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

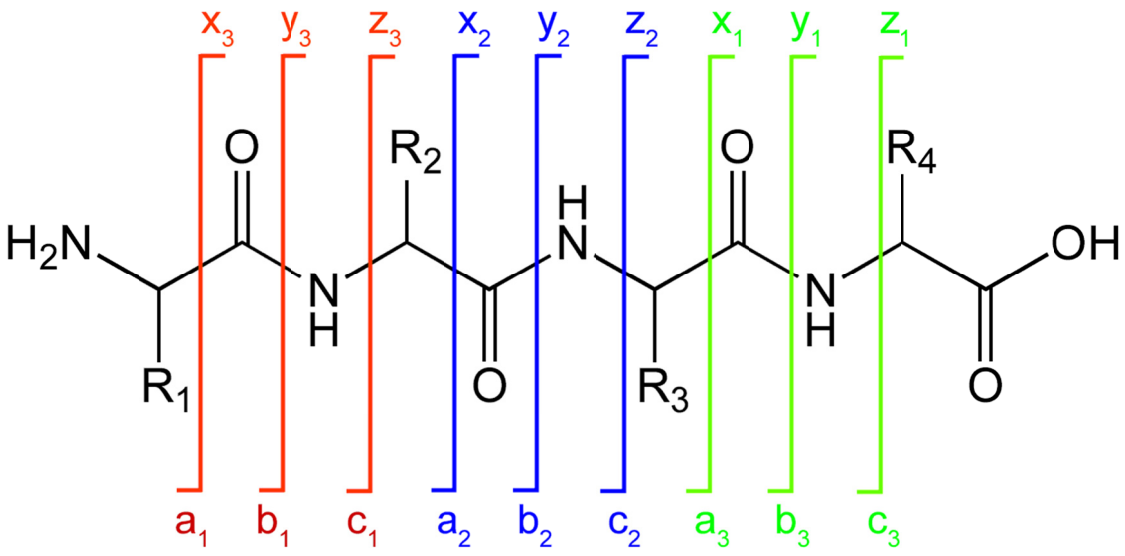
Development of two soft ionization methods, matrix-assisted laser desorption/ionization (MALDI)[1,2] and electrospray ionization (ESI)[3] has opened up the era of mass spectrometry (MS)-based proteomics. Using these soft ionization techniques, molecular weights of biopolymers are easily determined and 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.