Title
Advancing Academics Through Laboratory and Classroom Research: Probing Biomolecular Structure with Novel Mass Spectrometry Methods and Revisiting Student Learning Tools

Permalink
https://escholarship.org/uc/item/39r029dg

Author
Talbert, Lance Edward

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Advancing Academics Through Laboratory and Classroom Research: Probing
Biomolecular Structure with Novel Mass Spectrometry Methods and Revisiting Student
Learning Tools

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Lance E. Talbert

June 2019

Dissertation Committee:
      Dr. Ryan R. Julian, Chairperson
      Dr. Yinsheng Wang
      Dr. Leonard Mueller
      Dr. Jack Eichler

The Dissertation of Lance E. Talbert is approved:

_____

_____

_____

_____
                                    Committee Chairperson

University of California, Riverside

The text of this dissertation, in part or in full, is a reprint of the materials as they appear in the following publications:

1. L.E. Talbert, R.R. Julian, Directed-Backbone Dissociation Following Bond-Specific Carbon-Sulfur UVPD at 213 nm, J. Am. Soc. Mass Spectrom., (2018).
2. L.E. Talbert, X. Zhang, N. Hendricks, A. Alizadeh, R.R. Julian, Synthesis of new SS and CC bonds by photoinitiated radical recombination reactions in the gas phase, Int. J. Mass Spectrom., 441 (2019) 25–31.

ABSTRACT OF THE DISSERTATION

Advancing Academics Through Laboratory and Classroom Research: Probing
Biomolecular Structure with Novel Mass Spectrometry Methods and Revisiting Student
Learning Tools

by

Lance E. Talbert

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2019
Dr. Ryan R. Julian, Chairperson

The development of mass spectrometry (MS) as a tool for the characterization of
biological molecules has seen rapid growth over the past three decades. The structure of
a peptide or protein is key to its function and the role it plays within a biological
environment. Several tools are available for determination of primary structure, however
higher order structural characterization continues to be more challenging. As such, the
development of new analytical methodology for the structural characterization of peptides
and proteins is of significant interest.

Wavelength selection is key to modulate fragmentation and biomolecule
characterization when combining spectroscopy and MS. By coupling 213nm ultraviolet
photodissociation (UVPD) with MS, both bond-specific dissociation and traditional
nonspecific UVPD are observed. 213nm UVPD showed enhanced Carbon-Sulfur bond
dissociation, leading to an investigation of their potential as energy acceptors in an action
excitation energy transfer system. 266nm excitation of synthetic peptides containing

methionine and a native aromatic donor revealed low energy transfer efficiency, leading to an investigation of a common methionine analogue: selenomethionine. Examination of C-Se bonds within synthetic peptides revealed enhanced energy transfer efficiency. This indicates selenomethionine may prove useful for probing protein structure in the gas phase. Photoinitiated radical chemistry has proven useful for breaking covalent bonds but may also have a role in a different application: bond synthesis in the gas phase. To show this, 266nm photoactivation of peptides, peptide pairs, and peptide-noncovalent complexes that contained either S-S or C-I bonds create sulfur- and carbon-centered radicals. Following radical attack or radical migration, the formation of new S-S or C-C bonds in the gas phase.

The development and implementation of education tools, to improve student learning gains, is a key area of research in chemical education. A quarter-long concept mapping exercise was used in an effort to improve student conceptual understanding. Students participating in the quasi-experiment showed higher self-reported learning gains, and those that performed better on the concept mapping activity scored higher on concept inventory questions. Collectively, this work demonstrates that chemistry in the laboratory is just as important as chemistry in the classroom, and advances in both will lead to better scientific innovation.

Table of Contents

List of Figures

List of Tables

List of Schemes

xvi

List of Equations

Chapter 1: Introduction

1.1 Peptides and Proteins: Sequence and Structure

    Proteins continue to be one of the focal points of research in the study of biological systems. They are a major structural component of cells and organelles involved in many functions, ranging from catalyzing reactions to operating as transport channels for small ions. These macromolecules are made up of combinations of twenty amino acids, leading to many different combinations and an infinite amount of possible proteins. Biomolecules are characterized by their primary, secondary, tertiary, and quaternary structures. The primary structure refers to the specific connectivity of the amino acids, which is unique to each protein. Secondary structure is defined by the formation of alpha helices and beta sheets, with the tertiary structure being the specific interactions of such secondary motifs, or in other words, the three-dimensional protein's shape. Quaternary structure alludes to how a protein interacts with other proteins and the number of other protein subunits that come together to make a protein complex. The tasks that proteins perform are based on each of these structural classifications. This could include the arrangement of amino acids to form a particular catalytic domain, like those found in enzymes, or how the structure of the protein allows interaction with other proteins or metabolites to enable some downstream biological function. In fact, proteins play a role in all processes in living organisms, making characterization of their structure a necessity to understanding how they work. Depending on their environment, proteins can adopt many different structures. This leads to one important question: how can these structures be elucidated in order to gain an understanding of their function?

