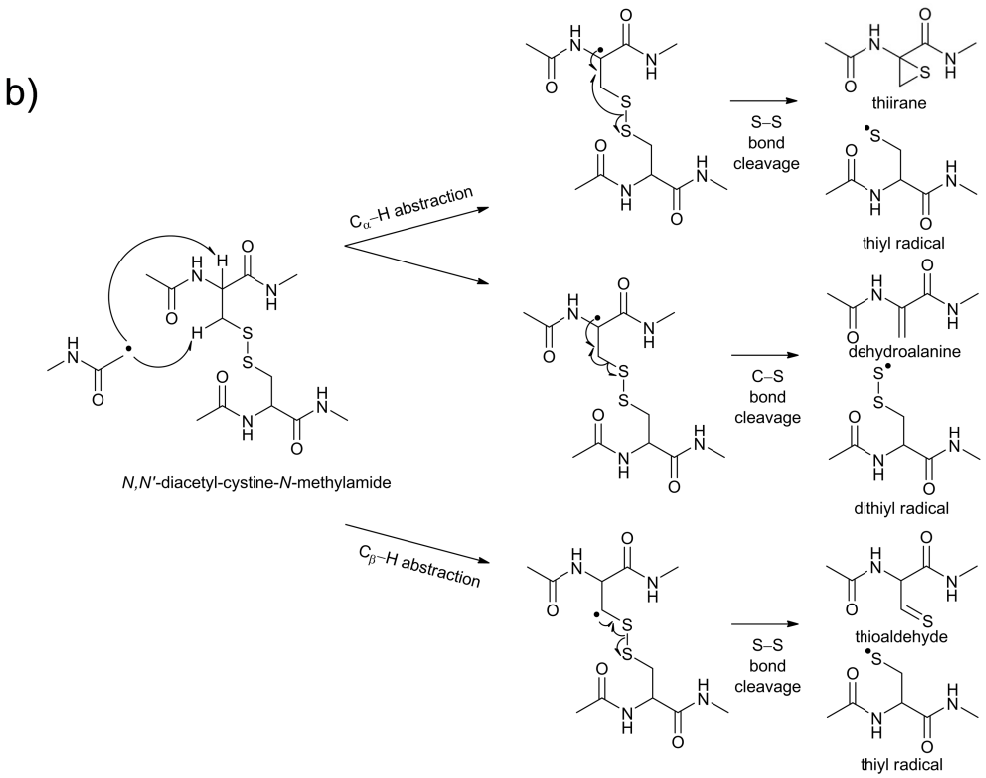
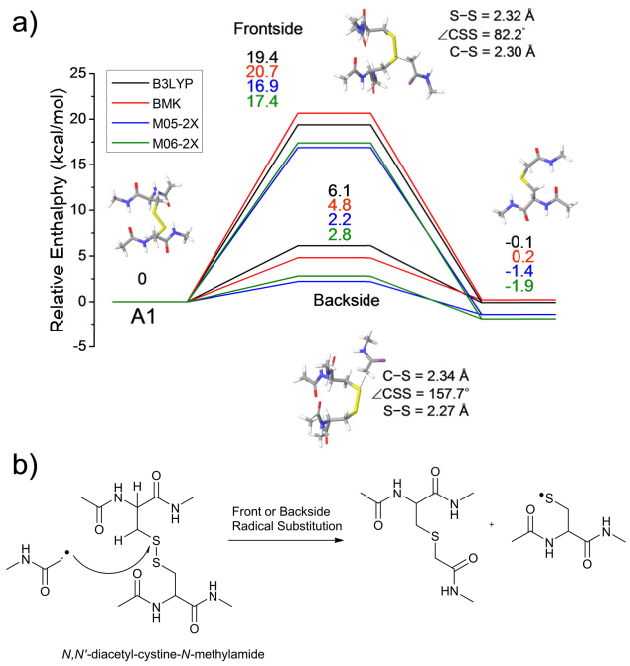

b)
N,N′-diacetyl-cystine-N-methylamide
Cα–H abstraction
Cβ–H abstraction
S–S bond cleavage
C–S bond cleavage
S–S bond cleavage
thiirane
thiyl radical
dehydroalanine
dthiyl radical
thioaldehyde
thiyl radical

Figure 3.10 Reaction energetics of S−S bond cleavage by direct radical substitution



a) Reaction energetics for S−S bond cleavage of *N,N′*-diacetyl-cystine-*N*-methylamide by direct radical substitution via front- or backside, showing relative enthalpies in kcal/mol. Geometry optimization and thermochemical calculation (298.15K and 1 atm) were performed using B3LYP/6-311++G(d,p) level of theory and single point energy refinement was performed using B3LYP (black), BMK (red), M05-2X (blue), and M06-2X (green) density functionals with the 6-311++G(3df,3pd) basis set, respectively. b) Schematic drawing of reaction mechanisms for direct radical substitution.

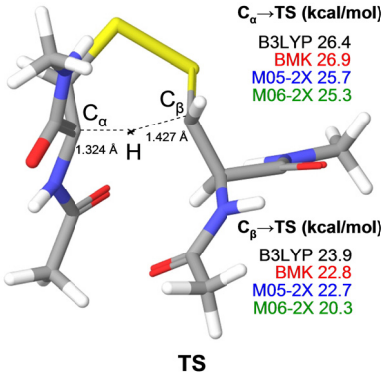
In addition, the energetics of the direct acetyl radical substitution to the sulfur atom, followed by S−S bond cleavage is considered (**Figure 3.10**). Formerly, the methyl radical substitution to dimetyldifulfide was studied using DFT.[181] Two distinctive transition states were reported via front- and backside attack of the methyl radical and this process turns out to be concerted. The good orbital overlap via the backside attack between the σ* orbitral of S−S bond and the singly occupied molecular orbital (SOMO) of the methyl radical lowers the barrier for S−S bond cleavage. For our case, the formation of the hypervalent sulfur radical by substitution of the acetyl radical group is investigated whether the process is concerted or consecutive. However, intermediate structures having no imaginary vibrational frequency (*i.e.*, non-transient state structures) for the hypervalent sulfur radical were not found. Therefore, concomitant addition and dissociation of a S−S bond by addition of the acetyl radical are assumed (a concerted reaction pathway) and their energetics are evaluated. Enthalpy changes for S−S bond cleavage via direct addition of the acetyl radical group are estimated to be −0.1, 0.2, −1.4 and −1.9 kcal/mol by the B3LYP, BMK, M05-2X and M06-2X/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) levels of theories, respectively. This process is a lot more favored (ΔH = ~0 kcal/mol) than what is observed in the β-hydrogen abstraction initiated process (ΔH = ~24 kcal/mol). Also, the barrier for the backside attack (~2-5 kcal/mol) is substantially lower than all other reaction pathways. However, it should be noted that the alignment of reactant residues, the acetyl radical and the disulfide bond, is of particular importance for the radical substitution reaction. The reaction barrier is a very sensitive on the incident angle of the incoming acetyl radical (frontside versus backside, **Figure 3.10a**). The conformational space of the gaseous peptide ion may also not be highly populated for the successful orbital overlap between the σ* orbitral of S−S bond and SOMO of the acetyl radical lowers the barrier for S−S bond cleavage. Thus the very small ranges of incident angles may play an important role for the reaction in actual model peptides. In this

regard, it is concluded that the contribution of the direct radical substitution pathway is estimated to be rather low for S−S bond cleavage.

The H-atom migration between α- and β-carbons via the six-membered ring transition state is also investigated. H-atom abstraction from sterically more accessible β-carbons would favorably occur by the N-terminal acetyl radical group. If the barrier for the H-atom migration between α- and β-carbons is reasonably lower, the resulting radicals can initiate various C−S or S−S bond cleavages (**Figure 3.10**).

Figure 3.11 depicts the structure of the transition state for radical migration in which the transferred H-atom is approximately centered between the α- and β-carbons. The calculated reaction barrier from the β-carbon is rather high, ~20-23 kcal/mol but it is still 4-7 kcal/mol lower than the overall endothermicity of S−S bond cleavage of the β-carbon radical. Therefore, it is energetically reasonable for the H-atom migration to occur from the β-carbon to the α-carbon of a disulfide bond and the resulting α-carbon radical can undergo further C−S or S−S bond cleavages.

Figure 3.11 Transition state structure of hydrogen migration between α- and β-carbons



The optimized transition state structure of hydrogen migration between α- and β-carbons and reaction barriers from each side showing relative enthalpies in kcal/mol. The geometry of the transition state was optimized using B3LYP/6-311++G(d,p) level of theory and was identified by the existence of single imaginary frequency from vibrational mode analysis. Single point energy refinement was performed by B3LYP, BMK, M05-2X, and M06-2X/6-311++G(3df,3pd) levels of theory, respectively. The presented reaction barriers include enthalpy corrections calculated by B3LYP/6-311++G(d,p) level of theory at 298.15K and 1 atm.

3.3.7. Disulfide Bond Cleavage by FRIPS versus ECD

As we discussed in the previous sections, the reaction mechanisms of FRIPS and ECD that lead to disulfide bond cleavage are different. However, the observed products from disulfide bond cleavage using FRIPS and ECD methods share a common feature by having the prominent products from S−S bond cleavage. On the other hand, ECD more prefers to cleave S−S bonds while FRIPS still fragments C−S bonds. The difference on relative abundances of the products from the C−S bond cleavage in each method could be partially attributed to the time scale of measurement. Considering the time scale of mass spectrometers used in this study (ion traps versus FT-ICR) in ion detection, the product distribution can be dominated by the lowest-energy pathway in longer time scale with a FT-ICR mass spectrometer. As a result, favored S−S bond cleavage fragments are prominent in ECD spectra whereas competitive C−S bond cleavage processes are observed in FRIPS spectra.

3.3.8. Reactivity of Disulfide Bond with Radicals

Both radical-driven peptide fragmentation methods, FRIPS and ECD produce abundant products from disulfide bond cleavage. It is also independent on its connectivity (*i.e.*, intra- or intermolecular linkages). The present experimental and theoretical results provide critical insights relating to selective disulfide bond cleavages effected using different experimental methodologies. Free-radical processes that cleave disulfide bonds in proteins are ubiquitous in nature,[183,188] resulting from a wide range of redox reactions *in vivo* in addition to the environmental hazards of ionizing and UV radiation. We believe that mechanistic aspects of disulfide bond cleavage in nature can be better understood by our fundamental biomimetic studies of model systems using well-constrained hydrogen-rich and hydrogen-deficient conditions.

3.4.Conclusion

The mechanism of disulfide bond cleavage by a regiospecific acetyl radical (FRIPS) is investigated and compared to that of a low-energy electron (ECD). The second generation TEMPO-based FRIPS reagent is derivatized to model peptides to generate the regiospecific acetyl radical group via collisional activation that predominantly leads to loss of TEMPO radical. Following collisional activation of the acetyl radical attached model peptides yields highly selective C−S and S−S bond cleavages in both inter- and intradisulfide linkages. Additional collisional activations of fragments from C−S and S−S bond cleavages generate sequence information of disulfide-linked peptide chains, allowing us to assign disulfide bond linkages between cysteine residues. Both C−S and S−S bond cleavage processes are mainly initiated by H-abstraction at the α-carbons. Following β- and γ-cleavages lead to C−S and S−S bond cleavages. Density functional theory calculations reveal the energetics of the C−S and S−S bond cleavage processes. Other possibilities (direct radical substitution and hydrogen shift between the α- and β-carbons) for explanation on the favored S−S bond cleavage in FRIPS are also suggested and discussed. Observed gas-phase fragmentation pathways can provide insights on biological processes associated with disulfide bond cleavages by reactive radical species and redox stresses.

3.5.Acknowledgment

This work was supported by the National Science Foundation through grant CHE-0416381 and the Beckman Institute at California Institute of Technology. Electron capture dissociation was performed using the LTQ-FT in the Proteome Exploration Laboratory, the Beckman Institute.

Computational resources were kindly provided by the Materials and Process Simulation Center at California Institute of Technology. C.H.S. acknowledges a fellowship from the Kwanjeong Educational Foundation.

Supporting Information Available:

1. Peptide fragmentation nomenclature used in this study for c- and z-type ions produced by ECD and ETD versus a- and x-type ions produced by FRIPS.

2. Details on the synthesis of the TEMPO-based FRIPS reagent and peptide conjugation.

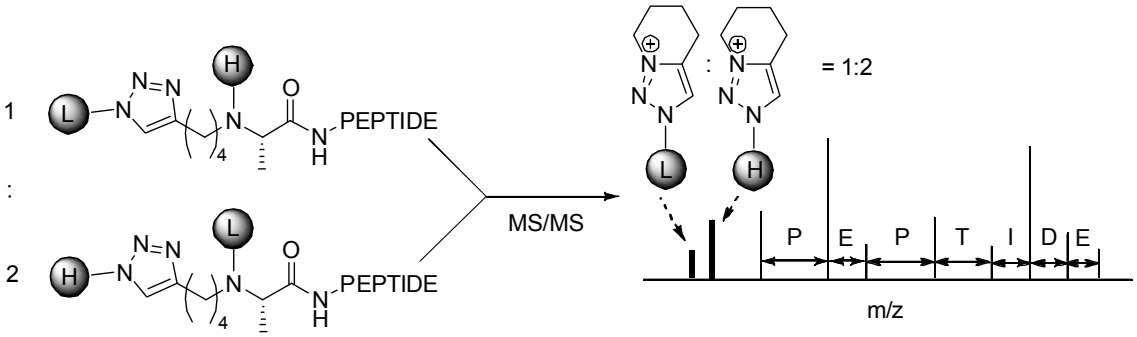
3. The precursor ion structure of TEMPO-CFIR/NCPR.

4. Designer Reagents for Mass Spectrometry-Based Proteomics: Click Chemistry Facilitates Synthesis of Amine-reactive Multiplexed Isobaric Tags for Protein Quantification

Abstract Quantitative mass spectrometry (MS)-based proteomics has been employed to investigate various biological processes by measuring the relative and absolute expression levels of proteins in cells. These studies are enabled by tandem mass spectrometry (MS/MS)-based isobaric tags, made popular by their capability for multiple sample comparison in a single MS experiment. Here, we develop novel isobaric tags for protein quantification, referred to as Caltech Isobaric Tags (CITs), which offer several advantages over other isobaric tags (*e.g.*, iTRAQ and TMT). These include 1) the formation of reporter ions based on a newly discovered low-energy fragmentation pathway, a nucleophilic displacement reaction by the 1,2,3-triazole ring, 2) an unlimited number of isobaric combinations of CIT reagents in principle, 3) an easily tunable reporter ion mass to access clear windows of m/z values not overlapping with peptide MS/MS fragments, and to avoid the low mass–cut off problems inherent in ion trap mass spectrometers, and 4) synthetic methodology that permits preparation of CITs with minimal expense and effort. These advantages are demonstrated by preparing duplex CIT reagents whose reporter ions appear at *m/z* 164 and 169. CIT reagents are applied to label a model tryptic peptide, protein mixture digests (bovine serum albumin, α and β caseins, ovalbumin, and lysozyme; enolase, aldolase, hemoglobin, creatine kinase, and alcohol dehydrogenase), and Cul1 protein complexes affinity-purified from HEK 293 cells with various ratios. The resulting CIT-labeled peptides are analyzed by either pulsed Q-dissociation (PQD) or higher energy collision dissociation (HCD) in LTQ-Orbitrap mass spectrometers. Heavy to light ratios of the CIT reporter ions provide excellent

quantitative results, indicating relative responses over a two-orders-of-magnitude dynamic range, in agreement with initial mixing ratios.

Graphical Abstract



4.1.Introduction

Recent achievements in mass spectrometry (MS)-based proteomics have provided essential methodologies for a deeper understanding of protein expressions in cells.[37,189] MS-based proteomics allows high-throughput identification and quantification of proteins of interest. Currently, state-of-the-art liquid chromatography (LC)-MS instruments can analyze the whole-cell yeast lysate within a day, identifying ~5000 proteins.[190,191] Quantitative approaches in MS-based proteomics are aiming to investigate the relative and absolute expression levels of proteins in cells.[16,17] By employing those approaches, various biological processes can be monitored by tracking changes in protein expression.[192]

The simplest method for quantitative mass spectrometric measurement is the label-free analysis. After successive runs of samples of interest under the same instrument conditions, protein abundances are determined by either integrating ion chromatograms or spectral counting of MS signals.[193] Yet, current label-free quantification approaches require highly consistent analyses, which are mostly hampered by fluctuations in ionization efficiencies and difficulties in subsequent data processing.[193,194]

In another label-free approach, selected reaction monitoring (SRM), and its extension plural multiple reaction monitoring (MRM),[195] examines the transitions (*i.e.*, one or more targeted fragment ions from the precursor ions) by scanning specific mass regions using triple quadruple mass spectrometers. This permits highly sensitive identification and concomitant quantification of peptides.[196-198] SRM, however, requires pre-knowledge of fragmentation behaviors of analytes, for which the transitions should be determined through tedious assays prior to actual SRM analyses. In addition, the high cost for the preparation of required synthetic peptides may limit its

wide application in shotgun proteomics involving complex mixtures. SRM also suffers from problems associated with fluctuation in ionization efficiency.

To address problems associated with label-free protein quantification, stable isotopes are incorporated into protein or peptide samples to be used as internal (or mutual) standards. A conceptual breakthrough was achieved by Aebersold and coworkers by introduction of isotope-coded tags that can be selectively labeled to peptide digests.[199] In this approach, cysteine-containing peptides from different sources are tagged by light or heavy isotope-coded affinity tag (ICAT) reagents, and enriched from the complex mixture utilizing an attached biotin affinity tag. Because these tagged peptides share the same physico-chemical properties, they are not differentiated by ionization and chromatography steps. Therefore, a simple comparison of MS signal intensities between light and heavy isotope-coded peptides directly yields the relative protein expression levels.

Another quantification approach takes advantage of *in vivo* incorporation of isotope labels. In SILAC (Stable Isotope Labeling with Amino acids in Cell culture),[200] two cell populations are grown with identical culture media except for stable isotope-labeled amino acids (*e.g.*, ^{13}C and/or ^{15}N labeled lysine and/or arginine). Resulting heavy isotope-coded cell populations behave equally as their light isotope-coded controls. After applying a perturbation to one of the light or heavy cell populations, the two samples are combined for MS analysis. Direct comparison of the peptide signals from the light- and heavy-labeled cell populations yields the relative protein expression levels. Unlike ICAT, which quantifies partial proteome by labeling cysteine-containing peptides, SILAC is capable of quantifying the global proteome, because all tryptic peptides are labeled with isotope-coded lysine and/or arginine.

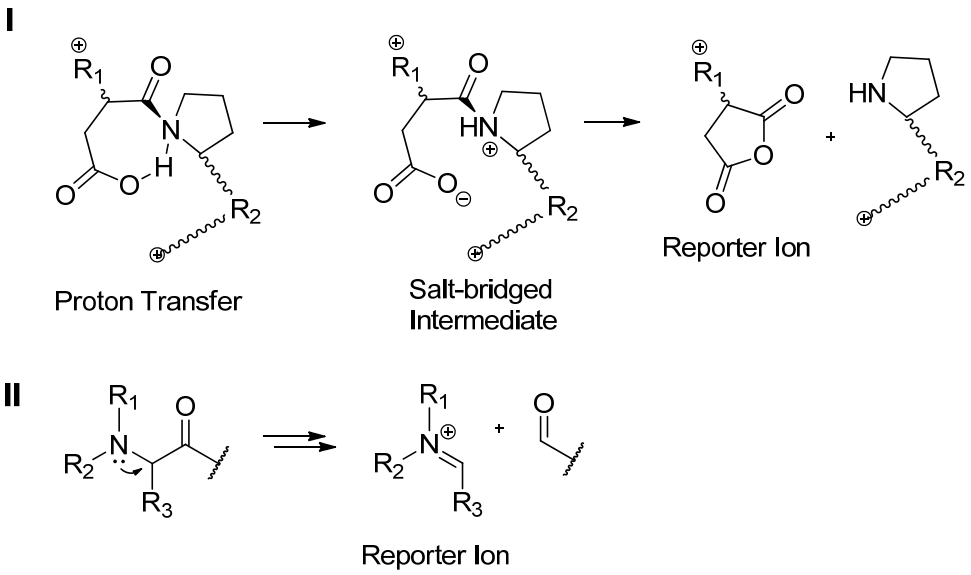
The LC-MS ion signals from isotope-coded peptides in both ICAT and SILAC methodologies are divided into two signals for each labeled peptide, causing an increase in the complexity of MS

scans and a reduction in the sensitivity of subsequent sequence analysis by MS/MS. As an alternative chemical-labeling method, isobaric tags such as Tandem Mass Tag (TMT)[18,201] and isobaric Tags for Relative and Absolute Quantification (iTRAQ)[19,202] were developed. An isobaric tag is composed of three parts: the reporter ion group, the mass balance group and the reactive group.[203] The combined mass of the reporter ion group and the mass balance group is isobaric despite each part having different masses. When the identical peptides from different biological experiments are labeled by isobaric tags, they possess the same masses, appearing as a single peak. Mass differentiated reporter ions are produced and detected in MS/MS scans, in which their relative intensities reflect the initial amounts of each peptide from the original sources. All backbone fragment ions are also isobaric, allowing simultaneous peptide sequencing and quantification. One advantage of isobaric tags over ICAT- and SILAC-based methodologies, especially when quantifying more than two system states, is that the combined signals from multiple biological samples reduce the complexity of the MS scans and increase the sensitivity for the following MS/MS analyses. Preparing each part of the tag with various isotopomers in principle allows for the facile construction of multiplexed reagents capable of quantifying multiple samples in a single MS analysis. The primary amines that are usually the target functional groups for isobaric tag labeling exist in virtually all peptides (*i.e.*, the N-terminal amine and ε-amine of lysine side-chain), enabling researchers to investigate the whole proteome. The isobaric tags have been employed for quantification of tissue samples and even isotopically non-transferable subjects such as human.[204]

In spite of the improvements achieved by isobaric tags, their applications have been limited by the cost of commercially available reagents.[205] In addition, the number of the current multiplexed isobaric tags (*e.g.*, iTRAQ) is limited to a maximum of eight[202] due to their inherent designs and difficulties in the synthesis of various isotope-coded functional groups. The low mass

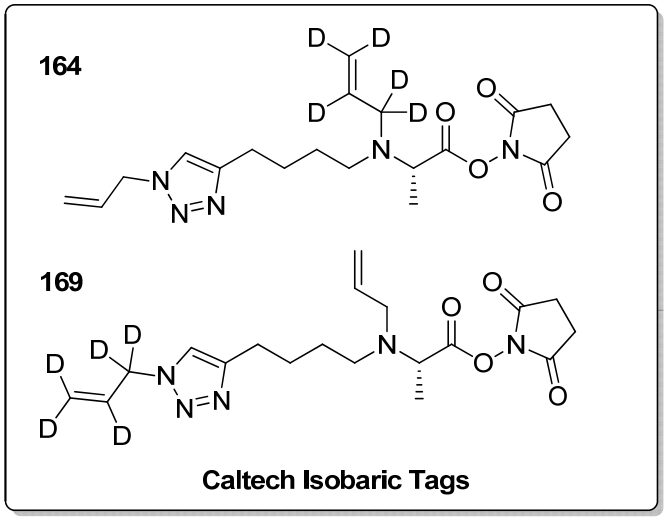
cut–off in resonance type ion trap mass spectrometers, one of the most popular proteomics platforms, also hinders the simultaneous monitoring of reporter ions and peptide sequence ions.

The design of new reagents is hindered by the fact that very few low-energy fragmentation pathways suitable for the production of the reporter ions are known in peptide tandem MS. One of the widely known fragmentation pathways, the preferential cleavage of the aspartic acid–proline peptide bond via a salt-bridged intermediate,[10,11] was applied in the first TMT report (**Scheme 4.1**, I).[18] In iTRAQ, the formation of the reporter ions proceeds through the facile *N*-methylpiperazine-acetyl bond−mediated cleavage process (**Scheme 4.1**, II). Most isobaric tags that have been proposed to provide cheaper synthetic routes are still based on similar tertiary amine-branched methylene-amide bonds.[206-208]



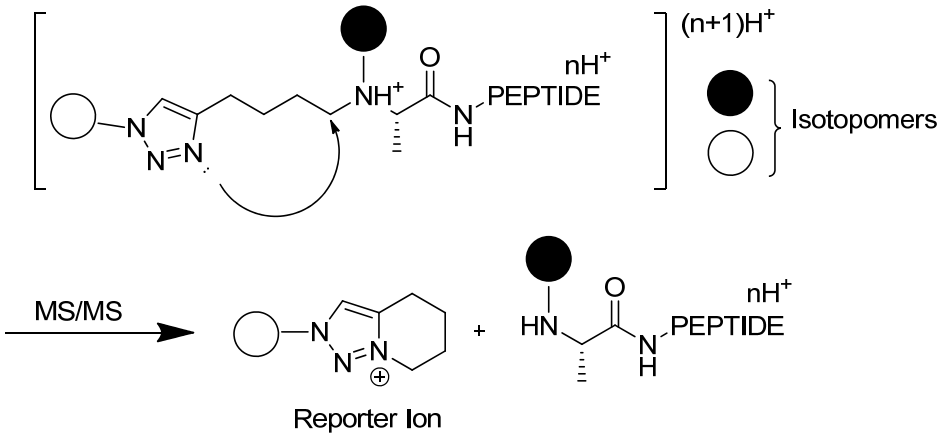
Scheme 4.1

Here, we report a novel isobaric reagent (**Scheme 4.2**), referred to as Caltech Isobaric Tag (CIT), which affords a significant advance in MS-based quantification methodology.



Scheme 4.2

CIT has many advantages over current isobaric labeling reagents. First, CIT is easy to produce and the possible number of multiplex isobaric tags, in theory, is unlimited. Furthermore, energetics for the reporter ion and peptide sequence ion formations are balanced, guaranteeing simultaneous quantification and sequencing of target peptides. CIT is inspired by the observation of a highly selective gas-phase fragmentation triggered by a nucleophilic substitution of N3 nitrogen in the 1,2,3-triazole ring (**Scheme 4.3**). The design of a duplex CIT reagent is described, and the applicability of the reagent is validated in model systems using various mass spectrometers.



Scheme 4.3

4.2.Experimental Sections

4.2.1. Materials

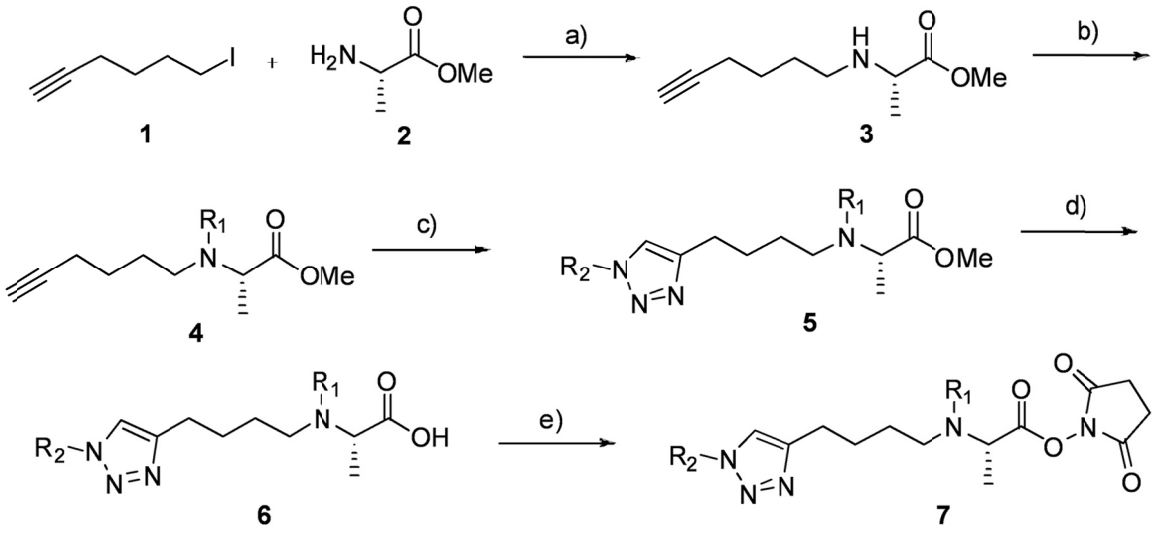
4.2.1.1.Chemicals

Allyl bromide-d_5 was purchased from C/D/N isotopes Inc. (Quebec, Canada). MagneHis Ni-Particles and sequencing grade trypsin was purchased from Promega (Madison, WI). The model peptide, VIP (residue 1-12), HSDAVFTDNYTR, was acquired from Anaspec (Fremont, CA). High Capacity Neutravidin Agarose Resin, n-dodecyl-β-D-maltoside, and SuperSignal West Dura Extended Duration chemiluminescent substrate were from Thermo Scientific (Rockford, IL). Lysyl endopeptidase (LysC) was from Wako Chemicals USA (Richmond, VA). Cell culture reagents, Flip-In T-REx 293 cells, plasmids and monoclonal antibodies for Cul1 and Cand1 were from Invitrogen (Carlsbad, CA). Plasmid DNA containing the human Cul1 sequence was purchased from Open Biosystems (Huntsville, AL). MLN4924 was a generous gift from Millennium: The Takeda Oncology Company (Cambridge, MA). All other general chemicals for buffers were purchased from Fisher Scientific (Hampton, NH), VWR International (West Chester, PA) and Sigma-Aldrich (St. Louis, MO) and used as received without further purification.

4.2.2. Synthesis of CITs

Synthetic schemes are summarized in **Figure 4.1**.

Figure 4.1 Synthesis of CIT reagents



Synthesis of CIT reagents. a) THF, K_2CO_3, TEAI, reflux, 18 h, 56%. b) THF, K_2CO_3, TEABr, reflux, 18 h, R_1=Allyl-d_0-bromide, 56%, R_1=Ally-d_5-bromide, 67%. c) 0.4 eq Na ascorbate, 0.1 eq CuSO4, 0.01 eq TBTA, DMSO/H_2O, RT, 4 h, R_2=Allyl-d_5-azide, 72% (heavy tag), R_2=Allyl-d_0-azide, 69% (light tag). d) 2M KOH, THF, RT, overnight, *quantitative* (heavy tag), 97% (light tag). e) TFA-NHS, DMF, overnight, 24% (heavy tag), 23% (light tag). THF = tetrahydrofuran, TEAI = tetraethylammonium iodide, TEABr = tetraethylammonium bromide, TBTA = tris[(1-*t*-butyl-1*H*-1,2,3-triazol-4-yl)methyl]amine, DMSO = dimethyl sulfoxide, TFA-NHS = trifluoroacetic *N*-hydroxysuccinimide ester, DMF = *N,N'*-dimethyl formamide.

4.2.2.1. 6-Iodo-hex-1-yne

In a flame-baked, one neck 250 mL round bottom flask, 40 mmol of 6-chloro-1-hexyne (~5mL) was dissolved in 25 mL of acetone with 80 mmol of sodium iodide (~12 g) and refluxed for 2 days with thin layer chromatography (TLC) check. After filteration, the mixture was properly diluted by diethylether. The organic layer was washed by water, $Na_2S_2O_4$ and brine and dried over $MgSO_4$. The resulting solution was concentrated by rotavap with caution (the product is slightly volatile). The desired product, 6-iodo-1-hexyne (5.58 g, 26.4 mmol) was acquired as a brownish oil. Yield: 66%

4.2.2.2. N-(5-hexynyl) L-alanine methyl ester

In a flame-baked, two neck 100 mL round bottom flask, 20 mmol of L-alanine methyl ester hydrochloride (2.8 g), 40 mmol of K_2CO_3 (5.53 g), and 20 mmol of tetrabutylammonium iodide (TBAI, 7.39 g) were charged under the stream of dry N_2 gas. 35 mL of tetrahydrofuran (THF) was slowly added and stirred for 15 min at room temperature. 2.6 mL of 6-iodo-hex-1-yne was added dropwise while the mixture was stirred. The reaction mixture was refluxed at ~70°C for 15-18 h with TLC check. After consumption of the starting material, the mixture was cooled to room temperature, diluted by diethylether, and filtered. The filtrate was further diluted by diethylether and filtered again. The solvent was removed by rotavap and the crude product was purified by flash chromatography on silica gel (1:2 = Hexane/EtOAc, 1% triethylamine) to give *N*-(5-hexynyl) L-alanine methyl ester (2.044 g, 11.2 mmol) as a yellow oil. Yield: 56%. R_f = 0.27 (1:1 = Hexane/EtOAc); ESI-MS $[M+H]^+$ = *m/z* 184.1; ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.70 (s, 3H), 3.32 (q, *J* = 7.0 Hz, 1H), 2.53 (m, 2H), 2.18 (m, 2H), 1.92 (t, *J* = 2.6 Hz, 1H), 1.56 (m, 5H),

1.27 (d, J = 7.0 Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 176.29, 84.21, 68.41, 56.61, 51.74, 47.41, 29.20, 26.01, 19.11, 18.23.

4.2.2.3. N,N-(5-hexynyl)(allyl-d_0) L-alanine methyl ester and N,N-(5-hexynyl)(allyl-d_5) L-alanine methyl ester

In a flame-baked, two neck 50 mL round bottom flask, 10 mL THF was charged under the stream of dry N_2 gas. 6 mmol of K_2CO_3 (0.83 g), and 6 mmol of tetrabutylammonium bromide (TBAB, 1.93 g) were slowly added and stirred for 15 min in room temperature. 3 mmol of *N*-(5-hexynyl) L-alanine methyl ester (0.55 g), and 4.5 mmol of allyl bromide-d_0 (0.544g, 0.38 mL) was slowly added dropwise using the syringe while the mixture was stirred. The reaction mixture was heated at ~55 °C and reacted for 8 h with TLC check. After consumption of the starting material, the mixture was cooled to room temperature, diluted by diethylether, filtered and repeated filteration twice to remove the remaining TBAB completely. The solvent was removed by rotavap and the crude product was purified by flash chromatography on silica gel (20:1 = Hexane/EtOAc) to give *N,N*-(5-hexynyl)(allyl-d_0) L-alanine methyl ester (0.378 g, 1.69 mmol) as a transparent oil. Yield: 56%. R_f = 0.4 (5:1 = Hexane/EtOAc); ESI-MS $[M+H]^+$ = *m/z* 224.2; ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.79 (m, 1H), 5.12 (m, 2H), 3.67 (s, 3H), 3.54 (q, J = 7.3 Hz, 1H), 3.18 (m, 2H), 2.53 (m, 2H), 2.18 (m, 2H), 1.93 (t, J = 2.7 Hz, 1H), 1.52 (m, 4H), 1.24 (d, J = 7.3 Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 174.45, 136.86, 116.78, 84.47, 68.24, 57.79, 54.20, 51.21, 49.85, 27.40, 26.07, 18.26, 14.88.

0.457g of *N,N*-(5-hexynyl)(allyl-d_5) L-alanine methyl ester (2.0 mmol) was obtained by the same procedure described above using 3.55 mmol of allyl bromide-d_5 (0.448 g). Yield: 67%. R_f = 0.4 (5:1 = Hexane/EtOAc); ESI-MS $[M+H]^+$ = *m/z* 229.3; ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.68 (s, 3H), 3.54 (q, J = 7.08 Hz, 1H), 2.53 (m, 2H), 2.18 (m, 2H), 1.93 (t, J = 2.68 Hz, 1H), 1.52 (m,

4H), 1.25 (d, J = 7.08 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 174.47, 136.16 (t, J = 23.48 Hz), 116.25 (quintet, J = 23.47 Hz), 84.47, 68.27, 57.76, 53.28 (quintet, J = 18.41 Hz), 51.22, 49.79, 27.39, 26.07, 18.26, 14.90.

4.2.2.4. Allyl-d_0 azide and allyl-d_5 azide

0.5 M NaN_3 in DMSO was prepared as described in the literature by stirring the mixture at room temperature overnight.[209] 1.1 eq of 0.5 M NaN_3 solution was mixed with the appropriate amount of allyl bromide (d_0/d_5) and stirred overnight. TLC was monitored for the complete consumption of the starting material and unwanted dimerization of allyl azides was not observed in this condition. The resulting mixtures were used for the next steps without further purification or analysis.

4.2.2.5. N,N-(4-(1-allyl-d_5-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine methyl ester and N,N-(4-(1-allyl-d_0-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine methyl ester

The literature procedure was followed with some modifications.[210] To *in situ* prepared allyl-d_5 azide solution (~1.2 eq), 1.7 mmol of *N,N*-(5-hexynyl)(allyl-d_0) L-alanine methyl ester (0.378 g), 0.17 mmol of $CuSO_4 \cdot 5H_2O$ (42.5 mg, 0.1 eq), 0.68 mmol of sodium ascorbate (134.7 mg, 0.4 eq), 0.017 mmol of tris[(1-*t*-butyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA) (7.3 mg, 0.01 eq), and additional 2 mL of DMSO were added and stirred for 2 h at room temperature. 2.4 mL of water was then added and stirred for additional 2 h with monitoring TLC. After the complete consumption of the starting material, 4 mL of 1M NH_4OH was added to remove residual CuN_3 and $(Cu)_2N_3$. The mixture was diluted by additional water and ethyl acetate. The aqueous layer turned to be blue by the coordination of ammonia to copper ions. The organic layer was separated, and further extracted by ethyl acetate three times. The combined organic layer was then washed

by brine, dried over $MgSO_4$ and concentrated by rotavap. The crude product was purified by flash chromatography on silica gel (1:1 = Hexane/EtOAc, 1% triethylamine) to give *N,N*-(4-(1-allyl-d_5-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine methyl ester (0.379 g, 1.22 mmol) as a transparent oil. Yield: 72%. R_f = 0.3 (1:1 = Hexane/EtOAc); ESI-MS $[M+H]^+$ = *m/z* 312.3, CID of $[M+H]^+$ produced *m/z* 169.1 fragment; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.25 (s, 1H), 5.76 (m, 1H), 5.14 (dd, *J* = 17.1, 1.22 Hz, 1H), 5.04 (d, *J* = 10 Hz, 1H), 3.64 (s, 3H), 3.51 (q, *J* = 7.08 Hz, 1H), 3.21 (dd, *J* = 14.6, 5.6 Hz, 1H), 3.09 (dd, *J* = 14.4, 6.9 Hz, 1H), 2.68 (t, *J* = 7.6 Hz, 2H), 2.52 (m, 2H), 1.54 (m, 4H), 1.21 (d, *J* = 7.08 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 174.46, 148.43, 136.81, 130.94 (t, *J* = 24.9 Hz), 120.40, 119.21 (quintet, *J* = 24.9 Hz), 116.82, 57.85, 54.25, 51.85 (quintet, *J* = 23.9 Hz), 51.24, 50.20, 27.97, 27.06, 25.54, 14.83.

1.59 mmol of *N,N*-(5-hexynyl)(allyl-d_5) L-alanine methyl ester (0.363 g) was used for the same reaction described above to give *N,N*-(4-(1-allyl-d_0-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine methyl ester (0.341 g, 1.10 mmol) as a transparent oil. Yield: 69%. R_f = 0.3 (1:1 = Hexane/EtOAc); ESI-MS $[M+H]^+$ = *m/z* 312.3, CID of $[M+H]^+$ produced *m/z* 164.1 fragment; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.25 (s, 1H), 5.98 (m, 1H), 5.30 (dd, *J* = 10, 0.9 Hz, 1H), 5.25 (d, *J* = 17.1 Hz, 1H), 4.91 (d, *J* = 6.1 Hz, 2H), 3.65 (s, 3H), 3.51 (q, *J* = 7.1 Hz, 1H), 2.69 (t, *J* = 7.6 Hz, 2H), 2.52 (m, 2H), 1.55 (m, 4H), 1.22 (d, *J* = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 174.49, 148.45, 136.12 (t, *J* = 23 Hz), 116.29 (quintet, *J* = 22.6 Hz), 120.43, 119.78, 57.84, 53.33 (quintet, *J* = 19.3 Hz), 52.55, 51.24, 50.15, 27.98, 27.07, 25.55, 14.85.

4.2.2.6. N,N-(4-(1-allyl-d_5-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine and N,N-(4-(1-allyl-d_0-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine

1.22 mmol of *N,N*-(4-(1-allyl-d_5-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine methyl ester (0.375 g) was charged to a 10mL one neck flask with 2 mL of THF and 2 mL of 2M KOH and

stirred at room temperature for 10 h. The reaction was monitored by TLC and ESI-MS until the starting material was completely consumed. Upon completion of the reaction, THF was removed by rotavap and the aqueous layer was neutralized by ~2 mL of 2M HCl. Water was then completely removed by rotavap and the residue was reconstituted by acetonitrile (ACN). Insoluble KCl salt was filtered and ACN was removed by rotavap. The free acid of the alanine derivative, *N,N*-(4-(1-allyl-d_5-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine was recovered as a pale yellow gleasy oil. Yield: *quantitative*. ESI-MS $[M+H]^+$ = *m/z* 298.1; ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.62 (br, 1H), 7.90 (d, *J* = 2.2 Hz, 1H), 6.05 (m, 1H), 5.53 (d, *J* = 17.1 Hz, 1H), 5.45 (d, *J* = 10.5 Hz, 1H), 4.18 (q, *J* = 7.1 Hz, 1H), 3.83 (m, 2H), 3.15 (br, 2H), 2.64 (t, *J* = 7.6 Hz, 2H), 1.69 (m, 4H), 1.52 (d, *J* = 6.9 Hz, 3H).

N,N-(4-(1-allyl-d_0-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine methyl ester (0.247 g, 0.79 mmol) was used for hydrolysis by the same procedure described above and 0.230 g of *N,N*-(4-(1-allyl-d_0-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine (0.773 mmol) was obtained as a pale yellow gleasy oil. Yield: 97%. ESI-MS $[M+H]^+$ = *m/z* 298.1; ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.54 (br, 1H), 7.89 (d, *J* = 4.2 Hz, 1H), 6.03 (m, 1H), 5.25 (m, 1H), 5.16 (m, 1H), 4.98 (m, 2H), 4.18 (q, *J* = 7.1 Hz, 1H), 3.15 (br, 2H), 2.64 (t, *J* = 7.3 Hz, 2H), 1.69 (m, 4H), 1.52 (d, *J* = 7.1 Hz, 3H).

4.2.2.7.N,N-(4-(1-allyl-d_5-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine N-hydroxysuccinimide ester and N,N-(4-(1-allyl-d_0-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine N-hydroxysuccinimide ester

In a flame-baked 50 mL one neck flask, 2.75g of *N*-hydroxysuccinimide was added to 14 mL of trifluoroacetic anhydride at room temperature under the stream of dry N_2 gas and stirred for 4 h. The solvent was removed by rotavap and further eliminated by highvac overnight. The white

crystal product, trifluoroacetic *N*-hydroxysuccinimide ester (TFA-NHS) was obtained, stored in the dry desiccator, and used just before activation of free acids.

In a flame-baked 50 mL one neck flask, 87 mg of *N,N*-(4-(1-allyl-d_5-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine (0.29 mmol) and 75 mg of TFA-NHS were added to 1 mL of dry DMF, and stirred overnight at room temperature. After the complete consumption of the starting material by monitoring TLC, the reaction mixture was separated by flash chromatography on silica gel (1:1 = Hexane/EtOAc) and yielded 28 mg of *N,N*-(4-(1-allyl-d_5-1H-1,2,3-triazol-4-yl)butyl)(allyl-d_0) L-alanine *N*-hydroxysuccinimide ester (~0.7 mmol) as a yellow oil. (Note: the poor recovery yields for the final products are observed due to the retained products in silica gel through the coordination of the highly *N*-substituted residues such as the 1,2,3-triazole and the tertiary amine. Other purification methods such as crystallization would improve the overall yield.) Yield: 24%. ESI-MS (100% ACN) $[M+H]^+$ = *m/z* 395.1. The stock solution of the heavy tag (*m/z* 169 reporter ion) was prepared without further analysis by adding 20 μL dry DMSO to 1 mg of the NHS-ester product into each vials, and stored in −80 °C. Each vial contains 1mg of the reagent and used for each labeling experiment appropriately.

The same procedure was used for NHS ester activation of 52 mg of *N,N*-(4-(1-allyl-d_0-1*H*-1,2,3-triazol-4-yl)butyl)(allyl-d_5) L-alanine (0.175 mmol) and yielded 16 mg of the NHS-ester product. Yield: 23%. ESI-MS (100% ACN) $[M+H]^+$ = *m/z* 395.1. The stock solution vials of the light tag (*m/z* 164 reporter ion) were prepared as described above and stored in −80 °C.

4.2.3. Synthesis of iTRAQ-113 Reagent

4.2.3.1. Methyl 2-(4-methylpiperazin-1-yl) acetate

In a flame-baked, two neck 100 mL round bottom flask, 5 mmol of 1-methyl piperazine (0.5 g), and 6 mmol of K_2CO_3 (0.83 g) were charged under the stream of dry N_2 gas. Fifteen mL of

tetrahydrofuran (THF) was slowly added and stirred for 15 min at room temperature. Zero point fivesix mL of methyl boromoacetate was added dropwise while the mixture was stirred. The mixture was reacted at room temperature for 15-18 h with TLC check. After consumption of the starting material, the mixture was cooled to room temperature, diluted by diethylether, and filtered. The solvent was removed by rotavap and the crude product was purified by flash chromatography on silica gel (5:1 = $CHCl_3$/MeOH) to methyl 2-(4-methylpiperazin-1-yl)aceate (0.111 g, 0.65 mmol) as a pale yellow oil. Yield: 13%. R_f = 0.36 (5:1 = $CHCl_3$/MeOH); ESI-MS $[M+H]^+$ = *m/z* 173.1, CID of $[M+H]^+$ produced *m/z* 113.0 fragment; ^1H NMR (300 MHz, $CDCl_3$): δ 3.68 (s, 3H), 3.18 (s, 2H), 2.52 (m, broad, 8H), 2.26 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.70, 59.36, 54.71, 52.94, 51.70, 45.91.

4.2.3.2. 2-(4-methylpiperazin-1-yl)acetic acid

51.3 mg of methyl 2-(4-methylpiperazin-1-yl)acetate was charged to a 10mL one neck flask with 2 mL of THF and 2 mL of 2M KOH and stirred at room temperature for 10 h. The reaction was monitored by TLC and ESI-MS until the starting material was completely consumed. Upon completion of the reaction, THF was removed by rotavap and the aqueous layer was neutralized by ~2 mL of 2M HCl. Water was then completely removed by rotavap and the residue was reconstituted by acetonitrile (ACN). Insoluble KCl salt was filtered and ACN was removed by rotavap. The free acid of the product was recovered as a transparent gleasy oil. Yield: *quantitative*. ESI-MS $[M+H]^+$ = *m/z* 159.1.

4.2.3.3. N-hydroxylsuccinimide 2-(4-methylpiperazin-1-yl)acetate ester (iTRAQ-113)

In a flame-baked 15 mL glass vial, ~0.65 mmol of 2-(4-methylpiperazin-1-yl)acetic acid and 33.0 mg of TFA-NHS were added to 0.5 mL of dry DMF and 101 mg of *N*,*N*-diisopropylethylamine, and stirred overnight at room temperature. After the complete consumption of the starting material by monitoring TLC, the reaction mixture was used for peptide labeling without further purification. ESI-MS (100% ACN) $[M+H]^+$ = *m/z* 256.1. The stock solution vials of the iTRAQ-113 reagent (*m/z* 113 reporter ion) were prepared as described above and stored in −80 °C.

4.2.4. Protein Mixture Digestion

Two sets of equal amounts (by weight) of protein mixtures, 1) proteins, bovine serum albumin, ovalbumin, α and β caseins, lysozyme and 2) enolase, aldolase, hemoglobin, creatine kinase and alcohol dehydrogenase were digested by trypsin (50:1, w/w) in 50 mM ammonium bicarbonate buffer for 15 h at 37 °C. The reaction was quenched by 5% formic acid, and the resulting peptide products were desalted with C_{18} spin columns (Satorious Stedim Biotech, Aubagne Cedex, France).

4.2.5. Affinity Purification and Digestion of Cul1 and Its Associated Proteins

The applicability of CIT to the quantification of protein complexes in the cell was evaluated with Cul1. Cul1 is a ubiquitin ligase that attaches a ubiquitin chain on target substrates for proteasome-catalyzed degradation.[211] Cul1 is a prototype of the cullin ligase family, and constitutes modular ligase complexes with many confirmed binding partners. The purification of Cul1 and its binding partners from the cells were carried out as described previously with minor

modifications.[212] Briefly, to facilitate the purification of Cul1, a HEK 293-derived stable cell line capable of expressing tagged Cul1 upon tetracycline treatment was constructed using the T-REx™ (Tetracycline-regulated Expression) system (Invitrogen, Carlsbad, CA). The tandem tag we adopted consists of a hexa-histidine sequence and a biotinylation signal sequence.[213] Biotinylation is catalyzed by endogenous biotin ligases, which are present in all eukaryotic cells.[214] A specific lysine residue in the biotinylation signal sequence functions as an acceptor site for biotin *in vivo*.[215]

Tagged Cul1 was induced with 1.0 μg/mL tetracycline for 4 h in experiments aimed to quantify Cul1 complexes with an initial 1:2 mixing ratio for the light and heavy CIT labeling. For the quantification of differentially expressed Cul1 in the cell, 0.5 or 2.0 μg/mL tetracycline were added to the growth medium for 1 h or 4 h, respectively. Twenty four hours after induction, cells were lysed for 30 min at 4 °C with the lysis buffer (0.050 M HEPES, pH 7.5, 0.0050 M Mg(OAc)2, 0.070 M KOAc, 10% glycerol, and 0.4% IGEPAL CA630). The lysate was centrifuged at 16,600 g at 4 °C for 20 min and the supernatant was used for purification. Tandem purification of tagged Cul1 and associated proteins was carried out using MagneHis Ni-particles from Promega and Streptavidin-coupled Dynabeads from Invitrogen. Purified proteins were subjected to sequential on-bead protease digestions, first with Lys-C (35 ng/mg lysate) for 4 h at 37 °C in 8 M urea, and then with trypsin (30 ng/mg lysate) for 12 h at 37 °C in 2 M urea. The resulting tryptic peptides were desalted with C_{18} spin columns.

4.2.6. CIT Labeling

The model peptide, VIP (residue 1-12, HSDAVFTDNYTR; 50 μg), was dissolved in 50 μL of 100 mM tetraethylammonium bicarbonate (pH 8.5) and 100 μL of ACN (66.7% organic phase) and labeled with 5 μL of 5 μg/μL DMSO stock solution of either light or heavy tag, by incubating

for 2 h at room temperature. The reaction was quenched by adding 50 μL of 100 mM hydroxylamine and incubated for 7 h at room temperature. The mixture was acidified by adding 4 μL of formic acid and completely dried by speedvac. The residue was reconstituted by 100 μL of 0.1% formic acid, desalted by the C_{18} desalting tip, and eluted to 100 μL of 0.1% formic acid, 50% ACN and 50% water. The CIT labeled VIP peptide eluent was properly diluted (~x20), and analyzed by various mass spectrometers.