1.2 Condensed Phase Methods for Peptide and Protein Characterization

X-ray crystallography has been around since the early 1900's and is often considered the "gold standard" for the characterization of small molecules to larger biological molecules. This technique continues to produce groundbreaking structural insights and, to date, there have been fourteen different Nobel Prizes awarded to the development and application of X-ray crystallography.[1] For instance, the double helix nature of DNA was postulated by James Watson and Francis Crick through X-ray fiber diffraction images of DNA molecules.[2] As proteins are essential for life, it was only a matter of time before the structure of common proteins was ascertained. In the 1960's, X-ray crystallography was utilized for the characterization of various abundant enzymes and proteins in human tissue and animal organs, including: hemoglobin, myoglobin, egg white lysozyme, chymotrypsin, and ribonuclease A.[3-7] It is evident that this is an extremely powerful technique for studying biological molecules; however, there are several caveats that pose significant drawbacks and limitations. The most obvious issue is the requirement that the protein in question can crystalize out of solution and whether the obtained protein will be homogenous. A second issue is the requirement of high concentrations of protein crystal samples for study, which can be quite difficult to obtain if the protein in question is not found abundantly in vivo. Finally, proteins have dynamic structures and are not found as rigid crystals, which raises the question of how representative these crystal structures are of the native protein.

Nuclear magnetic resonance (NMR) is another common condensed phase method which can overcome some of the shortcomings of X-ray crystallography by allowing for the characterization of proteins in solution.[8,9] NMR measures the magnetic moment of how the nuclei of atoms interact with an external magnetic field. The "spin" of these nuclei produces chemical shifts specific to the interactions of the nucleus within its local

environment, allowing for NMR to capture the dynamic behavior of a protein. Using computer software to deconvolute the data, these spectra can be interpreted to determine the secondary and tertiary structures of the protein. This makes NMR a powerful tool for characterizing protein structure, as dynamic behavior is essential for describing how the structure plays a role in its biological function. However, like X-ray crystallography, some issues exist with NMR, namely, the requirement of an ample amount of pure sample for analysis. Prior to analysis, the proteins must be well purified, as any contamination will introduce additional chemical signals, which can complicate obtained NMR spectra. NMR also has lower resolution than X-ray crystallography, which limits the size of the proteins which can be analyzed.

Cryo-electron microscopy (cryo-EM) is a relative newcomer to the field of biological structure determination and was developed in the 1970s.[10] Many of the previous issues identified with NMR and X-ray crystallography are not seen with cryo-EM, as near-atomic resolution can be obtained without the requirement of large amounts of pure protein sample or the need for protein crystals. Cryo-EM excels at analyzing large and even massive structures as big as 200 kDa at a resolution of 3 Å.[10] The difficulty with cryo-EM is being able to accurately characterize smaller proteins. Recent advances with charge-coupled devices (CCDs) have improved the limits of detection, enabling characterization of proteins as small as ~60 kDa.[11] While this technique is promising for the future of structural determination of large biomolecules, there are still many challenges facing the field, including sample preparation, low resolution of the images, and high levels of noise.

1.3 Mass Spectrometry and Coupled Mass Spectrometry Approaches

With the limitations established for the aforementioned methods, interest has shifted to mass spectrometry (MS) as a potential solution. In its early years, hard ionization

techniques, such as electron ionization, were the main approach, particularly for the study of organic molecules. In order to study biological systems, a softer ionization technique was desired to transfer biomolecules from the condensed phase to the gas phase. Over the last thirty years, advances in the development of soft ionization techniques, like electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), have led to increased analysis of biomolecular systems with MS. On the surface, MS offers accurate molecular weight determination of molecules as well as structural characterization. The benefits of using MS for the study of biological molecules include: rapid sample analysis, less complicated sample preparation, and femtomolar sensitivity. MS can also be coupled with popular techniques, like liquid chromatography, which provides ample incentive to incorporate MS into experimental procedures for biomolecular structure determination. Some other examples of MS-coupled techniques include chemical cross-linking, hydrogen-deuterium exchange, and ion mobility.

Chemical cross-linking and hydrogen-deuterium exchange (HDX) allow for the study of three-dimensional protein structure and protein-protein interactions. The focus of HDX is on the amide-hydrogens and the rate in which they exchange in deuterated solvent. Analysis by MS utilizes the difference in mass between hydrogen and deuterium. Protein structure can be obtained by monitoring specific amide hydrogen exchanges along the peptide backbone. Because HDX requires solvent accessible hydrogens, it is possible to differentiate between residues buried in the globular structure and residues exposed on the surface.[12-14] In chemical cross-linking, a protein is covalently modified with a cross linking agent. Primary amines are the most commonly targeted sites for covalent modification as they are found in the lysine side chain.[15,16] The protein can be subsequently digested and analyzed via MS. The digested protein will contain a number

of peptides that are covalently linked by the modifier, providing some insight into which residues were solvent accessible and close in space in the liquid phase structure. Both of these techniques have great compatibility with MS. However, a significant amount of data is produced in either experiment. This creates a bottleneck between data analysis and protein structure determination, making the use of these techniques time-consuming.