An aliquot of 1 μg of the protein digest was labeled by light or heavy CIT reagents under the same solvent system used for the model peptide, and mixed with the 1:1 ratio. The combined samples were desalted, and injected to a nanoLC-LTQ-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA) at Caltech. Approximately 3 μg of Cul1 digests were labeled by light or heavy CIT reagents with the 1:2 ratio under the same solvent system used for the model peptide. After conjugation, light and heavy CIT-labeled Cul1 digests were combined, desalted and injected to a nanoLC-LTQ-Orbitrap mass spectrometer at Caltech. For HCD/CID experiments, differentially expressed Cul1 digest samples were labeled by light or heavy CIT reagents. The resulting peptides were subject to a nanoLC-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, San Jose, CA) at UCLA for HCD/CID analyses.

4.2.7. Instruments

Matrix-assisted laser desorption/ionization time-of-flight (MALDI TOF) spectra were acquired using Voyager DE-Pro mass spectrometer (Applied Biosystems, Foster City, CA) in the reflectron mode with a 20 kV acceleration voltage, a 150 ns delay extraction time and a 75% grid voltage. A 0.5 μL sample of the derivatized peptide solution was mixed with 0.5 μL of 10 mg/mL CHCA matrix solution in 0.1% TFA, 50% ACN, and 50% H_2O and the mixed spots were dried and introduced to the mass spectrometer for analyses. For all spectra, 100 shots were averaged.

Beam-type CID experiments were performed using Micromass Q-TOF ultra-2 (Waters, UK) in positive ion mode using the Z-spray ion source with a 2.65 kV spraying voltage, a 15 V cone voltage, a 6 V extractor voltage and 22 V (for 3+ ions) or 52 V (for 2+ ions) collision voltages. For beam-type CID spectra, ~100 scans were averaged and used for data analysis.

PQD experiments via direct infusion for model peptide studies were performed by ion trap scans in an LTQ-FTICR mass spectrometer equipped with the Nanomate (Advion BioSciences Inc., Ithaca, NY, USA) nanospray unit. The spraying voltage was 1.4 kV and the gas pressure was 0.3 psi. Critical parameters of the mass spectrometer include capillary voltage 49 V, capillary temperature 200 °C and tube lens voltage 180 V. Other ion optic parameters were optimized by the auto-tune function in the LTQ tune program for maximizing the signal intensity. The parameters for PQD experiments are the precursor isolation window 2.0 *m/z*, collision energy 29%, isolation Q 0.70, and activation time 0.1 ms. For PQD spectra, 100 scans were recorded.

The samples were analyzed by a nanoflow HPLC, Proxeon easy-nLC-System (Proxeon Biosystems) coupled on-line via a nanoelectrospray ion source (Proxeon Biosystems) to a LTQ-Orbitrap mass spectrometer at Caltech. Samples were loaded onto a C_{18}-reversed phase column (15 cm long, 75 μm inner diameter, packed in-house with ReproSil-Pur C_{18}-AQ 3 μm resin (Dr. Maisch)) in buffer A (5% ACN, 0.2% formic acid) with a flow rate of 500 nl/min for 24 min and eluted with a linear gradient from 0% to 36% buffer B (80% ACN 0.2% formic acid) over 110 minutes, followed by 10 minutes at 100% buffer B, at a flow rate of 350 nl/min. The column was re-equilibrated with buffer A. Mass spectra were acquired in the positive ion mode applying data-dependent acquisition with automatic switching between survey scan and tandem mass spectrum acquisition. Samples were analyzed with a top 5 method; acquiring one Orbitrap survey scan in the mass range of *m/z* 400-1600 followed by MS/MS of the five most intense ions in the LTQ in the mass range of *m/z* 100-1600. The target value in the LTQ-Orbitrap was 500,000 for survey

scan at a resolution of 60,000 at *m/z* 400. Fragmentation in the LTQ was performed by Pulsed Q-Dissociation (PQD) with a target value of 5,000 ions. Selected sequenced ions were dynamically excluded for 30 s. General mass spectrometric conditions were: spray voltage, 2.4 kV; no sheath and auxiliary gas flow; ion transfer tube temperature, 200 °C; normalized collision energy (29%) using wide band activation mode for MS/MS. An activation of q = 0.55 and delay time of 0.4 ms were applied in MS/MS acquisitions.[216]

For HCD/CID analyses, CIT-labeled peptides were loaded and washed on a 100 μm×35 mm CVC microtech (Fontana, CA). Peptide trapping was performed using a New Objective Halo C_{18} 75μm × 100 mm, 90 Å, 2.7 μm (Woburn, MA) column by flushing a mobile phase of 0.1% formic acid in water (A). Peptides were subsequently eluted from the column at 300 μL/min using an Eksigent nanoLC 2D pump (Dublin, CA) with a 110 min gradient (0.1% formic acid in water (buffer A) and ACN containing 0.1% formic acid (buffer B); 0-30% phase B over 90min, 30-80% phase B over 20 min). The HPLC system was coupled to an LTQ-Orbitrap XL mass spectrometer and the source conditions were as follows: capillary temperature, 180 C; capillary voltage, 49 V; ESI spray voltage, 1.8 kV. The automatic gain control target was fixed at 5×10^5 ions for MS and 5×10^4 for MS/MS scans. The instrument was operated in data-dependent acquisition mode, with MS survey scan (*m/z* 400-1400*)* performed in the Orbitrap using a resolution set at 60,000. CID and HCD activations were performed on the 3 most abundant ions over 5000 counts (charge state +1 rejected) using normalized collision energies of 30 and 40, respectively, and detected using the linear ion trap. Ions with masses within 10 ppm of previously fragmented ions were excluded for 120s.

4.2.8. Data Processing

The raw files from LTQ-Orbitrap mass spectrometers were converted to mascot generic format (MGF) files using ReAdW4Mascot (version 20090305a, available from the National Institute of Standards and Technology at http://peptide.nist.gov/software/ReAdW4Mascot2_20090305a.zip). PQD of the CIT-labeled model peptide acquired for testing of the linear dependency in reporter ion formation was then analyzed using in-house software and best-fit lines were calculated using linear regression. Error bars are displayed for the middle 95% reported ion ratios. For the five protein mixture, a database was constructed containing the five protein sequences as well as a small contaminant protein database. For the Cul1 pull-down search, a target sequence database was constructed from the human IPI database (version 3.54) and a small containment protein database. A decoy database was constructed from the target following the protocol as described elsewhere.[217] The decoy database was then appended to the target and used to estimate the false discovery rate of the database search. The database search was performed using mascot (version 2.2.06, Matrix Science, http://www.matrixscience.com). The database search parameters were as follows: 0.5 Da fragment ion mass tolerance, 10 ppm precursor ion mass tolerance, trypsin enzyme specificity, up to two missed cleavages, fixed carbamidomethyl (57.02146 Da) modification of cysteine, variable modifications of oxidation (15.99491 Da) of methionine, carbamylation (43.005814 Da) of the N-terminal, and quantitation enabled. The mascot quantitation parameters were as follows: fixed the N-terminal modification of 279.210745 Da with reporter ions at *m/z* 164.1188 and 169.1502, respectively. Reported proteins have at least one unique peptide sequence and two peptide ratios. Reported peptide ratios are included for those peptides whose score is above the homology level and outlier peptide ratios are discarded using the Mascot auto outlier detection. Reported mascot protein quantitation ratios are the median of the top scoring peptide reporter ion ratios.

4.2.9. Density Functional Calculation

The formation of the reporter ion is simulated by the *N*-protonated *N,N*-dimethyl-4-(1-methyl-1*H*-1,2,3-triazol-4-yl)butan-1-amine. Initial coarse geometries were constructed by the MC/MM conformer search using Macromol 8.0 (Schrödinger Inc., Portland, OR, USA) as implemented in Maestro 8.0 (Schrödinger Inc., Portland, OR, USA). The OPLS 2005 was used for the force field model. Within 5 kcal/mol energy, all low energy conformers were initially recorded. Low-energy conformers were selected for further structure optimization by density functional theory (DFT). Each conformer was subject to a geometry optimization using Jaguar 7.5 (Schrödinger Inc., Portland, OR, USA) at the B3LYP/6-311++G(d,p) level. Thermochemical parameters of optimized conformers were estimated by vibrational frequency calculation at 1 atm and 298.15 K at the same level of theory. The transition state structures were searched using the QST method by interpolating initial guesses for reactants, products and transition states. All calculations were performed using computational resources kindly provided by the Material and Process simulation center at the Beckman Institute, Caltech.

4.3.Results and Discussion

4.3.1. Rationale of CIT Design

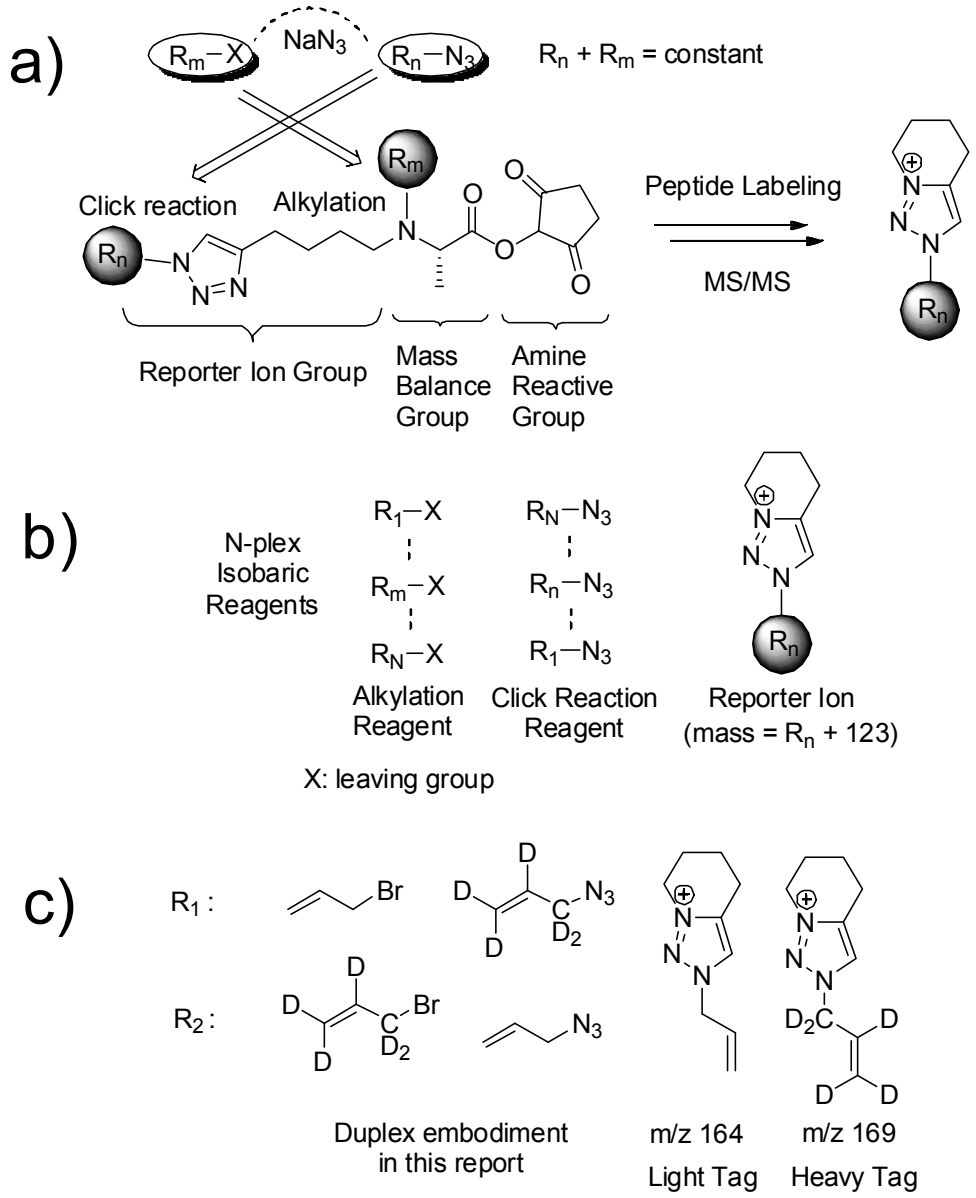
At the inception of this study, a key goal was to find an appropriate gas-phase fragmentation pathway for the formation of the reporter ions. At that time, we were interested in the application of bio-orthogonal azide-alkyne "click" cycloaddition reactions[218-220] to MS-based proteomics studies. Observation of a highly selective gas-phase fragmentation triggered by a nucleophilic substitution of N3 nitrogen in the 1,2,3-triazole ring competitive with the formation of b- and y-

type ions in CID of covalently labeled peptides inspired us to create novel isobaric tags (Scheme 3).

Figure 4.2 depicts the structure of CIT and the construction of the theoretical N-plex reagents. CIT is composed of three parts: the reporter ion group, the mass balance group and the amine reactive group found in other commercially available isobaric tags.

The major improvement of CIT that distinguishes it from other isobaric tags is the *modularization* of the isotope-coded residues, both for the reporter ion group and the mass balance group. Any groups (R) that do not contain other reactive or interfering functionalities can be implemented into the current CIT design by inserting a good leaving group such as bromine, iodine or tosylate. Via a simple S_N2 reaction in the mild conditions employing DMF/NaN_3, the isotope-coded reporter ion group can be easily prepared from an activated R group (**Figure 4.2**). Each isobaric pair of R_m−X (X: leaving group) and R_n−N_3 forms a building block for an isobaric tag with a certain reporter ion mass (R_n + 123 Da). By preparing a set of the N different isotope-coded R−X, it is possible to construct the N-plex isobaric reagents. This modularity of CIT significantly reduces the effort and cost of synthesis. This feature is made possible by the newly discovered low energy gas phase fragmentation pathway depicted in **Scheme 4.3**, which occurs regardless of the attached R groups. In addition, the mass of the reporter ion is *tunable*; this property enables us to bypass the mass cut-off problem in ion trap mass spectrometers and target open windows of m/z values normally found in peptide tandem MS (*e.g.*, sequence ions, immonium ions or internal fragments).

Figure 4.2 Design and structure of CIT



a) The components of N-plex CIT reagents: the reporter ion group, the mass balance group and the amine reactive group. b) Each reporter ion group and mass balance group can be prepared from a series of isotope-coded iodinated R_n groups. CIT labeled peptides are fragmented by various ion activation methods (*e.g.*, PQD, beam-type CID, and High energy Collisional Dissociation (HCD)), yielding the reporter ions whose masses are R_n + 123 Da. c) the duplex embodiment of the CIT reagents in this report by using allyl bromide-d_0 and d_5. Note that the reporter ion is formed regardless of the structure of the attached R_n or R_m groups.

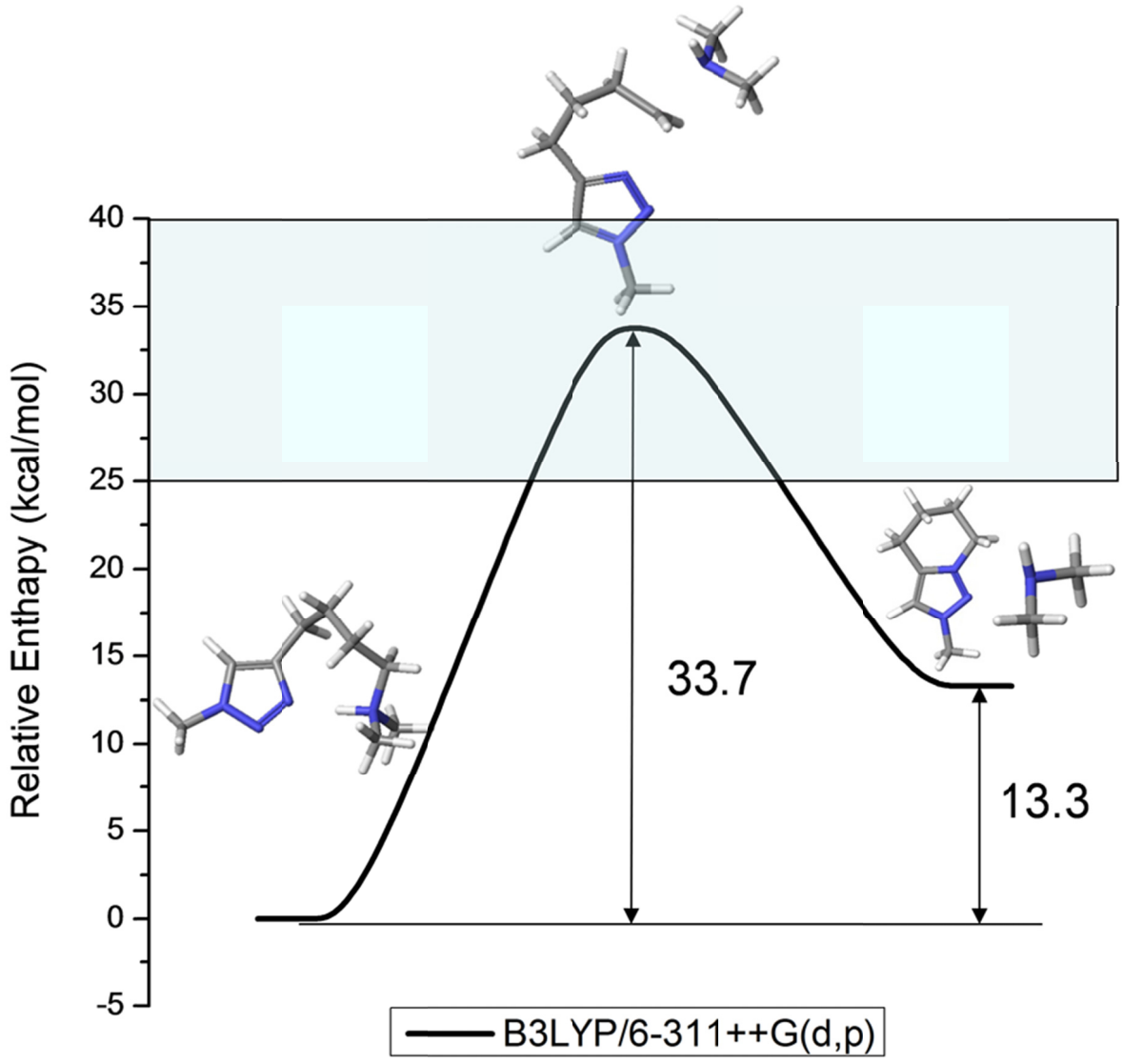
These new discoveries are applied to the creation of a prototype CIT using allyl bromide-d_0 and d_5 as the isotope-coded starting materials. Duplex CIT reagents were synthesized, with heavy and light isotopes having the reporter ions at *m/z* 164 and 169, respectively. The size of the overall modification by this duplex CIT reagent is 279 Da, which is not much larger than most of the commercially available isobaric tags (iTRAQ 4-plex, 144 Da; TMT 6-plex, 224 Da; iTRAQ 8-plex, 304 Da). Only effective collision between peptide amine and *N*-hydroxysuccinimide (NHS) of CIT induces the actual coupling reaction. Therefore, if the size of a certain isobaric tag is relatively large, the kinetics of conjugation can be adversely affected due to decreased frequency of the effective collision. Because the size of CIT is comparable to other isobaric tags, no significant slowdown in CIT labeling is expected.

We adopt the NHS group for facile amine-reactive coupling to peptides as in other commercially available isobaric tags. NHS has been popular in bioconjugation due to the compatibility with most biological buffer solutions. Most importantly, its target functional groups (N-termini of peptides and the ε-amine of lysine) are ubiquitous among tryptic peptides.

The reporter ion fragmentation pathway of CIT is proposed to occur as depicted in **Scheme 4.3**. In multiply protonated CIT-labeled peptides, the tertiary amine in the CIT reagent would be protonated due to its higher proton affinity than most backbone amides and amino acid side-chains. A nucleophilic attack of N3 of the 1,2,3-triazole ring to the C_α of the protonated *N*,*N*-alkylated alanine residue in the CIT reagent releases a stable quaternary ammonium reporter ion, forming a six-membered ring.

The energetics of reporter ion formation is investigated by density functional theory calculations (**Figure 4.3**). If this process is significantly favored compared to backbone fragmentation, less sequence information would be acquired by having fewer and weaker intensity b- and y-type ions in the MS/MS spectrum. It is desirable that activation parameters

Figure 4.3 Energetics of reporter ion formation



The model system, *N,N*-dimethyl-4-(1-methyl-1*H*-1,2,3-triazol-4-yl)butan-1-amine was chosen for calculation. Geometry optimization, thermochemical calculation were performed using B3LYP/6-311++G(d,p) level of theory. The shaded area indicates the range of enthalpies of activation for amide cleavage to form b- and y-type ions via collisional activation.

associated with reporter ion formation are balanced with those of backbone fragmentation. This ensures that accurate protein quantification is achieved while not reducing sequencing efficacy using MS/MS. In our calculation model, the formation of the reporter ion is simulated by the *N*-protonated *N,N*-dimethyl-4-(1-methyl-1*H*-1,2,3-triazol-4-yl)butan-1-amine. At the B3LYP/6-311++G(d,p) level of theory, the reaction barrier and enthalpy at 1 atm and 298.15 K are determined as 33.7 and 13.3 kcal/mol, respectively (**Figure 4.3**). The usual reaction barrier for the amide bond cleavage is ranged from 25 to 40 kcal/mol.[161] Therefore, it is expected that most of the backbone cleavages occur as efficiently as the formation of the reporter ion.

4.3.2. MS/MS of CIT-Labeled Peptides

The light or heavy duplex CIT reagents were labeled to the model tryptic peptide, HSDAVFTDNYTR. The masses of CIT-labeled peptides are 279 Da larger than the original peptides as expected (**Figure 4.4**). The labeling yields of light and heavy CIT reagents are both ~99% estimated by the peak height comparison between unmodified and CIT-labeled peaks in the MALDI TOF MS spectra (**Figure 4.4**). The exact masses of light- and heavy-labeled peptides are identical, appearing as one peak in all mass spectrometric analyses.

Two collisional activation methods, beam-type CID by qTOF and pulsed-Q dissociation (PQD) by a linear ion trap in the LTQ-Orbitrap mass spectrometers, were utilized to fragment CIT-labeled peptides. Beam-type CID of the 2:3 mixture of the light and heavy CIT-labeled peptides in qTOF generates abundant reporter ions at *m/z* 164.1 and 169.1 as well as sequence ions, confirming N-terminal labeling (**Figure 4.5**). CID of the triply protonated precursor ion yields abundant backbone fragment ions along with reporter ions, but only few backbone fragments are observed for the doubly protonated precursor ion. These results are presumably caused by sequestering of mobile protons at the CIT and arginine residues, increasing the reaction barrier

for backbone cleavage. The heavy to light ratio (H/L) of reporter ions is determined to be 1.50 for the triply protonated peptides and 1.26 for the doubly protonated peptides (**Figure 4.5**).

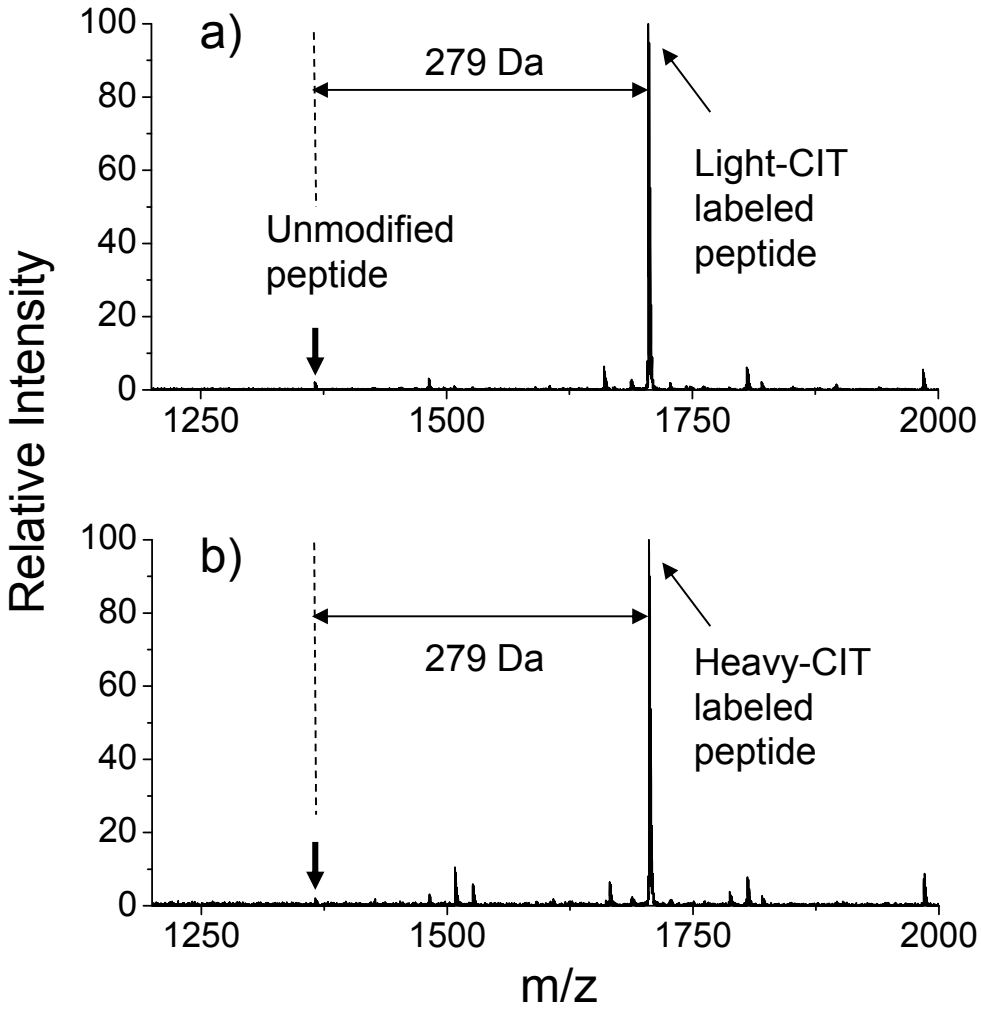
PQD of the 1:1 mixture of light and heavy CIT-labeled peptides in LTQ-Orbitrap produces abundant reporter ions and sequence ions (**Figures 4.6a-b**). More backbone cleavages are observed in PQD of the doubly protonated precursor ions than in beam-type CID. This feature of PQD is attractive for the bottom-up MS-based proteomics, considering dominant doubly protonated peptide ions generated by electrospray ionization (ESI). The H/L ratio of the CIT reporter ion is determined to be ~0.8 in both +2 and +3 charge states (**Figures 4.6a-b**), which may include the initial experimental mixing error.[221]

For the comparison of the energetics for the formation of the reporter ions, the iTRAQ-113 reagent is labeled to the same model peptide and the resulting peptides are subject to PQD (**Figures 4.6c-d**). The overall sequence coverage of the iTRAQ-labeled peptide by PQD is very similar to that of the CIT-labeled peptides. Yet, the relative reporter ion intensities in the PQD spectra of the doubly and triply protonated iTRAQ-labeled peptide ions are lower compared to those of the CIT-labeled peptides (**Figures 4.6c-d**). This result indicates that the process for reporter ion formation of CIT is slight favored over that of iTRAQ, enabling more reliable quantification using abundant CIT reporter ions.

To test the dynamic range of CIT, we monitored reporter ion formation upon mixing various ratios of light- and heavy-labeled peptides using PQD in an LTQ-Orbitrap mass spectrometer. The PQD spectra of doubly and triply charged CIT-labeled peptides were recorded in profile and centroid modes, and the intensities or areas of the reporter ions are used to plot the linear dependency on the initial mixing ratio (**Figure 4.7**, log2-log2 plot). All of the methods for data processing show a good correlation ($R^2 \sim 0.99$) between the initial mixing ratio of light and heavy CIT-labeled peptides. The overall linearity (slopes = ~1.0) and quality of fitting (R^2 = ~0.99)

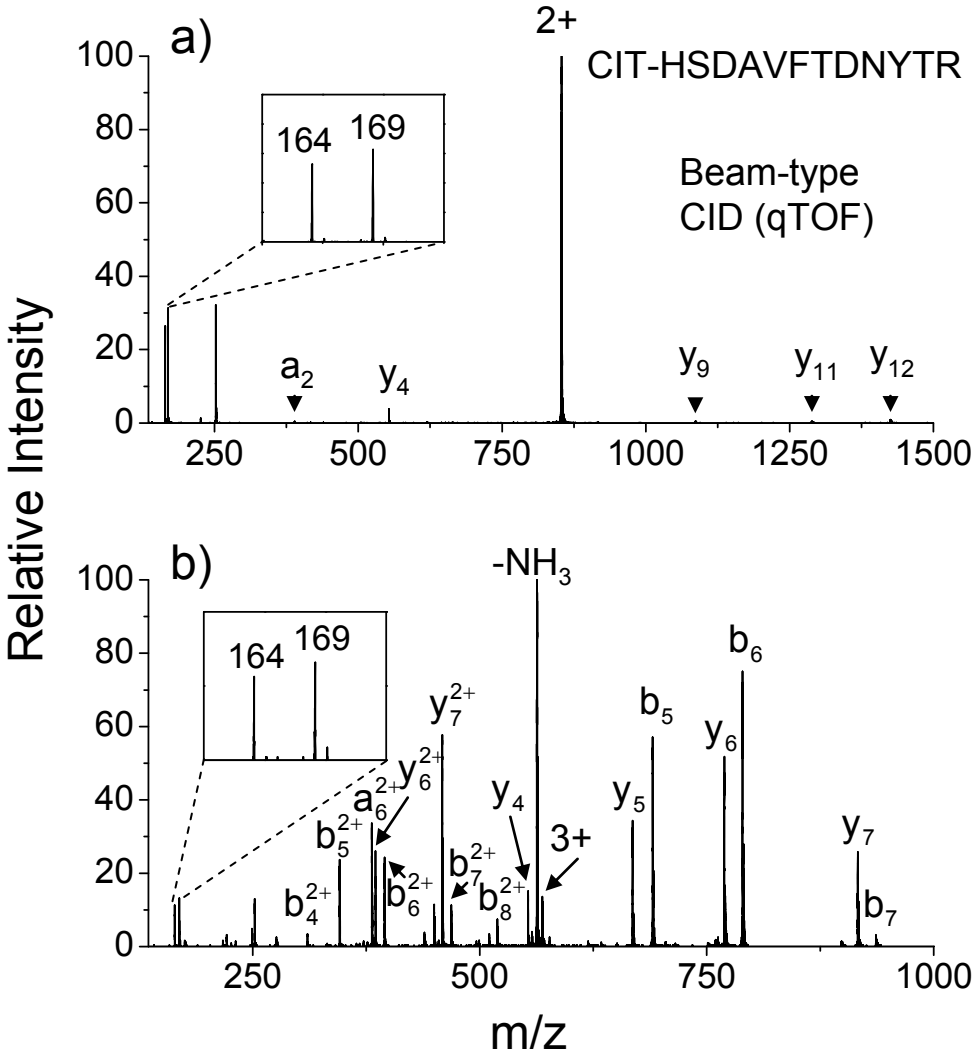
indicate a very good linear response of the CIT reagents from 1/9 to 9 H/L mixing ratios. This demonstrates roughly a two-orders-of-magnitude dynamic range performance of the CIT reagents in the relative quantification.

Figure 4.4 MALDI TOF MS spectra of CIT-labeled peptides



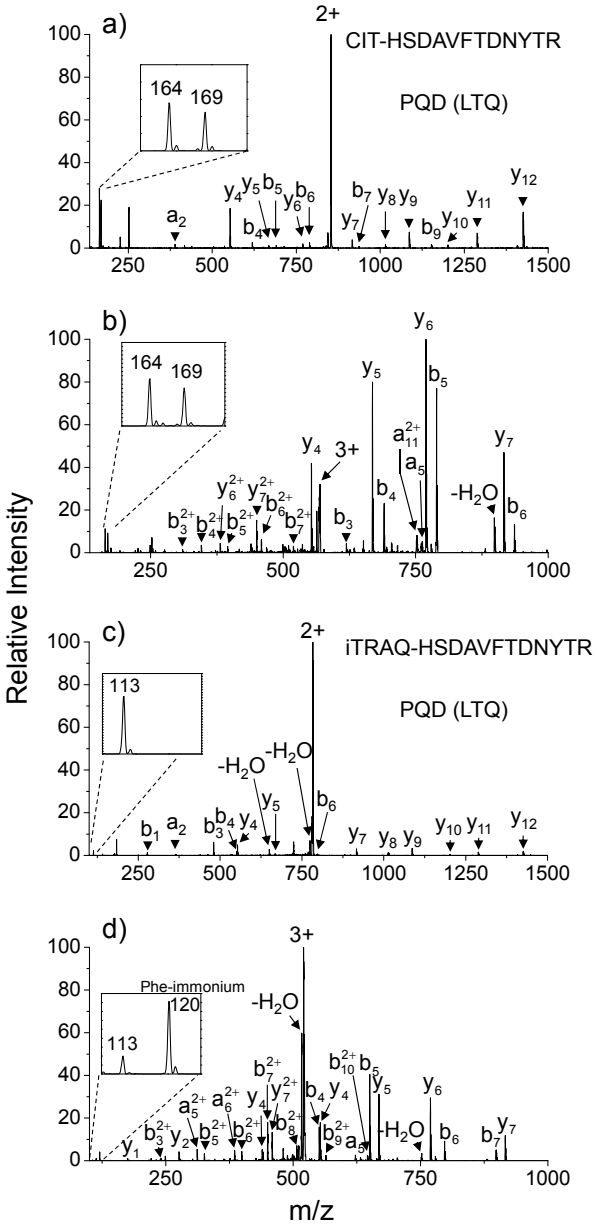
MALDI TOF MS spectra of a) Light and b) heavy CIT reagent labeling of the model peptide, VIP (1-12), HSDAVFTDNYTR. The labeling reaction was performed for 2 hr and quenched by 0.1 M hydroxylamine. The conversion yield is approximately ~99%. Quenching by 0.1 M hydroxylamine reverses unwanted byproducts which contain CIT reagent conjugation on tyrosine residues. Some of the impurities are observed but their contributions are appropriately considered for the calculation of H/L ratios.

Figure 4.5 Beam-type CID of CIT-labeled peptides



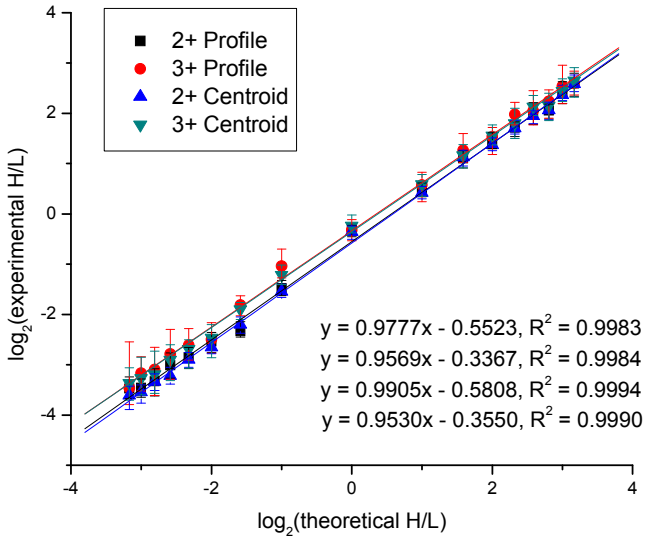
Beam-type CID of a) doubly and b) triply charged CIT-labeled model peptide ions, HSDAVFTDNYTR with 2:3 = light:heavy ratio. Both a) and b) show abundant reporter ions at *m/z* 164 and 169, respectively. Due to relatively short fragmentation time in qTOF, CID of the doubly charged peptide ions in a) shows poor sequence ion formation. CID of the triply charged peptide ions in b) contains rich sequence ion information. Experimentally observed reporter ion ratios (H/L) were 1.26 for 2+ and 1.499 for 3+ (theoretical: 1.5).

Figure 4.6 PQD of CIT-labeled peptides



PQD of doubly and triply protonated CIT- and iTRAQ-113-labeled model peptide ions, HSDAVFTDNYTR with 1:1 = light:heavy ratio. Both a) and b) show abundant reporter ions at *m/z* 164 and 169, respectively, whereas the intensities of the reporter ions at *m/z* 113 in c) and d) are relatively small, indicating the favored energetics of the fragmentation pathway used for the CIT reporter ions. Due to relatively longer fragmentation time in an ion trap compared to qTOF, PQD of the doubly peptide ions in a) and c) generates more sequence ions than that of beam type CID in **Figure 4.5a**. PQD of the triply protonated peptide ions in b) and d) also yields many sequence ions. Experimentally observed reporter ion ratios (H/L) were 0.80 for 2+ and 0.81 for 3+ (theoretical: 1.0). These deviations may result from initial experimental mixing errors.

Figure 4.7 Linearity test of CIT reporter ions



The linear fitting trend lines obtained by calculating the log2 of summations of a) the integration of all areas of isotopes in each reporter ions (*m/z* 164, 165, and 166 for the light tag; *m/z* 168, 169, 170, and 171 for the heavy tag), b) the integration of only *m/z* 164 and 169 peak areas, c) peak heights at *m/z* 164, 165, and 166 for the light tag, and at *m/z* 168, 169, 170, and 171 for the heavy tag, and d) peak heights only at *m/z* 164 and 169 for y-axis and the log2 of intended initial mixing ratios for x-axis. Relatively large (~0.4-0.5) y-axis intercepts in all figures are originated from systematic sources such as initial experimental mixing errors (See a note in the reference). Therefore, the overall linearity (slopes = ~1.0) and quality of fitting (R^2 = ~0.99) are not affected.

4.3.3. Chromatographic Separation

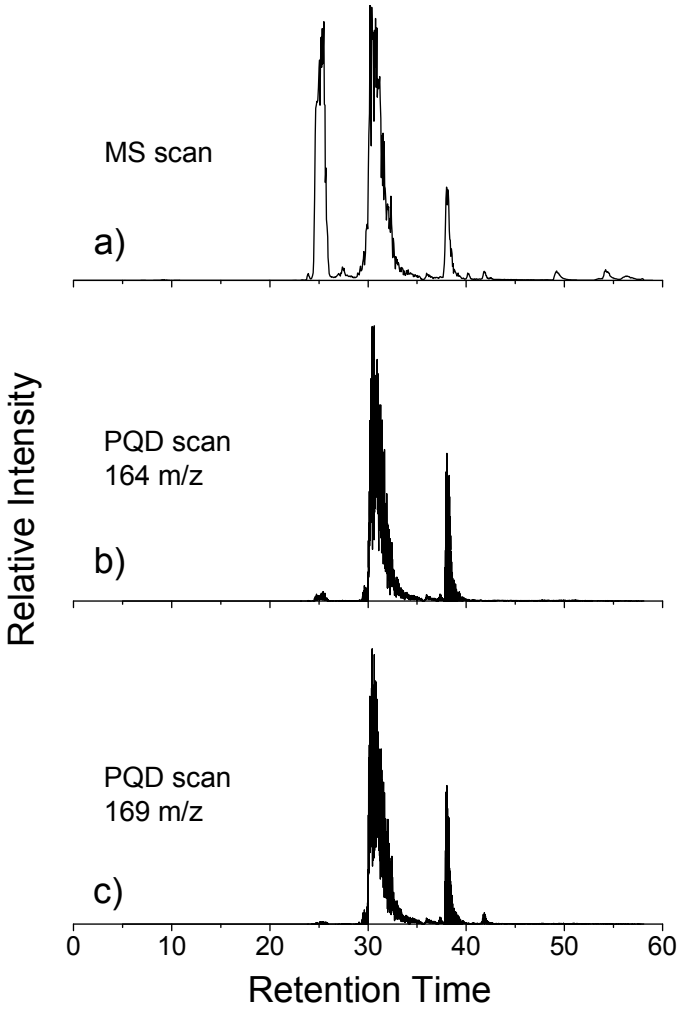
The tailing of deuterated peptides in LC elution profiles has been reported previously and may affect the accuracy of quantification.[206,222,223] To address this possibility, the retention times in nanoLC for both light and heavy peptide are measured. As seen in **Figure 4.8**, ion current diagrams for both reporter ions appear identical with no apparent tailing effect. This result indicates that both light and heavy CIT-labeled peptides have the same chromatographic properties, validating the suitability of the CIT reagents for protein quantification in the LC-MS platform.

4.3.4. Protein Labeling

The applicability of the CIT reagent is tested with model systems involving protein mixtures. Protein digests prepared from the mixtures of 1) bovine serum albumin, ovalbumin, α and β caseins and lysozyme and 2) enolase, aldolase, hemoglobin, creatine kinase and alcohol dehydrogenase were used for CIT labeling with the initial H/L mixing ratio of 1:1. The nanoLC-LTQ-Orbitrap analyses generally reproduce the initial mixing ratio, in which hundreds of peptides tagged by CIT reagents are quantified (**Table 4.1**). Note that geometric standard deviations of Mascot reported mixing ratios are relatively high. Judging from the same phenomenon observed independently in the previous report using PQD for the quantification of iTRAQ-labeled peptides,[224] we believe that the relatively high geometric standard deviation observed is, in large part, induced by poor performance of PQD.

Next, the CIT reagent is applied to quantify biologically relevant samples. Cul1 is a ubiquitin ligase that forms a large protein complex with dozens of known binding partners.[211] This protein complex was purified from the cell, and quantified using CIT after tryptic digest. To facilitate

Figure 4.8 Chromatographic identity of light and heavy CIT-labeled peptides



The nanoLC chromatograms of a) MS1 scans, b) *m/z* 164 reporter ion, and c) *m/z* 169 reporter ion observed in MS/MS scans generated by light and heavy CIT-labeled model peptides, HSDAVFTDNYTR. The base peaks in all chromatograms are related to CIT-labeled model peptides. Note that b) and c) are identical, indicating the same chromatographic property of light and heavy CIT-labeled peptides. The peak at 25 min in a) is a non-labeled model peptide. The peaks appearing around 37 min in b) and c) are from CIT-labeled peptide fragments, AVFTDNYTR.

purification of the Cul1 complex, we constructed a stable cell line that expresses tandem-tagged Cul1 upon tetracycline treatment.[212] Trypsin digests of Cul1 protein complexes affinity-purified from the HEK 293 cell line were split with the ratio of 2:1 for labeling with heavy or light CIT reagents. The labeled samples were combined, and the resulting mixture was analyzed by PQD in the LTQ-Orbitrap. **Table 4.2** lists identified proteins with the H/L ratios determined by Mascot. The calculated medians are close to 2 for all identified proteins, indicating that CIT is suitable for quantification of complex biological samples.

Lastly, the applicability of the CIT reagent for the quantification of relative protein expression levels *in vivo* is investigated. In this study, different amounts of Cul1 was expressed in stable HEK 293 cells by treating two populations of cells with either 0.5 μg/mL for 1 h or 2.0 μg/mL for 4 h. These two samples were subject to Western blot analysis. As shown in **Figure 4.9**, the level of Cul1 expression differs by ratios of 5.26 (Lane 2 / Lane1), and 5.20 (Lane 4 / Lane 3), respectively. For MS analyses, Cul1 was purified from the two differentially expressed samples and digested by Lys-C/trypsin. After CIT labeling, the resulting peptides were analyzed by HCD/CID in an LTQ-Orbitrap mass spectrometer. The median for the H/L ratio of 12 Cul1 tryptic peptides is 5.60, which agrees well with the ratio determined by Western blot experiments. These results demonstrate that the CIT-based quantification is an accurate, reliable methodology for the determination of protein abundance involving complex *in vivo* samples.

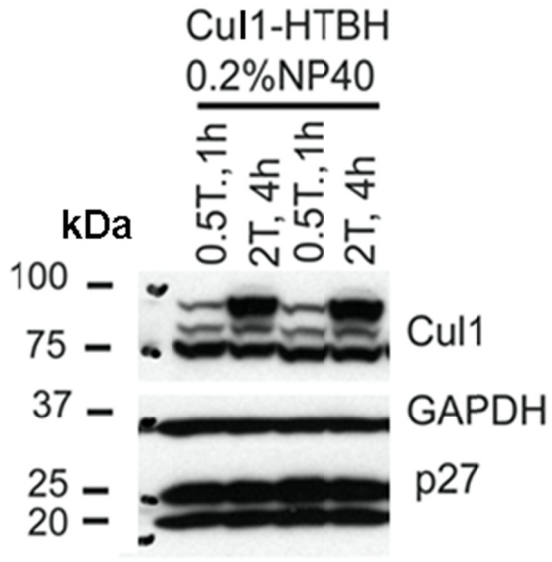
Table 4.1 Mascot Quantification Results of CIT-labeled Protein Mixtures

Mixture Set	Protein	H/L ratio	# of Peptide Hits	Geom. Std.
1	BSA	0.949	45	1.930
	Ovalbumin	0.872	44	1.791
	Lysozyme	1.049	4	1.262
	Alpha-S1-casein	0.887	27	2.086
	Beta-casein	1.063	7	2.232
2	Enolase 1, 2	0.951, 0.991	14, 10	1.651, 1.478
	Aldolase	0.929	18	1.818
	Hemoglobin alpha, beta units	0.790, 0.717	13, 8	1.757, 1.953
	Creatine Kinase	1.094	18	1.818
	Alcohol Dehydrogenase	1.131	2	2.431

Table 4.2 Mascot Quantification Results of Cul1 Complex

Protein	H/L ratio	Geom. Std.	# of Peptide Hits
CUL1	1.820	59	2.494
CAND1	1.779	35	N/A
COPS4	2.047	15	1.859
COPS5	1.635	6	1.749
COPS8	1.474	6	1.428
GPS1	1.642	6	1.123
DCUN1D1	2.834	2	1.079
COPS2	1.357	3	1.827
COPS6	2.277	2	1.156
FBXL18	1.734	3	1.031
SKP1	2.024	4	1.762
FBXO3	1.958	1	N/A
COPS7A	1.895	3	1.449
FBXO42	1.200	2	1.3344

Figure 4.9 Western blot analysis of cross-linked Cul1



Western blot analysis of differentially expressed Cul1 from HEK 293 cells. The amounts of GAPDH and p27 were analyzed as reference proteins. Differential induction was performed by adding 0.5 or 2.0 μg/mL tetracycline to the growth medium for 1 h or 4 h, respectively (Lane 1 and 2). Lane 3 and 4 are the replicates of Lane 1 and 2, respectively. The relative Cul1 expression level was 1:5.16 for Lane 1 and Lane 2, and 1:5.20 for Lane 3 and Lane 4. HCD/CID of CIT-labeled Cul1 digests yielded median H/L ratio of 5.60 using 12 peptides. The geometric standard deviation was 2.15.

4.4.Conclusion

A novel isobaric tag is developed for protein quantification, referred to as Caltech Isobaric Tag (CIT), with excellent demonstrated performance in a range of typical proteomics investigations employing model systems. The design of the CIT reagents is based on a novel gas-phase fragmentation pathway reported here for the first time. In this pathway, a nucleophilic attack of N3 of the 1,2,3-triazole ring releases a stable quaternary ammonium reporter ion with concomitant formation of a six-membered ring. The mass of the reporter ion can be easily tuned by varying azide groups in the preparation of the 1,2,3-triazole ring via Copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC), better known as click reaction. The number of the possible isobaric tags is determined by the number of the isotope-tagged azide groups. These azides can be prepared from halogenated alkyl groups, which are also used for the alkylation of the linker amino acids, reducing both the cost of reagents and effort required for the synthesis of isobaric tags. This modular feature expands the possible number of combinations of CIT reagents. The properties of CIT reagents can be tuned by using larger isotope-coded halogenated alkyls that yield higher *m/z* reporter ions and these avoid the low mass–cut off problems normally associated with ion trap mass spectrometers. Mixtures of light and heavy CIT-labeled model peptides showed good linear correlations with a two-orders-of-magnitude dynamic rage. Observed ratios of the light and heavy CIT-labeled protein digests from the mixtures of 1) bovine serum albumin, ovalbumin, α and β caseins and lysozyme and 2) enolase, aldolase, hemoglobin, creatine kinase and alcohol dehydrogenase also exhibited good agreement with the initial mixing ratios. Lastly, we have demonstrated the applicability of CIT reagents in quantifying complex biological samples using affinity-purified Cul1 ubiquitin ligase complexes from HEK 293 cells.

4.5.Acknowledgement

This work was supported by the National Science Foundation through grant CHE-0416381 and the Beckman Institute at California Institute of Technology. Computational resources for density functional theory results were kindly provided by the Materials and Process Simulation Center at California Institute of Technology. C.H.S. acknowledges a fellowship from the Kwanjeong Educational Foundation. J.E.L. was supported by the Ruth L. Kirschstein NRSA fellowship from the NIH (CA138126).

5. Designer Reagents for Mass Spectrometry-Based Proteomics: Clickable Cross-Linkers for Elucidation of Protein Structures and Interactions

ABSTRACT We present novel homobifunctional amine-reactive clickable cross-linkers (CXLs) for investigation of three-dimensional protein structures and protein-protein interactions (PPIs). CXLs afford distinct advantages over other currently available cross-linkers, including (1) their small size and cationic nature in physiological pH for good water solubility and cell permeability, (2) an alkyne group for bio-orthogonal conjugation to a biotin-azide or other affinity tags via click reaction for enrichment of cross-linked peptides, (3) a highly selective nucleophilic displacement reaction by the resultant 1,2,3-triazole yielding a reporter ion for fast screening of cross-linked peptides, (4) a readily modifiable chain length between the amine-reactive groups to act as molecular "rulers" for enhanced elucidation of structural details. Ubiquitin, a small and lysine-abundant protein, is used as a model system to demonstrate protein structural studies using CXLs. To validate the sensitivity of our approach, biotin-azide labeling and subsequent enrichment of cross-linked peptides are performed for cross-linked ubiquitin digests mixed with yeast cell lysates. Cross-linked peptides are successfully detected and identified by collision induced dissociation (CID) and electron transfer dissociation (ETD) in LTQ-FTICR and LTQ-Orbitrap mass spectrometers. The observed cross-linked lysine residues and their connectivities are consistent with the high resolution X-ray crystallography structure of ubiqutin. The application of CXLs in more complex systems (*e.g. in vivo* cross-linking) is illustrated by detection of Cul1 complexes in HEK 293 cells, confirming good water solubility and cell permeability.

KEYWORDS Chemical Cross-linkers, Cross-linked Peptide, Protein Structure, Protein-Protein Interaction, Mass Spectrometry, Copper-Catalyzed Azide-Alkyne Cycloaddition, Reporter Ion.

5.1.Introduction

Elucidation of three-dimensional structures of protein complexes and protein-protein interactions (PPIs) is one of the central goals in current biological research. Proteins bind to each other to carry out specific biological functions by forming various protein complexes.,[225,226] On average, proteins *in vivo* act not alone but rather as part of a protein complex comprising 10 protein subunits in the cell.[227] Proteolysis by poly-ubiquitination is a good example for a functional protein complex.[228] The development of new reagents and methods for identification of binding partners and their interfaces is an enabling part of proteomic science.

Chemical cross-linkers have been widely employed in analysis of three-dimensional protein structures and protein-protein interactions (PPIs).[229,230] For identification of cross-linked proteins, traditional experimental methodologies including affinity-based chromatography and Western blot have been performed. However, no detailed structural information relative to the nature of specific protein interfaces is revealed in these experiments. Full atomistic structures of isolated proteins and their complexes can be obtained from NMR spectroscopy and X-ray crystallography but these methodologies usually require large amounts of sample for analysis. Crystallization of diffraction quality protein complexes is often the bottleneck in structure determination by X-ray crystallography.

Recently, mass spectrometry (MS)-based analysis has allowed detection of binding partners and specific contacting residues in more sensitive ways.[21,23,24,231-235] *In vitro* cross-linking and enzymatic digestion produce cross-linked peptides containing spatial information between residues reactive with the cross-linker. This topological information constrains relative distances of amino acid residues, thus aiding in the reconstruction of protein complex subunits.

For investigation of *in vivo* PPIs, protein complex immunoprecipitation (*i.e.*, co-IP or "pull-down") is often performed to recover strongly interacting partners, such as an enzyme bound to its inhibitor. Co-IP requires the use of several antibodies to validate putative binding partners by

running successive rounds of experiments. Alternatively, affinity tags can be infused into genes of target proteins to permit efficient purification from cell lysates.[236] However, many of the important signaling pathways are believed to be relayed via weak interactions that occur at the outside of strongly bound core protein complexes,[226] and co-IP often fails to identify those weak binding partners. Chemical cross-linking has been performed to freeze weak interactions by forming covalent bonds, and then sample analysis is usually combined with other targeted protein purification techniques.[213,237] Some previous reports have shown promising results by taking an *in vivo* chemical cross-linking strategy for investigation of PPIs.[237,238]

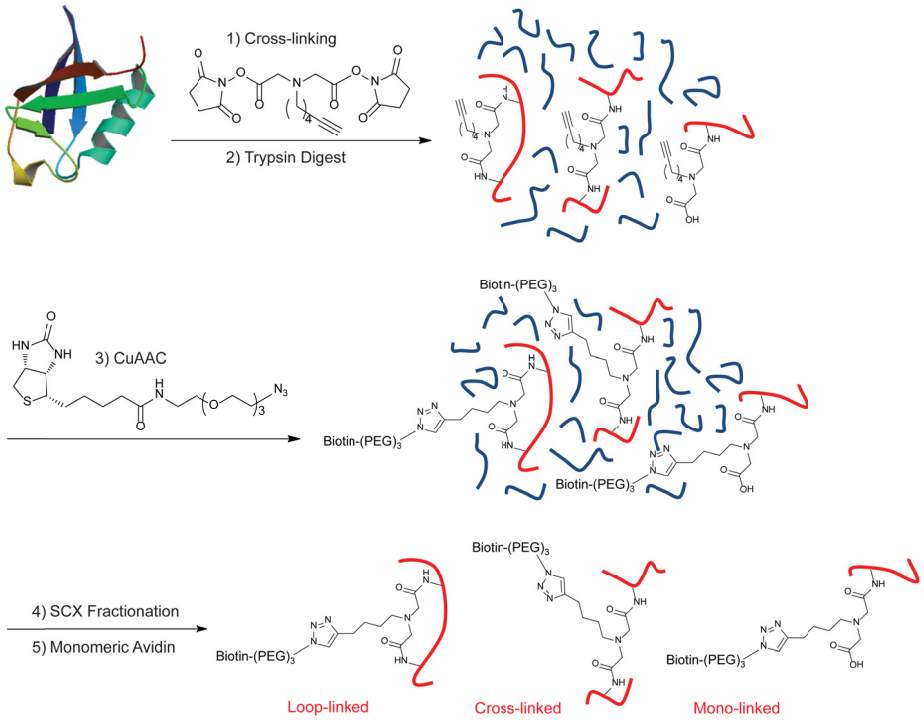
Nonetheless, most chemical cross-linking studies have been limited to purified proteins and specific targets. The formidable complexity of protein interaction networks greatly hinders identification of PPIs using cross-linking strategies at the systems level. The complexity of protein samples steeply increases upon cross-linking, roughly proportional to the half of the square of the numbers of enzymatically cleaved peptides in the cell.[234] It leads to two practical problems associated with computational and experimental challenges. A protein cross-link searching program accounting for PPIs in mammalian cells on a genome-wide scale is not available, and current computational platforms are limited to just a few proteins. Computational resources required for this type of unrestricted searching algorithm are therefore substantial. A recent study done by the Aebersold group tackled this problem by introducing a new searching algorithm, xQuest, which reduces the searching space by an upstream candidate-peptide search and by isotope-coded cross-linkers.[239] They successfully showed that cross-linked peptides are identified from a total *E. coli* lysate with an unrestricted database search.

For selective and sensitive detection of cross-linked peptides, functionalized chemical cross-linking reagents are required. Various designs including biotinylated[240-242], isotope-coded[243-246], fluorophore labeled[20,22,247], mass-tag labeled[248], amidinating[249] and chromophore labeled[250] cross-linking reagents have been reported. However, the addition of functional groups can often cause the cross-linker to become very bulky or less cell-permeable, thus not very effective for *in vivo*

cross-linking.[238] To reduce the total size of the cross-linker, separation of the cross-linking step from conjugation of affinity tags is one effective strategy.

New cross-linking and enrichment strategies for separation of the cross-linking reaction from enrichment steps have recently been developed based on bio-orthogonal chemistries such as the azide-alkyne "click" cycloaddition[218-220] and Staudinger ligation[251] using alkyne[252,253] or azide[254,255] tagged cross-linkers. Azides and alkynes are not naturally found in proteins, peptides, nucleic acids or glycans. The orthogonality of azides and alkynes to biological processes (*i.e.*, competing reactions) is a significant advantage of this approach, and holds a great promise for protein cross-linking studies. Moreover, the "click" cycloaddition can be performed under aqueous conditions, allowing the enrichment of cross-linked products by conjugation of an appropriate affinity tag.

Here we report a novel clickable cross-linker (CXL) that addresses and overcomes many of the current challenges described above (**Scheme 5.1**).



Scheme 5.1

CXLs offer several advantages over previous amine-reactive cross-linkers: (1) small size, (2) good cell permeability, (3) water solubility, (4) ease of synthesis, (5) alkyne group for bio-orthogonal conjugation to a biotin-azide affinity label via copper-catalyzed azide-alkyne cycloaddition (CuAAC, or click reaction) for enrichment of cross-linked peptides by avidin-biotin chromatography, (6) a highly selective nucleophilic displacement reaction by the resultant 1,2,3-triazole in the gas phase, yielding a reporter ion at *m/z* 525.3 for fast screening of cross-linked peptides and (7) synthetically tunable chain length enabling the preparation of a group of CXLs to acquire enhanced structural information using various lengths as molecular "rulers". Ubiquitin, a small and lysine-abundant protein, is chosen as our model system for three-dimensional protein structure study using a hexynyl CXL. To demonstrate the non-denaturing feature of CXLs in solution, secondary structures of cross-linked ubiquitin are monitored using circular dichroism spectrometry. After cross-linked peptides are labeled with biotin-azide by CuAAC, they are enriched by strong cation exchange fractionation (SCX) and avidin affinity chromatograpy. Cross-linked peptides from ubiquitin are analyzed by collisional induced dissociation (CID) and electron transfer dissociation (ETD) in LTQ-FTICR and LTQ-Orbitrap mass spectrometers. Monitoring the gas-phase reporter ion at *m/z* 525.3 by CID allows fast screening of results with high confidence. Highly charged cross-linked peptide ions are observed due to basic residues in CXL (*i.e.* tertiary amine and the 1,2,3-triazole), yielding high quality ETD spectra and improved sequencing. The application of CXL in more complex systems is also tested *in vivo* by cross-linking of Cul1, a ubiquitin ligase E3, in HEK293 cells, confirming good water solubility and cell permeability.

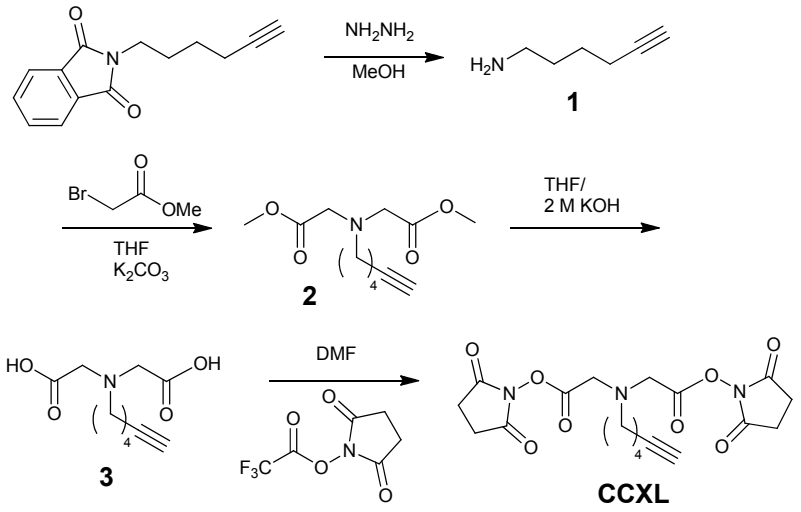
5.2.Experimental Section

5.2.1. Materials

N-(5-Hexynyl)phthalimide, anhydrous diethyl ether (Et_2O), lithium aluminum hydride (1 M in Et_2O), *N*-hydroxysuccinimide, methyl bromoacetate, trifluoroacetic anhydride, tetraethylammonium bicarbonate, tris(2-carboxyethyl)phosphine, and bovine ubiquitin were obtained from Sigma-Aldrich Co. (St. Louis, MO). The hydrophilic ligand, tris[(hydroxyethyl-triazolyl)methyl]amine (TBTA-OH) was a kind gift from the Sharpless and Fokin groups at The Scripps Research Institute. The model peptide, Ac-AAKAAAAAKAR (98% purity), was obtained from Biomer-Tech (Pleasanton, CA). Yeast cell lysate was a generous gift from the Coon group at the University of Wisconsin-Madison. Monomeric avidin resin kit was purchased from Pierce (Rockford, IL). Biotin-$(PEG)_3$-azide was acquired from Berry & Associates, Inc. (Dexter, MI). Sequencing grade trypsin was obtained from Promega (Madison, WI). OMIX-C18 desalting tips (100 μL capacity) were purchased from Varian, Inc. (Palo Alto, CA). Microcon YM-3K spin filter units and C18-ZipTip (10 μL scale) were purchased from Millipore (Billerica, MA). Microspin SCX columns were acquired from the Nest Group, Inc. (Southborough, MA). High Capacity Neutravidin Agarose Resin, n-dodecyl-β-D-maltoside, and SuperSignal West Dura Extended Duration chemiluminescent substrate were from Thermo Scientific (Rockford, IL). Lysyl endopeptidase (LysC) was from Wako Chemicals USA (Richmond, VA). Cell culture reagents, Flip-In T-REx 293 cells, plasmids, and monoclonal antibodies for Cul1 and Cand1 were from Invitrogen (Carlsbad, CA). Plasmid DNA containing the human Cul1 sequence was purchased from Open Biosystems (Huntsville, AL). MLN4924 was a generous gift from Millennium: The Takeda Oncology Company (Cambridge, MA). All other general chemicals for buffers were purchased from Fisher Scientific (Hampton, NH), EMD (Gibbstown, NJ), VWR International (West Chester, PA), and Sigma-Aldrich (St. Louis, MO). All chemicals mentioned above were used as received without further purification.

5.2.2. Synthesis of Clickable Cross-Linker (CXL)

Overall synthetic steps are summarized in **Scheme 5.2**.



Scheme 5.2

5.2.2.1. 6-Amino-hex-1-yne (1)

The amino alkyne was prepared from 2-(hex-5-ynyl)isoindoline-1,3-dione and hydrazine as described in the literature without modification.[256] The crude product was purified by flash chromatography (silica gel, dimethylene chloride:methanol = 3:1 ~ 1:1) to yield 6-amino-hexyne as a pale greenish yellow oil. Yield: 30%. ^{1}H NMR spectra is reproduced as reported.[257] ESI-MS $[M+H]^+$ *m/z* 98.1.

5.2.2.2. Dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate (2)

The 6-amino-hex-1-yne (0.3 g) was added to the stirring solution of 20 mL THF, 2 eq K_2CO_3, and 2.4 eq methyl bromoacetate. The mixture was further stirred at room temperature for 3 h under a stream of dry N_2. The reaction was monitored by thin layer chromatography (TLC) using hexanes:ethyl acetate (= 2:1) as the mobile phase, and the mixture was filtered after completion of the reaction. The filtrate was concentrated and purified by flash chromatography (silica gel, hexanes:ethyl acetate = 1:1). The final product, dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate, was

concentrated by rotary evaporation and acquired as a transparent oil. Yield: 59%. ESI-MS $[M+H]^+$ *m/z* 242.1, ^1H NMR ($CDCl_3$) δ 3.71 (s, 6H), 3.56 (s, 4H), 2.73 (t, 2H), 2.22 (m, 2H), 1.94 (t, 1H), 1.57 (m, 4H).

5.2.2.3. 2,2'-(hex-5-ynylazanediyl)diacetic acid (3)

To the obtained ~1.4 g dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate was added 20 mL THF and 20 mL of 2 M KOH. The mixture was stirred overnight and monitored by TLC. The organic layer was separated, and the aqueous layer was quenched by addition of 20 mL of 2 M HCl. The solvent (H_2O) was completely removed by rotary evaporation, and the resulting solid was dissolved into acetonitrile (ACN). The insoluble KCl salt was filtered, and the filtrate was concentrated by rotary evaporation. The final product of free acid was obtained as a greasy transparent oil. Yield: quantitative. ESI-MS: $[M+H]^+$ *m/z* 214.1. To obtain the hydrochloride salt, an additional 10 mL of 2 M HCl was added before removal of the solvent. Dimethyl formamide (DMF, 3 × 30 mL) was added to the resulting solid and filtered. The hydrochloride salt was obtained as a white solid after concentration under reduced pressure. Yield: quantitative.

5.2.2.4. NHS-activated 2,2'-(hex-5-ynylazanediyl)diacetic acid (CCXL)

N-hydroxysuccinimide trifluoroacetate was prepared by stirring *N*-hydroxysuccinimide (NHS) and 4 eq trifluoroacetic anhydride for 5 h. The mixture was concentrated under reduced pressure and further dried under high vacuum overnight. The product was obtained as a white, highly hygroscopic solid and stored in an anhydrous desiccator before use. The obtained ~1.1 g 2,2'-(hex-5-ynylazanediyl)diacetic acid was activated by 2.4 eq *N*-hydroxysuccinicimide trifluoroacetate in 10 mL anhydrous DMF under a stream of dry N_2. The mixture was stirred overnight and monitored by TLC using hexanes:ethyl acetate (= 2:1) as the mobile phase. After completion of the reaction, the mixture was concentrated to ~500 μL by rotary evaporation and subjected to flash chromatography using hexanes:ethyl acetate (= 2:1) as the mobile phase. The

final product, NHS-activated 2,2'-(hex-5-ynylazanediyl)diacetic acid (CCXL) was concentrated by rotary evaporation, and obtained as a pale yellow oil. Several 200 µL aliquots of 50 mM stock solution dissolved in anhydrous dimethyl sulfoxide (DMSO) were prepared and stored at –80 °C. The sealed stock aliquots were opened immediately before use, and NHS activation was verified by ESI-MS in 100% ACN. ESI-MS $[M+H]^+$ *m/z* 408.1, $[M+Na]^+$ *m/z* 420.0. (Note: the reaction yield can be improved by adding stoichiometric equivalent of triethylamine.)

5.2.3. Cross-Linking of a Model Peptide

A 50 µg portion of the model peptide, Ac-AAKAAAAAKAR (98% purity), was dissolved in 50 µL of HPLC grade H_2O. A mixture of 5 µL CXL stock solution (10 µg/µL in DMSO), 5 µL Ac-AAKAAAAAKAR stock solution (10 µg/µL), and 15 µL ACN was prepared. The mixture was allowed to react at room temperature for 1 h. The reaction was terminated by adding 5 µL formic acid (FA). The solvent was completely removed by speed-vac, and the residue was reconstituted in 100 µL of 0.1% FA (aq) with additional 2 µL FA to further acidify. The resulting solution was desalted using an OMIX-C18 tip (100 µL capacity) following the standard procedure. The cross-linked peptide (~50 µg) was eluted in 100 µL solution of 0.1% FA, 50% ACN, and 50% H_2O, and 5 µL of the eluted cross-linked peptide solution was diluted to 5 µM by 0.1% FA, 50% ACN, and 50% H_2O, and analyzed by a LCQ ion trap mass spectrometer. The remaining cross-linked peptide solution was dried for click reaction.

CuAAC with the biotin-(triethyleneglycol)-azide (biotin-$(PEG)_3$-azide) was performed as follows: Ten µg of the CXL cross-linked Ac-AAKAAAAAKAR peptide were dissolved in 100 mM tetraethylammonium bicarbonate (TEAB) at pH 8.5, 250 µM hydrophilic ligand tris[(hydroxyethyl-triazolyl)methyl]amine (TBTA-OH), 2.5 mM $CuSO_4$•$5H_2O$, 5 mM tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, and 1 mM biotin-$(PEG)_3$-azide in a total volume of 100 µL containing 99% H_2O and 1% DMSO (from the TBTA-OH stock). The mixture was reacted for 2 h with gentle mixing at 37 °C and quenched by 5% FA (aq). The solvent was

removed by speed-vac, and the residue was desalted by OMIX-C18 tip as described above. The eluent was properly diluted to 10 μM with 0.1% FA, 50% ACN, and 50% H_2O and directly infused to the LCQ ion trap mass spectrometer for analysis.

5.2.4. Cross-Linking of Ubiquitin

The general reaction summary of cross-linking of ubiquitin (PDB ID: 1UBQ) is shown in **Scheme 5.1**. Twenty μg of ubiquitin were dissolved in 200 μL of 1X PBS (pH = 7.4) and 1.2 μL of 50 mM CXL stock solution in DMSO was added and reacted for 30 min at room temperature. The reaction was quenched by 50 μL of 100 mM Tris-HCl buffer (pH = 8.5) and incubated for 15 min. The cross-linked ubiquitin was concentrated to ~30 μL, and the buffer was exchanged to 100 mM ammonium bicarbonate at pH 8.5 using Microcon YM-3K spin filter units. The trypsin digest reaction volume was adjusted by adding 185.5 μL of 100 mM ammonium bicarbonate buffer containing 2 M urea, and 2.5 μL of 100 mM $CaCl_2$. Two μL of 0.5 μg/μL trypsin in 5 mM acetic acid (proteins:trypsin = 20:1 w/w) was added and incubated for 15-18 h at 37 °C. The reaction was terminated by addition of 5% FA (aq). The resulting tryptic digest was desalted by OMIX-C18 tip and a 1 μg portion was injected into a nanoLC-LTQ-FTICR mass spectrometer for analysis.

Forty μg of the cross-linked tryptic digest of ubiquitin were subjected to click reaction by combining resulting peptides from two identical cross-linking experiments. The desalted tryptic digest was dissolved in 100 mM TEAB, 250 μM hydrophilic ligand TBTA-OH, 2.5 mM $CuSO_4 \bullet 5H_2O$, 5 mM TCEP hydrochloride, and 1 mM biotin-$(PEG)_3$-azide in a total volume of 100 μL containing 99% H_2O and 1% DMSO (from TBTA-OH stock). An additional sample was prepared by mixing 50 μg cross-linked digest of ubiquitin with 50 μg yeast cell lysate and subjected to similar click reaction conditions to demonstrate enrichment from a complex sample. Mixtures were reacted at 37 °C for 12 h with gentle shaking. Reactions were quenched by addition of 5% FA (aq).

Microspin SCX columns (200 µL scale, with 50 µL of the bed volume for SCX material) were used for removal of excessive TBTA-OH and biotin-$(PEG)_3$-azide. A 10 µg portion of the peptides from the click reaction (25 µL) was dried to completeness by speed-vac, and the residue was reconstituted with 0.5% FA, 5% ACN (aq). Microspin SCX columns were prepared by applying 4 bed volumes (200 µL) of MeOH then H_2O, respectively. Activation of the SCX material was performed by 200 µL of 500 mM ammonium acetate and incubated for 1 h at room temperature. After activation, the spin columns were washed by H_2O and equilibrated with 0.5% FA, 5% ACN (aq). The peptide sample solution was applied to the spin column and flushed twice to bind completely. The spin column was washed by 400 µL of 0.5% FA, 5% ACN (aq), which corresponds to at least 8 bed volumes of the SCX material. The peptides were fractionated by 400 µL of 50, 250, and 500 mM ammonium acetate in 0.5% FA, 25% ACN (aq), and additional 500 mM ammonium acetate solution was used for complete elution of highly charged cross-linked peptides. Each fraction was desalted by C18-ZipTip following manufacturer instructions, and eluents were dried by speed-vac. The residues were reconstituted with 5 µL of 0.2% FA (aq) and injected to a nanoLC-LTQ-FTICR mass spectrometer for analysis.

Monomeric avidin-biotin affinity chromatography was performed using the batch style procedure according to the manufacturer manual with modifications. Peptide samples after click reaction were eluted without SCX fractionation using either 500 mM ammonium acetate in 0.5% FA, 25% ACN (aq) or 50 mM ammonium acetate in 0.1% TFA, 25% ACN (aq). The SCX eluents were dried by speed-vac and readjusted to 1X PBS at the same concentration range used in the cross-linking reaction. By incubating the mixture of the monomeric avidin resin and peptide samples at room temperature or 4 °C for 12 h under the gentle mixing, the biotin-PEG_3-azide conjugated peptides were bound to monomeric avidin. Unmodified peptides were washed away by flushing 4 bed volume capacity of PBS, 100 mM Tris-buffer (pH 7.4), 100 mM ammonium bicarbonate (pH 7.4), and water. The final products of interest were eluted by 0.4%

TFA, 50% ACN (aq). An aliquot of the eluent was analyzed by a nanoLC-LTQ-FTICR mass spectrometer.

5.2.5. *In vivo* Cross-Linking of Cul1

The applicability of CXL to *in vivo* cross-linking was evaluated with Cul1. Cul1 is a ubiquitin ligase that attaches a ubiquitin chain on target substrates for proteasome-catalyzed degradation.[211] Cul1 is a prototype of the cullin ligase family, and constitutes modular ligase complexes with other binding partners. The *in vivo* cross-linking by CXL and following Western blot analysis were carried out as described previously with minor modifications.[212] Briefly, to facilitate the purification of Cul1, a HEK 293-derived stable cell line capable of expressing tagged Cul1 upon tetracycline treatment was constructed using the T-REx^{TM} (Tetracycline-regulated Expression) system (Invitrogen, Carlsbad, CA). The tandem tag we adopted consists of a hexa-histidine sequence and a biotinylation signal sequence.[213] Biotinylation is catalyzed by endogenous biotin ligases, which are present in all eukaryotic cells.[214] A specific lysine residue in the biotinylation signal sequence functions as an acceptor site for biotin *in vivo*.[215]

Tagged Cul1 was induced with 1.0 μg/mL tetracycline for 4 h in experiments for *in vivo* cross-linking. Twenty four hours after induction, cells were subject to *in vivo* cross-linking by treating 0, 0.1, 0.2, 0.5, and 1.0 mM of CXL, respectively and incubated for 1h at 37 °C. After the completion of the cross-linking, cells were lysed for 30 min at 4 °C with the lysis buffer (0.050 M HEPES, pH 7.5, 0.0050 M Mg(OAc)2, 0.070 M KOAc, 10% glycerol, and 0.4% IGEPAL CA630). The lysate was centrifuged at 16,600 g at 4 °C for 20 min and the supernatant was used for Western blot analysis.

5.2.6. Mass Spectrometry

The CXL cross-linked Ac-AAKAAAAAKAR model peptide was analyzed by a LCQ-deca XP ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The sample was directly

infused by the standard electrospray ionization unit with the constant flow at 3 μL/min. The critical instrumental parameters were set up as follows: the spraying voltage at 3.0 kV, capillary voltage at 25 V, capillary temperature at 200 °C, and tube lens voltage at −25 V. Fifty scans were recorded for each spectrum.

The cross-linked peptides from ubiquitin were analyzed by a nanoflow HPLC (Waters Co.) coupled on-line via a home-built nanoelectrospray ion source to a LTQ-FTICR mass spectrometer (Thermo Fisher Scientific). Samples in 5 μL of 0.2% FA (aq) were loaded onto a C_{18}-reversed phase column (15 cm long, 100 μm inner diameter, packed in-house with Magic C_{18}-AQ 5 μm resin (Michrom Bioresources) in buffer A (2% ACN, 0.2% FA) with a flow rate of 250 nl/min for 24 min and eluted with a linear gradient from 0% to 36% buffer B (98% ACN, 0.2% FA) over 110 min, followed by 10 min at 100% buffer B, at a flow rate of 250 nl/min. The column was re-equilibrated with buffer A. Mass spectra were acquired in the positive ion mode applying data-dependent acquisition with automatic switching between survey scan and tandem mass spectrum acquisition. Samples were analyzed with a top 10 method; acquiring one FTICR survey scan in the mass range of *m/z* 400-1600 followed by MS/MS of the ten most intense ions in the LTQ. The target ion value in the LTQ-FTICR was 500,000 for survey scan at a resolution of 50,000 at *m/z* 400. Fragmentation in the LTQ was performed by CID with a target value of 5,000 ions. Selected sequenced ions were dynamically excluded for 30 s. Critical mass spectrometric parameters were: spray voltage, 2.4 kV; no sheath and auxiliary gas flow; ion transfer tube temperature, 200 °C; normalized collision energy (35%) for MS/MS.

5.2.7. Circular Dichroism Spectrometry

The CXL cross-linked ubiquitin samples were analyzed by an Aviv Model 62A DS Circular Dichroism (CD) Spectrometer (Lakewood, NJ) at room temperature. The critical instrumental parameters included: acquisition range = 200 to 300 nm, step size = 1.00 nm, averaging time = 5 s, bandwidth = 1 nm, and path length = 1 mm. One scan was recorded for each spectrum. For

CCXL cross-linked samples, 1 to 4 µL of 50 mM CCXL stock solution in DMSO was added to 196 to 199 µL of 0.1 mg/mL ubiquitin in 1X PBS pH 7.4 to result in a final concentration of 0.25 to 1 mM CCXL in 200 µL. CD spectra were collected upon addition of CCXL (t = 0) and after 30 min of cross-linking reaction (t = 0.5 h).

5.2.8. xQuest Search

The raw files from the LTQ-FTICR mass spectrometer were converted to Mascot generic format (MGF) files using ReAdW4Mascot (version 20090305a, available from the National Institute of Standards and Technology at http://peptide.nist.gov/software/ReAdW4Mascot2_20090305a.zip), and all but the 150 most intense fragment ions were filtered out. The cross-linked peptide search was performed using xQuest (http://prottools.ethz.ch/orinner/public/htdocs/xquest).[239] The database search parameters were as follows: 0.5 Da fragment ion mass tolerance; 0.3 Da common fragment ion mass tolerance; 10 ppm precursor ion mass tolerance; trypsin enzyme specificity (up to two missed cleavages); fixed carbamidomethyl (57.02146 Da) modification of cysteine; and variable modifications of methionine oxidation (15.99491 Da), cross-linked primary amines (177.07898 Da), mono-linked dead ends (195.08954 Da), cross-linked and biotin-$(PEG)_3$-azide clicked primary amines (621.295568 Da), mono-linked and biotin-$(PEG)_3$-azide clicked primary amines (639.305584 Da), and reporter ion (525.28537 Da). Both reporter ion–filtered and non–filtered MGF files by xQuest search were tested, but identical results were reported. Reduced MS/MS scans for xQuest search resulted in faster searching for the reporter ion-filtered MGF files.

5.2.9. X-ray Crystal Structure Analysis

Human ubiquitin (PDB ID: 1UBQ. Note that the sequences of human and bovine ubiquitin are identical.) structure was analyzed by UCSF Chimera (version 1.5.3rc).[258] For each pair of cross-

linked lysine residue, the distances between the alpha carbons (C_α–C_α) and between nitrogens of ε-amines (NZ–NZ) were measured.

5.3.Results and Discussion

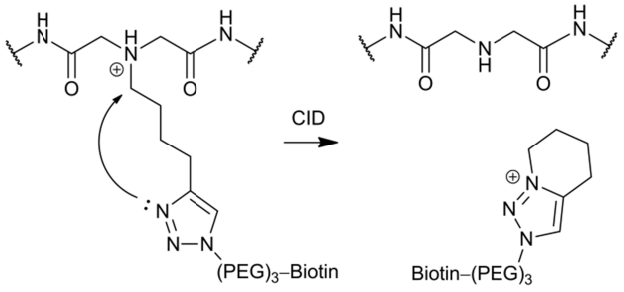
Before presenting our development to address the problems described above, it is worthy to review general aspects of chemical cross-linking strategies used in this study and structures of peptide products resulting from cross-linking. Structurally, there are three types of cross-linked peptides: intermolecular cross-linked, intramolecular cross-linked and mono-linked peptides (**Scheme 5.1**). Intermolecular cross-linked peptides (*i.e.*, cross-linking between proteins or between two residues within a protein), herein referred to as just "*cross-linked*" peptides for convenience, only exist sub-stoichiometrically compared to other *loop-linked* (*i.e.*, intramolecular modified peptides), *mono-linked* (*i.e.*, "dead-end" modified peptides) and dominantly unmodified linear peptides (**Scheme 5.1**). Among these species, only cross-linked peptides provide meaningful information for PPIs.

In this study, a novel clickable cross-linker (CXL) that incorporates an alkyne tag to enable enrichment of cross-linked peptides after coupling to an affinity tag via CuAAC is designed, synthesized, and evaluated. To demonstrate the efficiency of the new reagent, cross-linking reactions are first monitored with a model peptide, Ac-AAKAAAAAKAR (loop-linking), and the resulting cross-linked peptide is subject to CID to investigate their gas-phase fragmentation patterns. Second, cross-linking of ubiquitin as a model protein is examined by Circular Dichloism (CD) spectrometry and subject to tryptic digestion, followed by analysis of the resulting cross-linked peptides using both CID and ETD. Further evaluation of the enrichment capability of CXL is performed using CuAAC to attach biotin-$(PEG)_3$-azide, followed by biotin-avidin affinity chromatography to separate and identify cross-linked peptides from a digested yeast cell lysate.

Lastly, HEK293 cells are cross-linked *in vivo* and Western blot analysis of Cul1 is performed to monitor the cell permeability and water solubility of CXL.

5.3.1. Model Peptide Cross-Linking

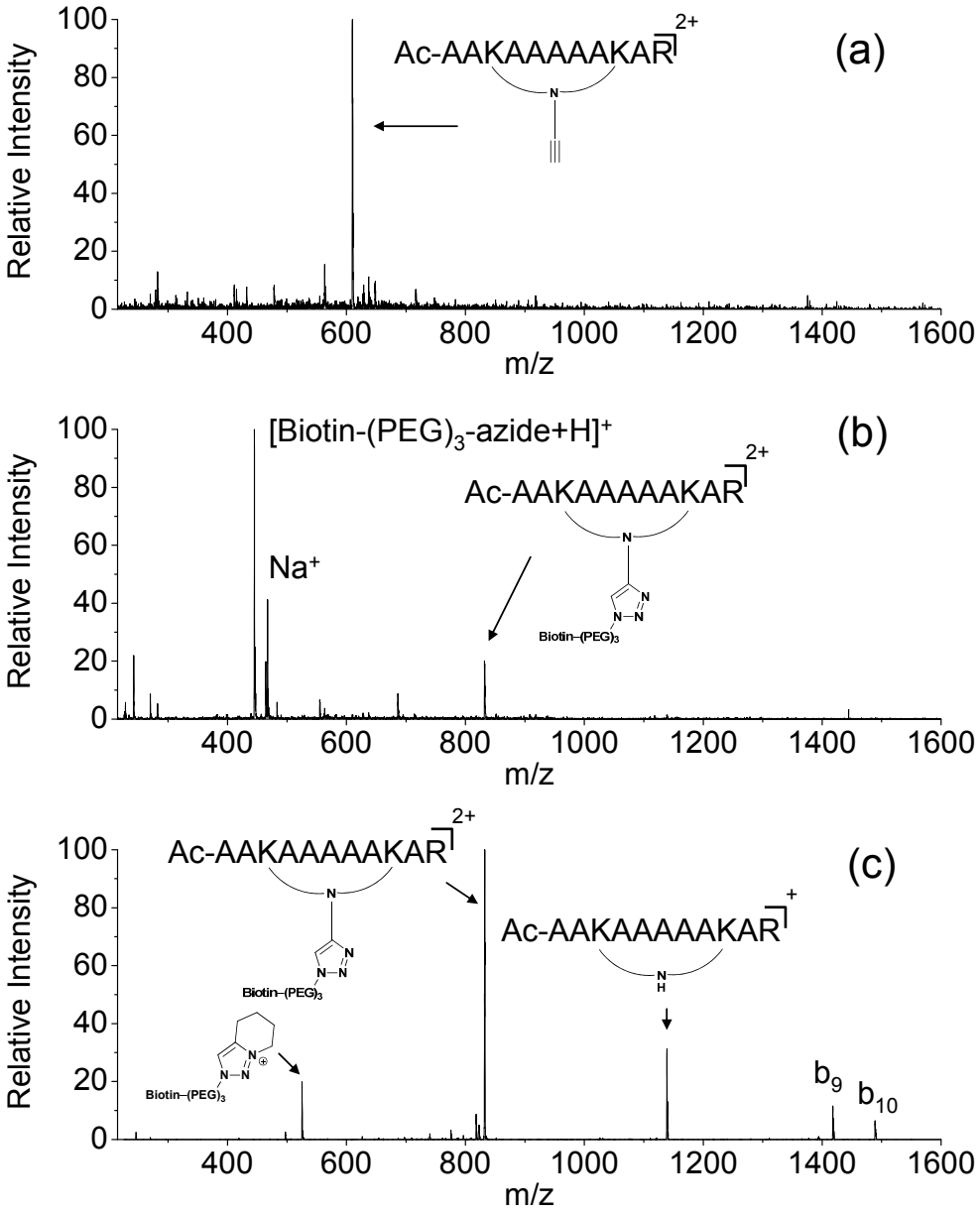
ESI-MS and CID spectra of the cross-linked model peptide Ac-AAKAAAAAKAR are shown in **Figure 5.1**. The two lysine residues in the model peptide Ac-AAKAAAAAKAR are cross-linked by CXL (*m/z* 609, **Figure 5.1a**). Protonation sites are expected to be the arginine side chain and the central tertiary amine in the cross-linker. The 1,2,3-triazole product from conjugation of biotin-$(PEG)_3$-azide via CuAAC corresponds to the doubly charged ion at *m/z* 832 in **Figure 5.1b**. No precursor ion (*m/z* 609) is observed, indicating quantitative conversion via CuAAC (**Figure 5.1b**). CID of the biotin-$(PEG)_3$-azide conjugated peptide dication yields two backbone fragments along with the reporter ion at *m/z* 525.3. This product ion is generated by a nucleophilic attack forming a six membered ring as depicted in **Scheme 5.3**.



Scheme 5.3

It is thus shown that the resulting b- and y-type ions generated by CID of cross-linked peptides allow the efficient sequencing of cross-linked peptide chains without ambiguity. In this regard, it is important that the fragmentation pathway of the CIT reporter ion occurs in competition with those of other b-and y-type ions, yielding rich sequence information.

Figure 5.1 ESI-MS and CID spectra of the cross-linked, clicked model peptide

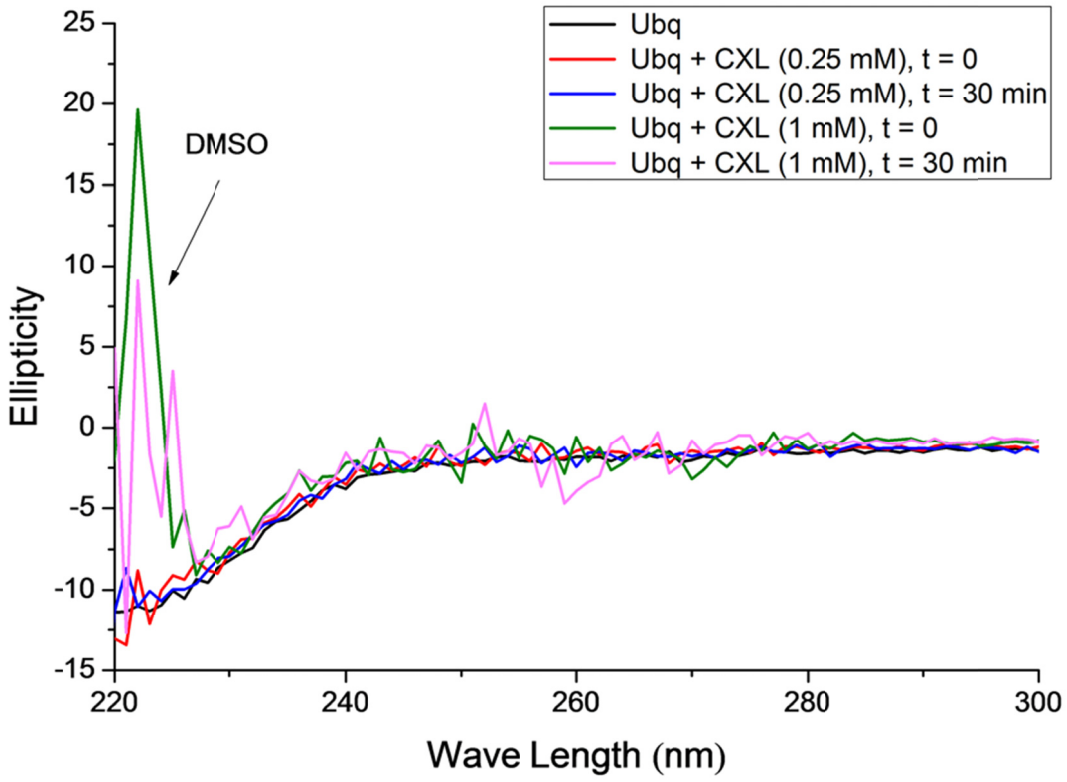


The ESI-MS spectra of the cross-linked (a), and clicked (b) Ac-AAKAAAAAKAR peptide. The CID spectrum (c) of the cross-linked and clicked Ac-AAKAAAAAKAR peptide. The reporter ion at *m/z* 525.3 is observed along with backbone fragments in the CID spectrum. Note that in b, CuAAC with biotin-$(PEG)_3$-azide proceeded almost quantitatively, showing no cross-linked precursor peptide ion at *m/z* 609.

5.3.2. Structural Analysis by Circular Dichroism

Circular dichroism (CD) spectrometry is a very useful analytical technique for quick investigation of secondary structure, folding, and binding properties of peptides and proteins.[259] Cross-linking reactions with proteins should be efficient without structural perturbation or denaturation under the working concentration of the cross-linker. To probe the structural effect, ubiquitin cross-linked by CXL is examined by CD spectrometry. The CD spectra of native ubiquitin were previously recorded, yielding the secondary structure analysis of 6% α-helix, 10% β-sheet and 84% random structures, which can be observed as increasingly negative ellipticity over the range of 225 to 240 nm.[260] In this work, the CD spectra were acquired in the presence of 0.25 or 1 mM CCXL (t = 0) and after a duration of 30 min (t = 30 min). The CD spectra of CCXL cross-linked ubiquitin samples are depicted in **Figure 5.2**, showing no significant change in various conditions. These results suggest that the secondary structure is essentially unchanged by the addition of CXL or its cross-linking reaction. Therefore, the contribution of CCXL on ubiquitin secondary structure seems to be minimal or undetectable using CD spectrometry. After this confirmation, we proceeded to analyze cross-linked peptides obtained by tryptic digestion of cross-linked ubiquitin using mass spectrometers.

Figure 5.2 CD spectra of CXL cross-linked ubiquitin



The CD spectra of ubiquitin cross-linked by various concentrations and reaction times of CXL. No significant change is observed among ubiquitin samples that are native (black), cross-linked at 0.25 mM (red) or 1 mM (green) of CXL and incubated for 30 min after the initiation of cross-linking reaction (blue for 0.25 mM and pink for 1 mM CXL). Abundant peaks from 220 to 230 nm result from the increased DMSO portion in 1 mM CXL cross-linked ubiquitin samples (from 1% to 4%). The CD signal fluctuation in 1 mM CXL experiments is mainly caused by the light scattering with the increased concentrations of small molecules such as cross-linkers and DMSO.

5.3.3. Ubiquitin Cross-Linking

Ubiquitin cross-linked peptides and their conjugates with biotin-$(PEG)_3$-azide were analyzed by a nanoLC-LTQ-FTICR mass spectrometer, and the resulting LC/MS and MS/MS profiles were searched against xQuest. All types of loop-linked, mono-linked and cross-linked peptides (**Scheme 5.1**) that are found before click reaction are summarized in **Tables 5.1** and **5.2**. Modified lysines in mono-linked peptides indicate solvent accessible residues. Loop-linked peptides are generated by cross-linking of two lysine residues that are closely positioned and where there are no unmodified lysine or arginine residues for tryptic cleavage between them. Cross-linked peptides have one or more unmodified lysine or arginine residues between cross-linked lysine residues, yielding two peptide chains covalently joined by the cross-linker. These intermolecular cross-linked peptides constrain the topology of surface exposed lysine residues, which is of particular interest for the construction of low resolution three-dimensional protein structures.

Table 5.1 Mono- and Loop-linked Peptides from Ubiquitin

Sequence	Before Click	After Click & SCX Fractionation	Reporter Ion
^{1}MQIFV^{6}K^TLTG^{11}K	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K	Yes, 2+, 3+, 4+	Yes, 3+, 4+, 250 mM	Yes
^{12}TITLEVEPSDTIENV^{27}K^A^{29}K	Yes, 3+	Yes, 3+, 250 mM	Yes
^{28}A^{29}K^IQD^{33}K	Yes, 2+	No	N/A
^{28}A^{29}K^IQDKEGIPPDQQ^{42}R	Yes, 3+	No	N/A
^{28}AKIQD^{33}K^EGIPPDQQ^{42}R	Yes, 3+	Yes, 4+, 250 mM	Yes
^{30}IQD^{33}K^EGIPPDQQ^{42}R	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
^{43}LIFAG^{48}K^QLEDG^{54}R	Yes, 2+, 3+	Yes, 2+, 3+, 4+, 50, 250 mM	Yes
^{43}LIFAG^{48}K^QLEDGRTLSDYNIQ^{63}K	Yes, 2+, 3+, 4+	Yes, 3+, 4+, 250 mM	Yes
^{55}TLSDYNIQ^{63}K^ESTLHLVL^{72}R	Yes, 2+, 3+	Yes, 3+, 4+, 250 mM	Yes
^{1}MQIFV^{6}K^TLTG^{11}K^	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
^{28}A^{29}K^IQD^{33}K^EGIPPDQQ^{42}R	Yes, 2+, 3+	Yes, 3+, 250 mM	Yes

^: linked residues. Superscript numbers are the residue numbers of amino acids of ubiquitin.

Table 5.2 Cross-linked Peptides from Ubiquitin

chain	chain	Before Click	After Click & SCX Fractionation	Reporter Ion	Avidin Enrichment	C_α distance (Å)	NZ distance (Å)
^{43}LIFAG^{48}K^QLEDG^{54}R	^{1}MQIFV^{6}K^TLTG^{11}K	Yes, 3+, 4+, 5+	No	N/A	No	15.14	17.82
^{1}MQIFV^{6}K^TLTG^{11}K	^{55}TLSDYNIQ^{63}K^ESTLHLVL^{72}R	Yes, 5+, 6+	No	N/A	No	15.01	20.31
^{30}IQD^{33}K^EGIPPDQQ^{42}R	^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K	Yes, 4+, 5+	Yes, 4+, 5+, 250 mM, 1st 500 mM, 2nd 500 mM	Yes	Yes, 4+, 5+	12.85	7.15
^{43}LIFAG^{48}K^QLEDG^{54}R	^{43}LIFAG^{48}K^QLEDG^{54}R	Yes, 3+, 4+, 5+	Yes, 5+, 2nd 500 mM	Yes	Yes, 5+	N/A	N/A
^{43}LIFAG^{48}K^QLEDG^{54}R	^{55}TLSDYNIQ^{63}K^ESTLHLVL^{72}R	Yes, 6+	No	N/A	No	17.89	19.85
^{28}A^{29}K^IQD^{33}K	^{30}IQD^{33}K^EGIPPDQQ^{42}R	No	No	Yes	Yes, 5+	6.24	9.09

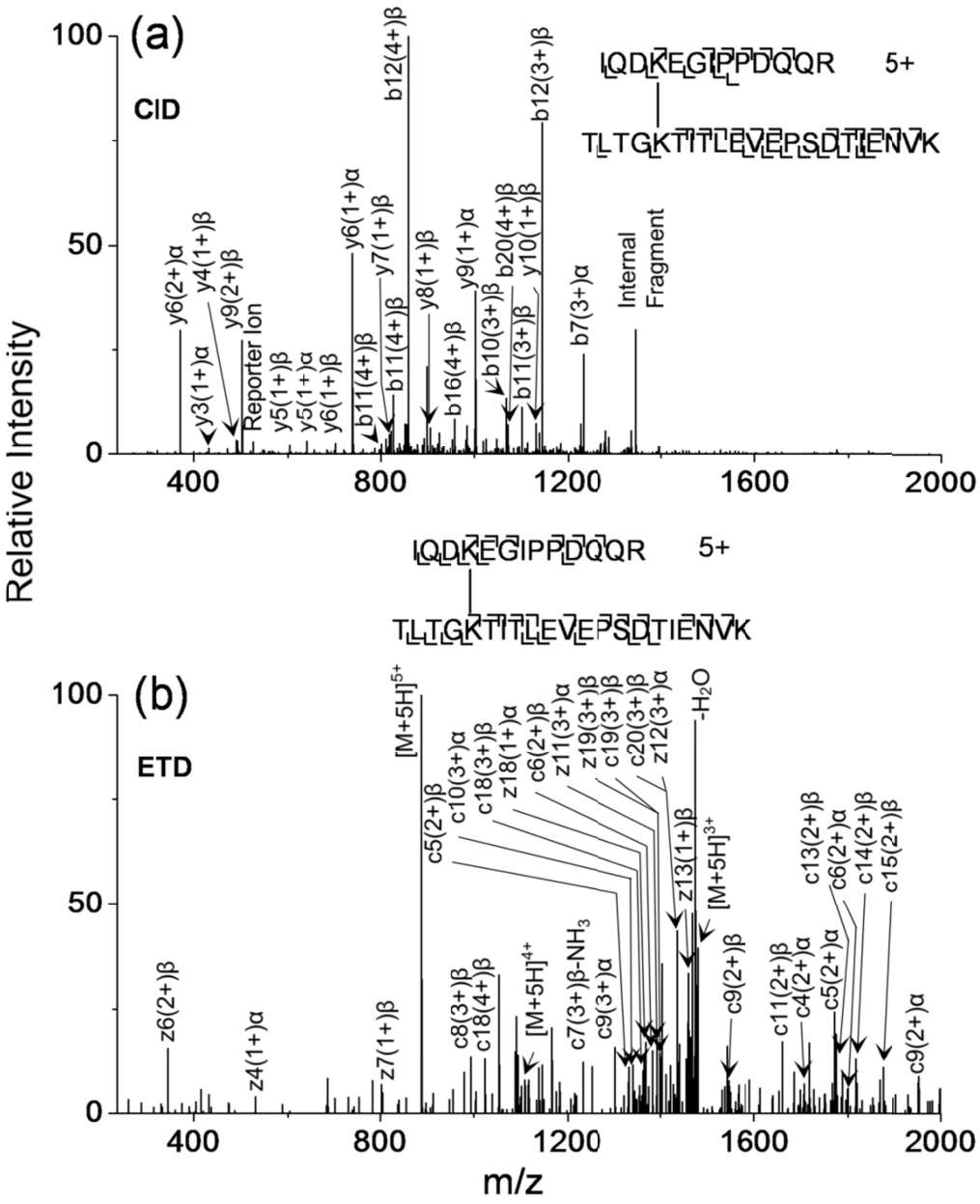
^: cross-linked residues. Superscript numbers are the residue numbers of amino acids of ubiquitin.

5.3.4. CID and ETD of Cross-Linked Peptides

Fragmentation of cross-linked peptides is investigated by CID and ETD. As an example, the CID spectrum of the 5+ charged ^{30}IQD^{33}K^EGIPPDQQ^{42}R–^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K ion is shown in **Figure 5.3**. Abundant b- and y-type ions are generated and cover many sequences in each peptide chain. The diagnostic reporter ion is found at *m/z* 525.3 without any ambiguity, and it does not overlap with other backbone fragments. ETD of the 5+ charged peptide ion also produces many c- and z-type ions due to its high charge state (**Figure 5.3**). It should be noted that the charge-reduced molecular ion species are less abundant compared to the ETD spectra of the previously reported cross-linker by Chowdhury *et al.* (Click-enabled Linker for Interacting Proteins, or CLIP).[252] With CLIP, the nitro group (NO_2) is inserted for water solubility, and neutral loss of NO_2 by CID can be used as a diagnostic peak. In ETD, the high electron affinity of the nitro group can initially trap a transferred electron, followed by proton transfer.[261] The resulting nitronic radical stabilizes the charge-reduced species and prevents further fragmentation to form c- and z-type ions, yielding abundant charge-reduced species. With CXLs, there are no occurrences of specific chemical bonds or residues that can generate unexpected neutral losses or stable charge-reduced species in ETD, so efficient electron-based dissociation processes dominate. Generally, CuAAC with biotin-$(PEG)_3$-azide increases the charge states of peptides (**Tables 5.1** and **5.2**). The tertiary amine and 1,2,3-triazole ring moderately increase the overall proton affinity for all types of cross-linked peptides. This feature apparently provides highly charged precursor ions for ETD with augmented fragmentation yields.

Note that all clicked peptides yield the reporter ion at *m/z* 525.3 by CID. Pre-screening of MS/MS scans by this particular ion provides identical searching hits as compared to the non-filtered MS/MS scan searching for every type of mono-, loop- and cross-linked peptides in *all experiments* performed in this work. Prefiltering of MS/MS scans can significantly reduce the requirement of computational resources, especially for systems level database searching, by

Figure 5.3 CID and ETD of the 5+ charged cross-linked peptide



CID (a) and ETD (b) of the 5+ charged cross-linked peptide, IQD^KEGIPPDQQR-TLTG^KTITLEVEPSDTIENVK where ^K is the cross-linked residue. (a) Although the molecular weight (4431 Da) of the cross-linked peptide is relatively large to be fragmented by CID, good coverage of b- and y-type ions is observed. The reporter ion at *m/z* 525.3 is also observed with acceptable yield. (b) The high charge state of the cross-linked peptide yields abundant c- and z-type ions in the ETD spectrum. This shows potential for the general usage of CID and ETD for efficient sequencing of cross-linked peptides.

cutting down the number of candidate MS/MS spectra. This robust diagnostic reporter ion offers a technical improvement over most of the previously reported cross-linkers and affinity enrichment schemes.

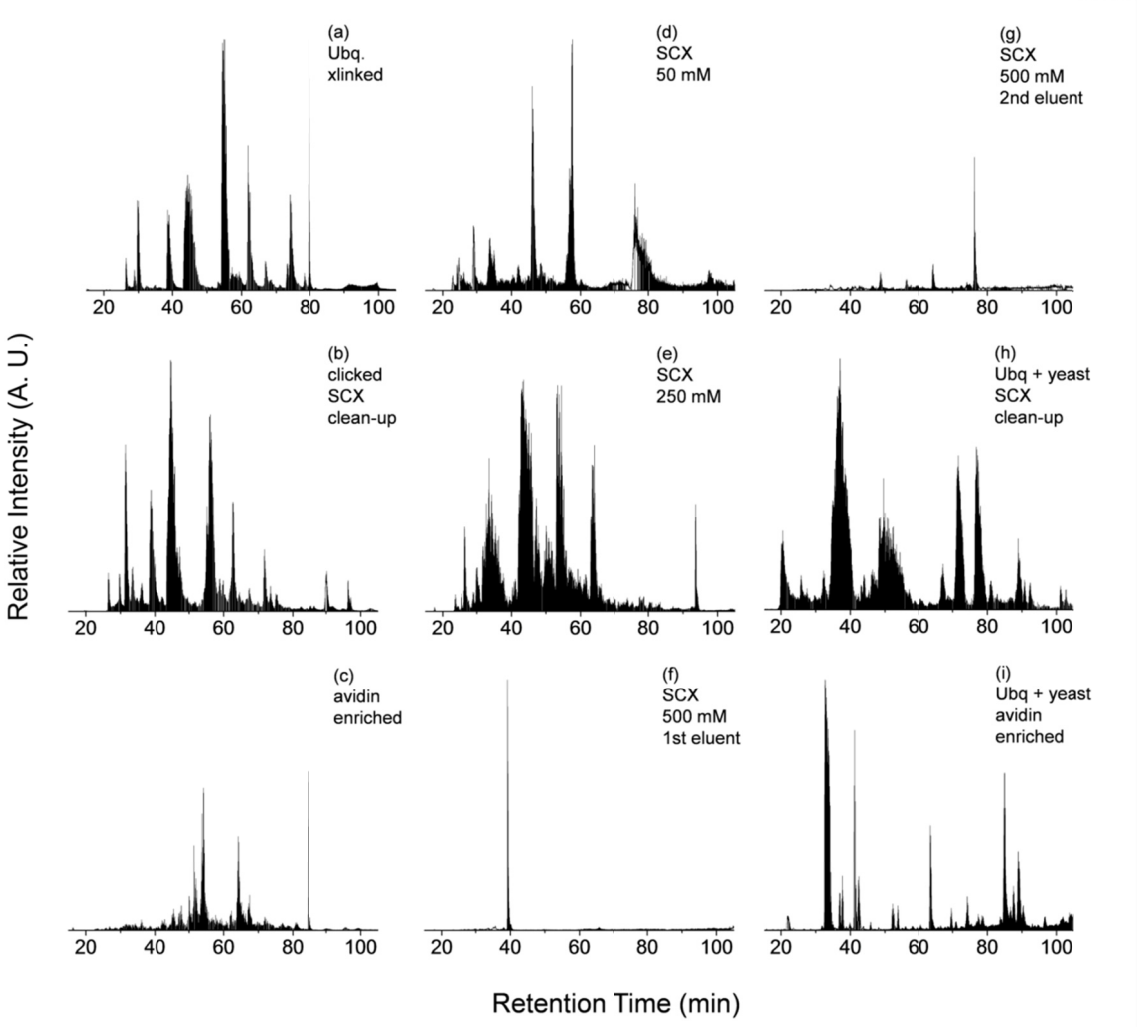
5.3.5. Sample Clean-up Following Click Reaction

CuAAC is a widely used bioconjugation reaction. However, seamless integration of CuAAC into the downstream proteomics workflows can be challenging due to the persistence of residual chemical reagents such as Cu^{2+} ions, ligands (*e.g.*, TBTA), and coupling reagents (azide or alkyne).[255] Those impurities often adversely impact the ionization efficiency of target peptides. Effective sample clean-up procedure after CuAAC is therefore very important for successful sample analysis by MS.

For this purpose, SCX, followed by C18 desalting, is the adopted clean-up procedure for the removal of non-ionic species following CuAAC. The hydrophobic TBTA reagent also has been replaced by its hydrophilic counterpart, TBTA-OH. Poor binding of TBTA-OH to the C18 matrix would result in its effective removal via C18 desalting column. Copper ions can be removed during the same desalting step.

Figure 5.4 depicts LC-MS profiles of ubiquitin cross-linked peptide samples following CuAAC. TBTA-OH is not detected in any LC-MS profile, confirming its successful removal. Excess biotin-$(PEG)_3$-azide molecules are mostly removed by SCX clean-up due to its poor interaction with SCX matrix (**Figure 5.4**), but not completely eliminated. The protonated biotin-$(PEG)_3$-azide (*m/z* 445.2) is eluted along with peptides (Retention Time around 55.17 min in **Figure 5.4b** and 46.22 min in **Figure 5.4d**), but the mass to charge ratio does not overlap with those of other cross-linked peptides, ensuring no disturbance on the sequencing. Based on its ion signal, the residual amount of biotin-$(PEG)_3$-azide after SCX clean-up seems to be less significant, and appears not to suppress peptide ionization (**Figure 5.4b**). For complete removal of the residual azide affinity tags, cleavable biotin tags can be employed along with streptavidin

Figure 5.4 LC-MS total ion current (TIC) chromatograms of cross-linked peptides



LC-MS TIC chromatograms of (a) cross-linked ubiquitin digest, (b) cross-linked, biotin-azide clicked ubiquitin digest after SCX clean-up without fractionation, (c) cross-linked, biotin-azide clicked, avidin enriched ubiquitin digest, (d) 50, (e) 250, (f) 1st 500, and (g) 2nd 500 mM NH4OAc SCX eluents of cross-linked, biotin-azide clicked ubiquitin digests, (h) biotin-azide clicked 1:1 mixture of cross-linked ubiquitin and yeast cell lysate by weight after SCX clean-up without fractionation, and (i) avidin-enriched 1:1 mixture of cross-linked ubiquitin and yeast cell lysate.

magnetic resin, which suffer less nonspecific bindings.[262] This may allow us to perform much clean elution via chemical cleavage of affinity tags.

5.3.6. Peptide Fractionation by SCX

Highly charged species tend to more strongly bind to SCX matrix. Therefore, SCX can be used for sample fractionation by discriminating the charge states of analytes. Primary sample fractionation by SCX was demonstrated for sensitive detection of cross-linked peptides from abundant linear peptides by Rinner *et al.*[239] In this report, peptide fractionation of ubiquitin cross-linked peptides was performed by sequential increase of the salt concentration during the SCX elution step. **Figure 5.4d-g** show LC-MS total ion current (TIC) chromatograms of differentially eluted ubiquitin cross-linked peptide samples by applying salt gradients, 50 mM, 250 mM, 1st 500 mM and 2nd 500 mM ammonium acetate, 0.5% FA, respectively (**Figure 5.4d-g**), or direct elution using 500 mM ammonium acetate, 0.5% FA with no fractionation (**Figure 5.4b**) during the SCX elution steps. As summarized in **Tables 5.1** and **5.2**, highly charged cross-linked peptides are eluted in the high concentration region. However, some of the cross-linked peptides are also co-eluted with other linear peptides at 250 mM salt concentration (**Table 5.2**, **Figure 5.4e**). Thus, marginal separation of cross-linked peptides is achieved by SCX fractionation. For further optimization of the separation, (an) additional elution step(s) using intermediate salt concentrations between 50 to 250 mM can be performed. Due to the low complexity of the ubiquitin cross-linked sample, SCX fractionation by itself is sufficient for separation and identification of cross-linked peptides from other linear peptides (**Figure 5.4d-g**, **Table 5.1** and **5.2**). As pointed out by Rinner et al,[239] however, the ultimate test for enrichment capability is found in the application of CXLs to higher complex systems such as *in vivo* cross-linking in eukaryotic cells.

5.3.7. Avidin Affinity Chromatography

The general design of CXLs and their capacity for bio-orthogonal incorporation of an affinity tag via CuAAC are validated by a biotin-avidin enrichment strategy. Here, monomeric avidin affinity chromatography was employed for enrichment of cross-linked peptides from the simple ubiquitin cross-linked sample prepared without SCX fractionation (**Figure 5.4c**). Cross-linked peptides enriched by avidin affinity chromatography are summarized in **Table 5.2**. Two cross-linked peptides ^{30}IQD^{33}K^EGIPPDQQ^{42}R–^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K, and the homodimer of ^{43}LIFAG^{48}K^QLEDG^{54}R that are observed in SCX fractionation are also detected by avidin affinity chromatography (**Table 5.2**), confirming the value of SCX fractionation as a separation technique. Interestingly, one additional cross-linked peptide, ^{28}A^{29}K^IQD^{33}K–^{30}IQD^{33}K^EGIPPDQQ^{42}R is identified only by avidin affinity chromatography. This result demonstrates the sensitivity of affinity-based enrichment of cross-linked peptides, maximizing their detection.

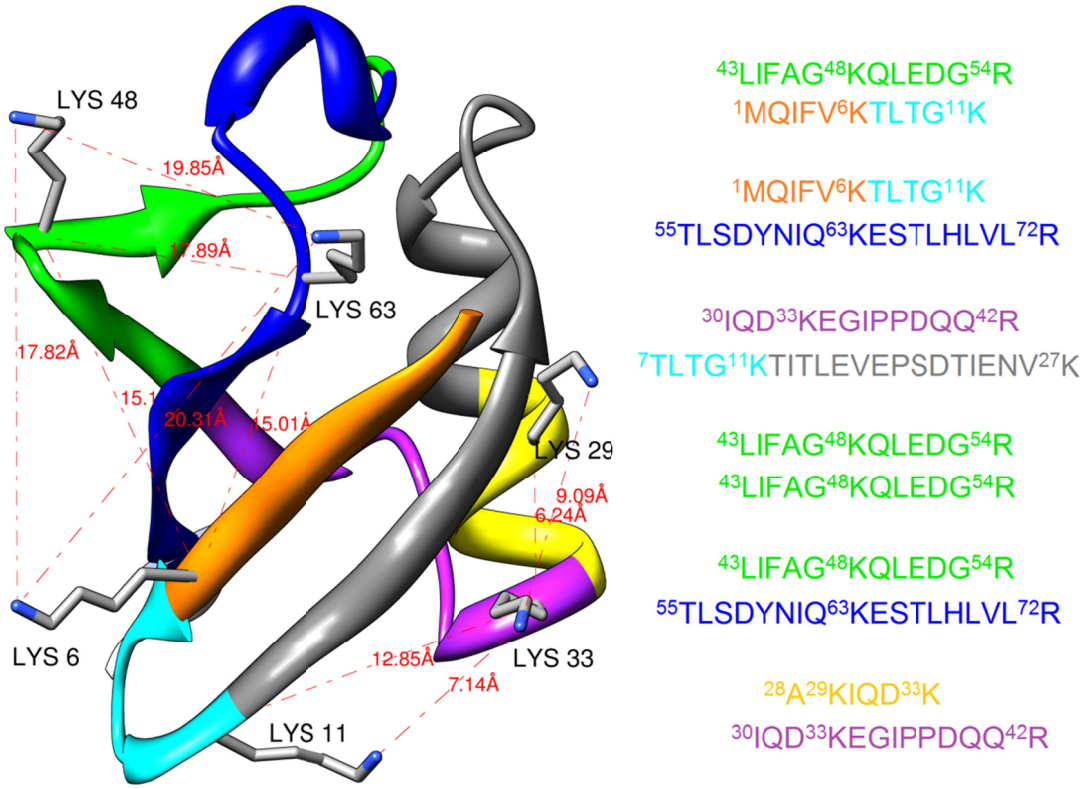
For an extreme test, a highly complex peptide sample prepared by mixing equal amounts of yeast cell lysates and ubiquitin cross-linked peptides by weight, respectively, was subjected to avidin affinity chromatography. **Figure 5.4h** and **i** depict LC-MS TIC chromatograms of the samples from SCX clean-up (**h**), and avidin enrichment (**i**) of this highly complex peptide mixture. Yeast peptides are mostly eliminated during avidin capture step, and only few are detected after enrichment. Other unmodified ubiquitin peptides that present no biotin tag are also mostly removed. The majority of the peaks in the LC-MS TIC chromatograms are singly charged impurities introduced after avidin affinity chromatography.[263] The two identified cross-linked peptides ^{30}IQD^{33}K^EGIPPDQQ^{42}R–^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K, and the homodimer of ^{43}LIFAG^{48}K^QLEDG^{54}R are reproduced those that are detected in the absence of yeast cell lysate (**Table 5.2**). This proves that the click reaction for labeling of an affinity tag is efficient, and that SCX combined with avidin affinity chromatography is an effective workflow for enrichment of cross-linked peptides even in extremely complicated environments.

5.3.8. Validation of Cross-Linked Residues

The ubiquitin lysine residues cross-linked by CXL are compared to the known X-ray crystal structure in **Figure 5.5** (PDB ID: 1UBQ). The observed lysine pairs are all located within 20 Å. This is consistent with previous reports[252,264,265] performed by other chemical cross-linkers that display a range of chain lengths similar to CXL. For example, in the ^{30}IQD^{33}K^EGIPPDQQ^{42}R–^{7}TLTG^{11}K^TITLEVEPSDTIENV^{27}K cross-linked peptide, the distance between alpha carbons in each lysine residue (K11 and K33) is 12.85 Å. The maximum length of the cross-linker in all-trans conformation is ~6.6 Å which is shorter by 6.25 Å. However, the distance between NZ atoms in the side-chains of lysine residues is only 7.15 Å. Considering the flexibility of the lysine side chain and thermal motions in proteins, the observation of cross-linking between K11 and K33 residues is reasonable. The homodimer of the ^{43}LIFAG^{48}K^QLEDG^{54}R peptide is also detected as in the previous report, indicating the formation of native ubiquitin homodimers in solution.[252]

By employing chemical reactions and subsequent sample treatments, less abundant cross-linked peptides can be lost (**Table 5.2**). Compared to missing cross-linked peptides, three cross-linked peptides detected after avidin enrichment have more closely positioned lysine pairs. Therefore, the result can be rationalized by the relationship between the spatial orientation of the lysine pairs (*i.e.*, opportunity for cross-linking) and the resulting copy number of the cross-linked peptides, and is also consistent with the crystal structure. Therefore, no further enrichment by clicking an affinity tag is necessary for the low complexity samples (*e.g.*, *in vitro* cross-linking of protein complexes that include dozens of known proteins) and SCX fractionation would be a still good choice for separation of cross-linked peptides without any sacrifice for sensitivity.

Figure 5.5 Ubiquitin X-ray crystal structure versus cross-linked peptides



Ubiquitin X-ray crystal structure (PDB: 1UBQ), presented in ribbon diagram and color coded to highlight the peptide chains that are cross-linked by CXL. Lysine residues are shown explicitly, and Cα-Cα and NZ-NZ distances (in Å) for the experimentally observed cross-linked peptides are denoted in red.

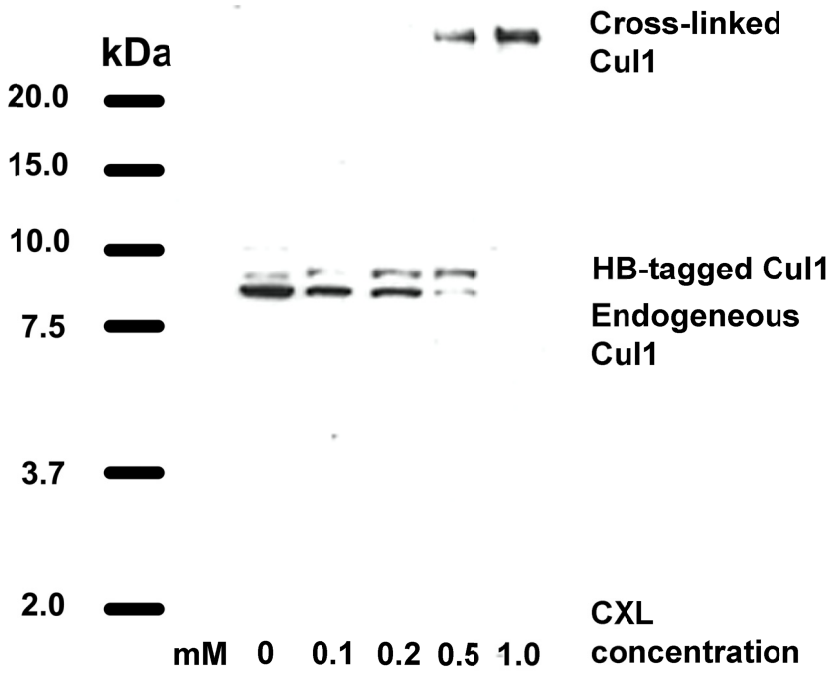
5.3.9. *In vivo* Cross-Linking of HEK 293

To test the cell permeability and water solubility of CXL, *in vivo* cross-linking of HEK 293 cells, followed by Western blot analysis of Cul1 were performed. Cul1 is a ubiquitin ligase that attaches a ubiquitin chain on target substrates for proteasome-catalyzed degradation.[211] Cul1 is a prototype of the cullin ligase family, and constitutes modular ligase complexes with other binding partners. During the design of CXL, we considered two features, the cationic nature of CXL in physiological pH and the small size, for efficient penetration of cross-linkers into cell membranes, which is a widely employed strategy in synthesis of drug delivery carriers using cationic polymers.[266] It should be firstly confirmed that CXL in working concentration ranges has no or minimal cytotoxic activity on cells. It was found that treatment of CXL up to 1 mM did not induce any significant cell toxicity and no visual change that may be caused by the entanglement of cell debris was observed. Secondly, if cytoplasmic Cul1 protein is cross-linked by CXL, the postulated cell permeability and water solubility from the design of CXL can be verified. **Figure 5.6** depicts the Western blot analysis of cross-linked Cul1 samples acquired from *in vivo* cross-linking of HEK 293 cells. From the observation of the higher molecular weight band, it is clear that Cul1 is cross-linked by CXL in 0.5~1.0 mM ranges. It also indicates that CXL is cell-permeable and soluble in PBS. In summary, we found that CXL is compatible with the cellular environment for *in vivo* cross-linking.

5.3.10. Application for Complex Systems

The encouraging results observed in this work from both *in vitro* and *in vivo* cross-linking experiments hold a promise for future application in systems level cross-linking experiments. As observed from CID and ETD of cross-linked peptides originated from ubiquitin cross-linking, tandem mass spectrometric analyses can provide useful information for sequencing of cross-linked peptides and identification of protein binding partners. Especially, the reporter ion

Figure 5.6 *In vivo* cross-linking of HEK 293 cells



In vivo cross-linking of HEK 293 and Western blot analysis. From 0 to 1.0 mM of CXL were applied to the cross-linking of Cul1. The observation of higher molecular weight bands over 20.0 kDa in 0.5 and 1.0 mM CXL lanes indicate successful cross-linking of cytoplasmic Cul1.

observed in the CID spectra is very useful for the reduction of the number of the MS/MS spectra that are subject to the data base searching. This feature would be particularly invaluable for the systems level study by saving the requirement of computational resources. The fragmentation pathway used in the formation of the reporter ion is actually universal when the 1,2,3-triazole ring is positioned through four methylene linker to the tertiary amine residue regardless of the structures of the attached affinity tags. Therefore, the highly selective reporter ion can be generated in a *mass-tunable* way by inserting various affinity tags. Cross-linkers with various chain lengths, and isotope-coded cleavable affinity tags[262] can be also prepared with the current synthetic scheme. By combining these versatile features, it is possible to explore future applications of CXLs to obtain more information about protein structure and PPIs that are targeted for more complex systems.

5.4.Conclusions

In this work, a highly versatile clickable cross-linker (CXL) that overcomes many limitations of the previous functionalized cross-linkers is developed for selective and sensitive identification of cross-linked peptides from complex mixtures. CXL is designed to have good solubility, cell permeability and an alkyne functional group for attachment of an affinity tag via CuAAC. Several advantages of CXLs have been validated using model peptides, a model protein, ubiquitin, and a real biological system using HEK293 cells. Especially, the application of CXL for analysis of complex systems is demonstrated by successful fractionation and enrichment of cross-linked peptides, and MS/MS scan filtering using the reporter ion at *m/z* 525.3. With these promising results, we are exploring the possibilities in application of CXLs in more complex systems.

5.5.Acknowledgements

This work was supported by the National Science Foundation through grant CHE-0416381 and the Beckman Institute at California Institute of Technology. Computational resources for density functional theory results were kindly provided by the Materials and Process Simulation Center at California Institute of Technology. We are also grateful to Bert Lai for technical assistance with the CD experiments. C.H.S. acknowledges a fellowship from the Kwanjeong Educational Foundation. J.E.L. was supported by the Ruth L. Kirschstein NRSA fellowship from the NIH (CA138126).

Bibliography

(1) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151.
(2) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 53.
(3) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64.
(4) Sleno, L.; Volmer, D. A. *J. Mass Spectrom.* **2004**, *39*, 1091.
(5) Cooks, R. G.; Beynon, J. H.; Caprioli, R. M.; Lester, G. R. *Metastable Ions* Amsterdam, The Netherlands, 1973.
(6) Marcus, R. A. *J. Chem. Phys.* **1952**, *20*, 359.
(7) Rosenstock, H. M.; Wallenstein, M. B.; Wahrhaftig, A. L.; Eyring, H. *Proc. Natl. Acad. Sci. U. S. A.* **1952**, *38*, 667.
(8) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
(9) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365.
(10) Qin, J.; Chait, B. T. *J. Am. Chem. Soc.* **1995**, *117*, 5411.
(11) Lee, S. W.; Kim, H. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 3188.
(12) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185.
(13) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265.
(14) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528.
(15) Hodyss, R.; Cox, H. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 12436.
(16) Ong, S.-E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252.
(17) Gstaiger, M.; Aebersold, R. *Nat. Rev. Genet.* **2009**, *10*, 617.
(18) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. *Anal. Chem.* **2003**, *75*, 1895.
(19) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154.
(20) Sinz, A.; Wang, K. *Biochemistry* **2001**, *40*, 7903.
(21) Sinz, A. *J. Mass Spectrom.* **2003**, *38*, 1225.
(22) Sinz, A.; Wang, K. *Anal. Biochem.* **2004**, *331*, 27.
(23) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25*, 663.
(24) Sinz, A. *Anal. Bioanal. Chem.* **2010**, *397*, 3433.
(25) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563.
(26) Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337.
(27) Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57.
(28) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201.
(29) Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949.
(30) O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 576.

(31) Lin, C.; O'Connor, P. B.; Cournoyer, J. J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1605.
(32) Senko, M. W.; Speir, J. P.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2801.
(33) Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* **2003**, *22*, 158.
(34) Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. *Anal. Chem.* **2000**, *72*, 552.
(35) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248.
(36) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809.
(37) Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212.
(38) Siuti, N.; Kelleher, N. L. *Nat. Methods* **2007**, *4*, 817.
(39) Kim, H. I.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2008**, *130*, 1245.
(40) Kim, H. I.; Beauchamp, J. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 157.
(41) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793.
(42) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19.
(43) Sweet, S. M. M.; Cooper, H. J. *Expert Rev. Proteomics* **2007**, *4*, 149.
(44) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2199.
(45) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431.
(46) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530.
(47) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161.
(48) Mormann, M.; Paulsen, H.; Peter-Katalinic, J. *Eur. J. Mass Spectrom.* **2005**, *11*, 497.
(49) Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J. C. *Proteomics* **2006**, *6*, 3993.
(50) Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 2901.
(51) Khidekel, N.; Ficarro, S. B.; Clark, P. M.; Bryan, M. C.; Swaney, D. L.; Rexach, J. E.; Sun, Y. E.; Coon, J. J.; Peters, E. C.; Hsieh-Wilson, L. C. *Nat. Chem. Biol.* **2007**, *3*, 339.
(52) Simon, M. D.; Chu, F. X.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. *Cell* **2007**, *128*, 1003.
(53) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857.
(54) Xia, Y.; Chrisman, P. A.; Erickson, D. E.; Liu, J.; Liang, X. R.; Londry, F. A.; Yang, M. J.; McLuckey, S. A. *Anal. Chem.* **2006**, *78*, 4146.
(55) Liang, X. R.; Xia, Y.; McLuckey, S. A. *Anal. Chem.* **2006**, *78*, 3208.
(56) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1.
(57) Hu, Q. Z.; Noll, R. J.; Li, H. Y.; Makarov, A.; Hardman, M.; Cooks, R. G. *J. Mass Spectrom.* **2005**, *40*, 430.
(58) McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J. *Anal. Chem.* **2007**, *79*, 3525.
(59) McAlister, G. C.; Berggren, W. T.; Griep-Raming, J.; Horning, S.; Makarov, A.; Phanstiel, D.; Stafford, G.; Swaney, D. L.; Syka, J. E. P.; Zabrouskov, V.; Coon, J. J. *J. Proteome Res.* **2008**, *7*, 3127.
(60) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245.
(61) Chen, X. H.; Turecek, F. *J. Am. Chem. Soc.* **2006**, *128*, 12520.

(62) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998.
(63) Jackson, S. N.; Dutta, S.; Woods, A. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 176.
(64) Breuker, K.; Oh, H. B.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14011.
(65) Patriksson, A.; Adams, C.; Kjeldsen, F.; Raber, J.; van der Spoel, D.; Zubarev, R. A. *Int. J. Mass Spectrom.* **2006**, *248*, 124.
(66) Rand, K. D.; Adams, C. M.; Zubarev, R. A.; Jorgensen, T. J. D. *J. Am. Chem. Soc.* **2008**, *130*, 1341.
(67) Hudgins, R. R.; Kleinnijenhuis, A. J.; Quinn, J. P.; Hendrickson, C. L.; Marto, J. A. In *Proceedings of the 50th ASMS Conference on Mass Spectrometry and Allied Topics* Orlando, FL, June, 2002.
(68) Iavarone, A. T.; Paech, K.; Williams, E. R. *Anal. Chem.* **2004**, *76*, 2231.
(69) Chamot-Rooke, J.; Malosse, C.; Frison, G.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2146.
(70) Xia, Y.; Gunawardena, H. P.; Erickson, D. E.; McLuckey, S. A. *J. Am. Chem. Soc.* **2007**, *129*, 12232.
(71) Hayakawa, S.; Matsubara, H.; Panja, S.; Hvelplund, P.; Nielsen, S. B.; Chen, X. H.; Turecek, F. *J. Am. Chem. Soc.* **2008**, *130*, 7645.
(72) Sawicka, A.; Skurski, P.; Hudgins, R. R.; Simons, J. *J. Phys. Chem. B* **2003**, *107*, 13505.
(73) Sobczyk, M.; Anusiewicz, W.; Berdys-Kochanska, J.; Sawicka, A.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 250.
(74) Anusiewicz, W.; Berdys-Kochanska, J.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 5801.
(75) Anusiewicz, I.; Berdys-Kochanska, J.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2006**, *110*, 1261.
(76) Sobczyk, M.; Simons, J. *J. Phys. Chem. B* **2006**, *110*, 7519.
(77) Sobczyk, M.; Simons, J. *Int. J. Mass Spectrom.* **2006**, *253*, 274.
(78) Skurski, P.; Sobczyk, M.; Jakowski, J.; Simons, J. *Int. J. Mass Spectrom.* **2007**, *265*, 197.
(79) Sobczyk, M.; Neff, D.; Simons, J. *Int. J. Mass Spectrom.* **2008**, *269*, 149.
(80) Neff, D.; Sobczyk, M.; Simons, J. *Int. J. Mass Spectrom.* **2008**, *276*, 91.
(81) Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353.
(82) Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954.
(83) Turecek, F.; Syrstad, E. A.; Seymour, J. L.; Chen, X. H.; Yao, C. X. *J. Mass Spectrom.* **2003**, *38*, 1093.
(84) Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208.
(85) Yao, C. X.; Turecek, F. *Phys. Chem. Chem. Phys.* **2005**, *7*, 912.
(86) Jones, J. W.; Sasaki, T.; Goodlett, D. R.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 432.
(87) Yao, C. X.; Syrstad, E. A.; Turecek, F. *J. Phys. Chem. A* **2007**, *111*, 4167.
(88) Hayakawa, S.; Hashimoto, M.; Matsubara, H.; Turecek, F. *J. Am. Chem. Soc.* **2007**, *129*, 7936.
(89) Turecek, F.; Chen, X. H.; Hao, C. T. *J. Am. Chem. Soc.* **2008**, *130*, 8818.
(90) Turecek, F.; Jones, J. W.; Towle, T.; Panja, S.; Nielsen, S. B.; Hvelplund, P.; Paizs, B. *J. Am. Chem. Soc.* **2008**, *130*, 14584.
(91) Vertical electron affinity of the amide π* orbital is ca. −2.5 eV. The Coulomb stabilization energy varies with distance R (Å) as 14.4eV Å/R (Å). See Seydou, M.; Modelli, A.; Lucas, B.;

Konate, K.; Desfrancois, C.; Schermann, J. P. *Eur. Phys. J.* D **2005**, *35*, 199; Zhang, H.; Lu, J. F.; Zhang, S. F.; Tang, K.; Zhou, Z. Y. *Chin. J. Struct. Chem.* **2007**, *26*, 1373.

(92) Laskin, J.; Futrell, J. H.; Chu, I. K. *J. Am. Chem. Soc.* **2007**, *129*, 9598.

(93) Frison, G.; van der Rest, G.; Turecek, F.; Besson, T.; Lemaire, J.; Maitre, P.; Chamot-Rooke, J. *J. Am. Chem. Soc.* **2008**, *130*, 14916.

(94) Prell, J. S.; O'Brien, J. T.; Holm, A. I. S.; Leib, R. D.; Donald, W. A.; Willialms, E. R. *J. Am. Chem. Soc.* **2008**, *130*, 12680.

(95) Al-Khalili, A.; Thomas, R.; Ehlerding, A.; Hellberg, F.; Geppert, W. D.; Zhaunerchyk, V.; af Ugglas, M.; Larsson, M.; Uggerud, E.; Vedde, J.; Adlhart, C.; Semaniak, J.; Kaminska, M.; Zubarev, R. A.; Kjeldsen, F.; Andersson, P. U.; Osterdahl, F.; Bednarska, V. A.; Paal, A. *J. Chem. Phys.* **2004**, *121*, 5700.

(96) Uggerud, E. *Int. J. Mass Spectrom.* **2004**, *234*, 45.

(97) Bakken, V.; Helgaker, T.; Uggerud, E. *Eur. J. Mass Spectrom.* **2004**, *10*, 625.

(98) Pouthier, V.; Tsybin, Y. O. *J. Chem. Phys.* **2008**, *129*.

(99) Li, X. J.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1514.

(100)Gunawardena, H. P.; Gorenstein, L.; Erickson, D. E.; Xia, Y.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2007**, *265*, 130.

(101)Chamot-Rooke, J.; van der Rest, G.; Dalleu, A.; Bay, S.; Lemoine, J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1405.

(102)Belyayev, M. A.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1428.

(103)Brown, R. S.; Lennon, J. J. *Anal. Chem.* **1995**, *67*, 3990.

(104)Schopfer, F. J.; Baker, P. R. S.; Freeman, B. A. *Trends Biochem.Sci.* **2003**, *28*, 646.

(105)Pacher, P.; Beckman, J. S.; Liaudet, L. *Physiol. Rev.* **2007**, *87*, 315.

(106)Han, C. C.; Balakumar, R. *Tetrahedron Lett.* **2006**, *47*, 8255.

(107)Tewari, N.; Nizar, H.; Mane, A.; George, V.; Prasad, M. *Synth. Commun.* **2006**, *36*, 1911.

(108)Adamczyk, M.; Gebler, J. C.; Wu, J. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1481.

(109)Knight, Z. A.; Schilling, B.; Row, R. H.; Kenski, D. M.; Gibson, B. W.; Shokat, K. M. *Nat. Biotechnol.* **2003**, *21*, 1047.

(110)Jalili, P. R.; Sharma, D.; Ball, H. L. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1007.

(111)Jordan, K. D.; Michejda, J. A.; Burrow, P. D. *J. Am. Chem. Soc.* **1976**, *98*, 7189.

(112)Wentworth, W. E.; Kao, L. W.; Becker, R. S. *J. Phys. Chem.* **1975**, *79*, 1161.

(113)Ziatkis, A.; Lee, C. K.; Wentworth, W. E.; Chen, E. C. M. *Anal. Chem.* **1983**, *55*, 1596.

(114)Wentworth, W. E.; Limero, T.; Chen, E. C. M. *J. Phys. Chem.* **1987**, *91*, 241.

(115)Dillow, G. W.; Kebarle, P. *J. Am. Chem. Soc.* **1989**, *111*, 5592.

(116)Chowdhury, S.; Kebarle, P. *J. Am. Chem. Soc.* **1986**, *108*, 5453.

(117)Desfrancois, C.; Periquet, V.; Lyapustina, S. A.; Lippa, T. P.; Robinson, D. W.; Bowen, K. H.; Nonaka, H.; Compton, R. N. *J. Chem. Phys.* **1999**, *111*, 4569.

(118)Fukuda, E. K.; McIver, R. T. *J. Am. Chem. Soc.* **1985**, *107*, 2291.

(119)Sanger, F. *Biochem. J.* **1945**, *39*, 507.

(120)Chen, X. H.; Anderson, V. E.; Chen, Y. H. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 448.

(121)Xie, Y. M.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432.

(122)Betowski, L. D.; Enlow, M.; Aue, D. H. *Int. J. Mass Spectrom.* **2006**, *255*, 123.

(123)Swaney, D. L.; McAlister, G. C.; Wirtala, M.; Schwartz, J. C.; Syka, J. E. P.; Coon, J. J. *Anal. Chem.* **2007**, *79*, 477.

(124)Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1990**, *62*, 1836.

(125)http://prospector.ucsf.edu
(126)Granovsky, A.A. PC GAMESS version 7.1, http://classic.chem.msu.su/gran/gamess/index.html
(127)Polasek, M.; Turecek, F. *J. Am. Chem. Soc.* **2000**, *122*, 9511.
(128)Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(129)Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(130)Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
(131)Zhao, Y.; Truhlar, D. G. *Theor. Chem. Account* **2008**, *120*, 215.
(132)Bode, B. M.; Gordon, M. S. *J. Mol. Graph.* **1998**, *16*, 133.
(133)Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *Angew. Chem.-Int. Edit.* **2006**, *45*, 5301.
(134)Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 113.
(135)Chalkley, R. J.; Brinkworth, C. S.; Burlingame, A. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1271.
(136)Shaffer, S. A.; Sadilek, M.; Turecek, F. *J. Org. Chem.* **1996**, *61*, 5234.
(137)Laskin, J.; Yang, Z. B.; Lam, C.; Chu, I. K. *Anal. Chem.* **2007**, *79*, 6607.
(138)Yu, W.; Vath, J. E.; Huberty, M. C.; Martin, S. A. *Anal. Chem.* **1993**, *65*, 3015.
(139)Sarver, A.; Scheffler, N. K.; Shetlar, M. D.; Gibson, B. W. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 439.
(140)Lee, S. J.; Lee, J. R.; Kim, Y. H.; Park, Y. S.; Park, S. I.; Park, H. S.; Kim, K. P. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2797.
(141)Salzano, A. M.; D'Ambrosio, C.; Scaloni, A. In *Nitric Oxide, Part F: Oxidative and Nitrosative Stress in Redox Regulation of Cell Signaling*; Elsevier Academic Press Inc: San Diego, 2008; Vol. 440, p 3.
(142)Frankevich, V. E.; Zhang, J.; Friess, S. D.; Dashtiev, M.; Zenobi, R. *Anal. Chem.* **2003**, *75*, 6063.
(143)Karas, M.; Gluckmann, M.; Schafer, J. *J. Mass Spectrom.* **2000**, *35*, 1.
(144)Knochenmuss, R.; Zenobi, R. *Chem. Rev.* **2003**, *103*, 441.
(145)Knochenmuss, R. *Analyst* **2006**, *131*, 966.
(146)Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413.
(147)Freiser, B. S.; Woodin, R. L.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1975**, *97*, 6893.
(148)Frazier, J. R.; Christophorou, L. G.; Carter, J. G.; Schweinler, H. C. *J. Chem. Phys.* **1978**, *69*, 3807.
(149)The energetics and MOs of the excited states for the rest of species are available in Supporting Information.
(150)Addario, V.; Guo, Y. Z.; Chu, I. K.; Ling, Y.; Ruggerio, G.; Rodriquez, C. F.; Hopkinson, A. C.; Siu, K. W. M. *Int. J. Mass Spectrom.* **2002**, *219*, 101.
(151)Sohn, C. H.; Loo, J. A.; Beauchamp, J. L. unpublished.
(152)Branden, C.; Tooze, J. *Introduction to protein structure*; 2nd ed.; Garland Pub.: New York, 1999.
(153)Creighton, T. E. *Bioessays* **1988**, *8*, 57.
(154)Betz, S. F. *Protein Sci.* **1993**, *2*, 1551.
(155)Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. *Biochemistry* **2000**, *39*, 4207.
(156)Kadokura, H.; Katzen, F.; Beckwith, J. *Annu. Rev. Biochem.* **2003**, *72*, 111.

(157)Tu, B. P.; Weissman, J. S. *J. Cell Biol.* **2004**, *164*, 341.
(158)Hogg, P. J. *Trends Biochem.Sci.* **2003**, *28*, 210.
(159)Bures, E. J.; Hui, J. O.; Young, Y.; Chow, D. T.; Katta, V.; Rohde, M. F.; Zeni, L.; Rosenfeld, R. D.; Stark, K. L.; Haniu, M. *Biochemistry* **1998**, *37*, 12172.
(160)Gorman, J. J.; Wallis, T. P.; Pitt, J. J. *Mass Spectrom. Rev.* **2002**, *21*, 183.
(161)Lioe, H.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1109.
(162)StephensonJr., J. L.; Cargile, B. J.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2040.
(163)Wells, J. M.; Stephenson, J. L.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2000**, *203*, A1.
(164)Jones, M. D.; Patterson, S. D.; Lu, H. S. *Anal. Chem.* **1998**, *70*, 136.
(165)Bilusich, D.; Bowie, J. H. *Mass Spectrom. Rev.* **2009**, *28*, 20.
(166)Mentinova, M.; Han, H. L.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2647.
(167)Zhang, M. X.; Kaltashov, I. A. *Anal. Chem.* **2006**, *78*, 4820.
(168)Lioe, H.; Duan, M.; O'Hair, R. A. J. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2727.
(169)Gunawardena, H. P.; O'Hair, R. A. J.; McLuckey, S. A. *J. Proteome Res.* **2006**, *5*, 2087.
(170)Simons, J. *Chem. Phys. Lett.* **2010**, *484*, 81.
(171)Turecek, F.; Polasek, M.; Frank, A. J.; Sadilek, M. *J. Am. Chem. Soc.* **2000**, *122*, 2361.
(172)Sawicka, A.; Berdys-Kochanska, J.; Skurski, P.; Simons, J. *Int. J. Quantum Chem.* **2005**, *102*, 838.
(173)Neff, D.; Smuczynska, S.; Simons, J. *Int. J. Mass Spectrom.* **2009**, *283*, 122.
(174)Simons, J. *J. Am. Chem. Soc.* **2010**, *132*, 7074.
(175)Lee, M.; Kang, M.; Moon, B.; Oh, H. B. *Analyst* **2009**, *134*, 1706.
(176)Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405.
(177)Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
(178)Caldwell, H.; Young, W. In *Handbook of Neurochemistry and Molecular Neurobiology*; Lajtha, A., Lim, R., Eds.; Springer U.S.: 2006, p 573.
(179)Sun, Q. Y.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. *J. Proteome Res.* **2009**, *8*, 958.
(180)Zhang, L. Y.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1378.
(181)Krenske, E. H.; Pryor, W. A.; Houk, K. N. *J. Org. Chem.* **2009**, *74*, 5356.
(182)Corey, E. J.; Hertler, W. R. *J. Am. Chem. Soc.* **1960**, *82*, 1657.
(183)Schöneich, C. *Chem. Res. Toxicol.* **2008**, *21*, 1175.
(184)Mozziconacci, O.; Williams, T. D.; Kerwin, B. A.; Schöneich, C. *J. Phys. Chem. B* **2008**, *112*, 15921.
(185)The thiyl radical products are not compared because those easily undergo subsequent neutral loss of thioaldehydes.
(186)Izgorodina, E. I.; Brittain, D. R. B.; Hodgson, J. L.; Krenske, E. H.; Lin, C. Y.; Namazian, M.; Coote, M. L. *J. Phys. Chem. A* **2007**, *111*, 10754.
(187)Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2008**, *112*, 1095.
(188)Houée-levin, C. In *Methods in Enzymology*; Chandan K. Sen, L. P., Ed.; Academic Press: 2002; Vol. Volume 353, p 35.
(189)Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198.
(190)Service, R. F. *Science* **2008**, *321*, 1758.
(191)de Godoy, L. M. F.; Olsen, J. V.; Cox, J.; Nielsen, M. L.; Hubner, N. C.; Frohlich, F.; Walther, T. C.; Mann, M. *Nature* **2008**, *455*, 1251.
(192)Schulze, W. X.; Usadel, B. *Annu. Rev. Plant Biol.* **2010**, *61*, 491.

(193)Griffin, N. M.; Yu, J.; Long, F.; Oh, P.; Shore, S.; Li, Y.; Koziol, J. A.; Schnitzer, J. E. *Nat Biotech* **2010**, *28*, 83.
(194)Clough, T.; Key, M.; Ott, I.; Ragg, S.; Schadow, G.; Vitek, O. *J. Proteome Res.* **2009**, *8*, 5275.
(195)Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. *Mol. Syst. Biol.* **2008**, *4*, 222.
(196)Picotti, P.; Bodenmiller, B.; Mueller, L. N.; Domon, B.; Aebersold, R. *Cell* **2009**, *138*, 795.
(197)Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R. *Nat. Methods* **2010**, *7*, 43.
(198)Picotti, P.; Aebersold, R.; Domon, B. *Mol. Cell. Proteomics* **2007**, *6*, 1589.
(199)Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994.
(200)Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376.
(201)Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J.-C. *Anal. Chem.* **2008**, *80*, 2921.
(202)Choe, L.; D'Ascenzo, M.; Relkin, N. R.; Pappin, D.; Ross, P.; Williamson, B.; Guertin, S.; Pribil, P.; Lee, K. H. *Proteomics* **2007**, *7*, 3651.
(203)Treumann, A.; Thiede, B. *Expert Rev. Proteomics* **2010**, *7*, 647.
(204)Ralhan, R.; DeSouza, L. V.; Matta, A.; Chandra Tripathi, S.; Ghanny, S.; DattaGupta, S.; Thakar, A.; Chauhan, S. S.; Siu, K. W. M. *J. Proteome Res.* **2008**, *8*, 300.
(205)Seo, J.; Suh, M. S.; Thangadurai, T. D.; Kim, J.; Rhee, Y. H.; Yoon, H. J.; Shin, S. K. *Anal. Chem.* **2008**, *80*, 6145.
(206)Zeng, D.; Li, S. *Chem. Commum.* **2009**, 3369.
(207)Xiang, F.; Ye, H.; Chen, R.; Fu, Q.; Li, L. *Anal. Chem.* **2010**, *82*, 2817.
(208)Zhang, J.; Wang, Y.; Li, S. *Anal. Chem.* **2010**, *82*, 7588.
(209)Alvarez, S. G.; Alvarez, M. T. *Synthesis* **1997**, 413.
(210)Kacprzak, K. *Synlett* **2005**, 943.
(211)Deshaies, R. J. *Annu. Rev. Cell Dev. Biol.* **1999**, *15*, 435.
(212)Lee, J. E.; Sweredoski, M. J.; Graham, R. L. J.; Kolawa, N. J.; Smith, G. T.; Hess, S.; Deshaies, R. J. *Mol. Cell. Proteomics* **2010**, *Electronic preprint. doi:10.1074/mcp.M110.006460*.
(213)Tagwerker, C.; Flick, K.; Cui, M.; Guerrero, C.; Dou, Y. M.; Auer, B.; Baldi, P.; Huang, L.; Kaiser, P. *Mol. Cell. Proteomics* **2006**, *5*, 737.
(214)Cronan, J. E. *J. Biol. Chem.* **1990**, *265*, 10327.
(215)Kulman, J. D.; Satake, M.; Harris, J. E. *Protein Expr. Purif.* **2007**, *52*, 320.
(216)Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T.; Sweetman, G.; Kuster, B. *Mol. Cell. Proteomics* **2008**, *7*, 1702.
(217)Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26*, 1367.
(218)Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.-Int. Edit.* **2002**, *41*, 2596.
(219)Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
(220)Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16793.
(221)Note that for the PQD experiment, light- and heavy-labeled peptides are prepared and purified separately by C18 ziptips, and later mixed for linear response testing with various mixing ratios. In the actual quantification experiment, the sample mixing step is prior to the C18 purification to avoid dissimilar sample loss. In the beam-type CID experiments, the samples were mixed before the C18 purification, showing no significant experimental error. This source of

initial mixing error owing to differential sample loss during the C18 purification appears as the y-axis intercepts in the linear fitting lines in Figure 4.7.

(222)Zhang, R. J.; Sioma, C. S.; Wang, S. H.; Regnier, F. E. *Anal. Chem.* **2001**, *73*, 5142.

(223)Yi, E. C.; Li, X.-j.; Cooke, K.; Lee, H.; Raught, B.; Page, A.; Aneliunas, V.; Hieter, P.; Goodlett, D. R.; Aebersold, R. *Proteomics* **2005**, *5*, 380.

(224)Kocher, T.; Pichler, P.; Schutzbier, M.; Stingl, C.; Kaul, A.; Teucher, N.; Hasenfuss, G.; Penninger, J. M.; Mechtler, K. *J. Proteome Res.* **2009**, *8*, 4743.

(225)Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. *Nature* **1999**, *402*, C47.

(226)Pereira-Leal, J. B.; Levy, E. D.; Teichmann, S. A. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2006**, *361*, 507.

(227)Sharan, R.; Suthram, S.; Kelley, R. M.; Kuhn, T.; McCuine, S.; Uetz, P.; Sittler, T.; Karp, R. M.; Ideker, T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1974.

(228)King, R. W.; Deshaies, R. J.; Peters, J.-M.; Kirschner, M. W. *Science* **1996**, *274*, 1652.

(229)Wong, S. S. *Chemistry of Protein Conjugation and Cross-Linking*; CRC Press, 1991.

(230)Phizicky, E. M.; Fields, S. *Microbiol. Rev.* **1995**, *59*, 94.

(231)Back, J. W.; de Jong, L.; Muijsers, A. O.; de Koster, C. G. *J. Mol. Biol.* **2003**, *331*, 303.

(232)Gingras, A. C.; Gstaiger, M.; Raught, B.; Aebersold, R. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 645.

(233)Lee, Y. J. *Molecular BioSystems* **2008**, *4*, 816.

(234)Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. *Mol. Cell. Proteomics* **2010**, *9*, 1634.

(235)Petrotchenko, E. V.; Borchers, C. H. *Mass Spectrom. Rev.* **2010**, *29*, 862.

(236)Collins, M. O.; Choudhary, J. S. *Curr. Opin. Biotechnol.* **2008**, *19*, 324.

(237)Guerrero, C.; Milenkovic, T.; Przulj, N.; Kaiser, P.; Huang, L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 13333.

(238)Zhang, H.; Tang, X.; Munske, G. R.; Tolic, N.; Anderson, G. A.; Bruce, J. E. *Mol. Cell. Proteomics* **2009**, *8*, 409.

(239)Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller, L.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. *Nat. Methods* **2008**, *5*, 315.

(240)Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyo, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125*, 2416.

(241)Tang, X. T.; Munske, G. R.; Siems, W. F.; Bruce, J. E. *Anal. Chem.* **2005**, *77*, 311.

(242)Kang, S.; Mou, L. Y.; Lanman, J.; Velu, S.; Brouillette, W. J.; Prevelige, P. E. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1719.

(243)Chu, F. X.; Mahrus, S.; Craik, C. S.; Burlingame, A. L. *J. Am. Chem. Soc.* **2006**, *128*, 10362.

(244)Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O. *Anal. Chem.* **2001**, *73*, 1927.

(245)Collins, C. J.; Schilling, B.; Young, M. L.; Dollinger, G.; Guy, R. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4023.

(246)Petrotchenko, E. V.; Olkhovik, V. K.; Borchers, C. H. *Mol. Cell. Proteomics* **2005**, *4*, 1167.

(247)Wine, R. N.; Dial, J. M.; Tomer, K. B.; Borchers, C. H. *Anal. Chem.* **2002**, *74*, 1939.

(248)Back, J. W.; Hartog, A. F.; Dekker, H. L.; Muijsers, A. O.; de Koning, L. J.; de Jong, L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 222.

(249)Lauber, M. A.; Reilly, J. P. *Anal. Chem.* **2010**, *82*, 7736.

(250)Gardner, M. W.; Vasicek, L. A.; Shabbir, S.; Anslyn, E. V.; Brodbelt, J. S. *Anal. Chem.* **2008**, *80*, 4807.

(251)Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007.
(252)Chowdhury, S. M.; Du, X.; Tolić, N.; Wu, S.; Moore, R. J.; Mayer, M. U.; Smith, R. D.; Adkins, J. N. *Anal. Chem.* **2009**, *81*, 5524.
(253)Trnka, M. J.; Burlingame, A. L. *Mol. Cell. Proteomics* **2010**, *9*, 2306.
(254)Nessen, M. A.; Kramer, G.; Back, J.; Baskin, J. M.; Smeenk, L. E. J.; de Koning, L. J.; van Maarseveen, J. H.; de Jong, L.; Bertozzi, C. R.; Hiemstra, H.; de Koster, C. G. *J. Proteome Res.* **2009**, *8*, 3702.
(255)Vellucci, D.; Kao, A.; Kaake, R. M.; Rychnovsky, S. D.; Huang, L. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1432.
(256)Nilsson, B. L.; Overman, L. E.; Read de Alaniz, J.; Rohde, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 11297.
(257)Rozkiewicz, D. I.; Janczewski, D.; Verboom, W.; Ravoo, B. J.; Reinhoudt, D. N. *Angew. Chem.-Int. Edit.* **2006**, *45*, 5292.
(258)Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
(259)Greenfield, N. J. *Nat. Protocols* **2007**, *1*, 2876.
(260)Jenson, J.; Goldstein, G.; Breslow, E. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1980**, *624*, 378.
(261)Sohn, C. H.; Chung, C. K.; Yin, S.; Ramachandran, P.; Loo, J. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2009**, *131*, 5444.
(262)Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. *J. Am. Chem. Soc.* **2010**, *132*, 18351.
(263)To avoid such unexpected impurities, streptavidin magnetic bead with cleavable biotin azide tags under MS friendly mild conditions can be employed.
(264)Kruppa, G. H.; Schoeniger, J.; Young, M. M. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155.
(265)Novak, P.; Young, M. M.; Schoeniger, J. S.; Kruppa, G. H. *Eur. J. Mass Spectrom.* **2003**, *9*, 623.
(266)Nam, Y. S.; Kang, H. S.; Park, J. Y.; Park, T. G.; Han, S. H.; Chang, I. S. *Biomaterials* **2003**, *24*, 2053.