Ion mobility differs from the previous techniques in that no modification of the protein is required prior to analysis by mass spectrometry. In ion mobility, ions are separated within a gas-filled drift tube by their collision cross sections (CCSs). Collisions of the neutral gas with the protein are dependent on its conformation as it travels through the drift tube. The number of collisions that occur for a given conformation is proportional to its drift time, which allows for correlation between drift time and the proteins' CCS. Ion mobility has shown significant utility in the analysis of protein complexes, as changes in the number of interacting subunits leads to drastic shifts in CCS.[17,18] While the utility of ion mobility mass-spectrometry (IM-MS) for structural elucidation has been well demonstrated,[19,20] there are a number of limitations in its use. The obtained CCS is a single data point which is representative of the entire structure of the biomolecule. This CCS is often similar or identical to multiple conformations for the given biomolecule, which complicates data analysis. Subsequent molecular modeling may lead to multiple structures which share similar CCSs but vary widely in their overall structure. Recent advances have also shown that by correlating the CCS to activation energy, it is possible to obtain more detailed substructural information.[21] As such, by coupling ion mobility with other techniques, further insight into the protein's structure can be obtained.

1.4 Ion Activation Methods for Biomolecule Structural Characterization

For proteomics, "bottom-up" and "top-down" methodologies are the most commonly utilized approaches for protein structural characterization in MS experiments. The bottom-up approach utilizes chemical or enzymatic digestion of the protein followed by MS analysis. In top-down proteomics, proteins are introduced into the mass spectrometer without any prior liquid phase preparation. Different activation methods, including collision induced dissociation (CID), electron transfer dissociation (ETD), electron capture dissociation (ECD), and ultraviolet photodissociation (UVPD), can be used to fragment the protein. The top-down and bottom-up approaches are often complimentary to each other and, when utilized in tandem, offer the greatest characterization potential.

Each of the aforementioned techniques results in different accessible fragmentation pathways and fragment ions, which are identified by their m/z ratio for characterization. Figure 1.1 shows a reaction coordinate diagram highlighting some of the different activation techniques. An activation technique can impart energy in several low energy steps, as is the case for infrared multiphoton dissociation (IRMPD) or CID, or a single high energy step, as is seen with UVPD. The difference in energy deposition on to the biomolecule allows for variation in the types of fragment ions produced as well as their abundance. The fragmentation of the peptide or protein backbone is dependent not only on the energy introduced from the activation technique but also the primary sequence and charge state. Figure 1.2 shows the location of each bond cleavage along a peptide backbone and its corresponding fragment ion produced. Cleavage of the C-C, C-N, and N-C_α leads to the formation of a/x, b/y, and c/z ion pairs respectively. Fragment ions containing the N-terminus of the peptide are labeled as a, b, or c, while the inclusion of the C-terminus in the fragment ion leads to a label of x, y, or z. The following paragraphs

will briefly cover the most common techniques utilized for structural characterization of

peptides and proteins.
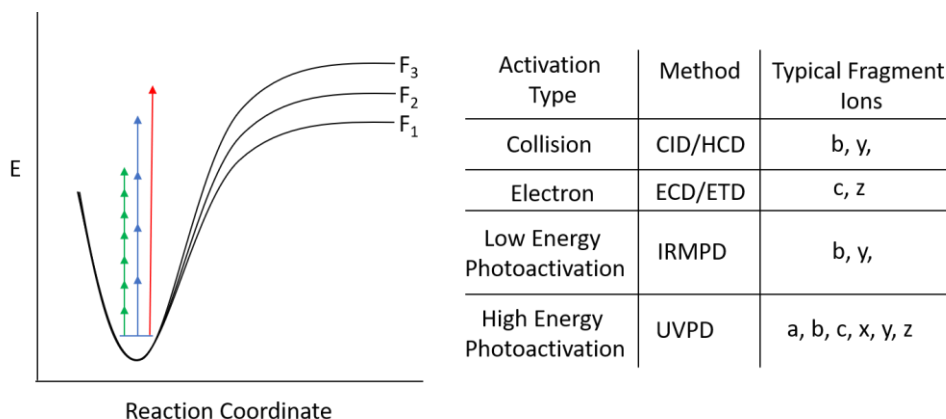
Figure 1.1: Simplified reaction coordinate diagram for various activation techniques and the energy imparted by the activation onto the peptide. The red arrow corresponds to energy deposition in a single high energy activation step, while blue and green require multiple lower energy activation events for fragmentation to occur. The chart on the right shows common methods for the various activation types as well as typical fragment ions observed.
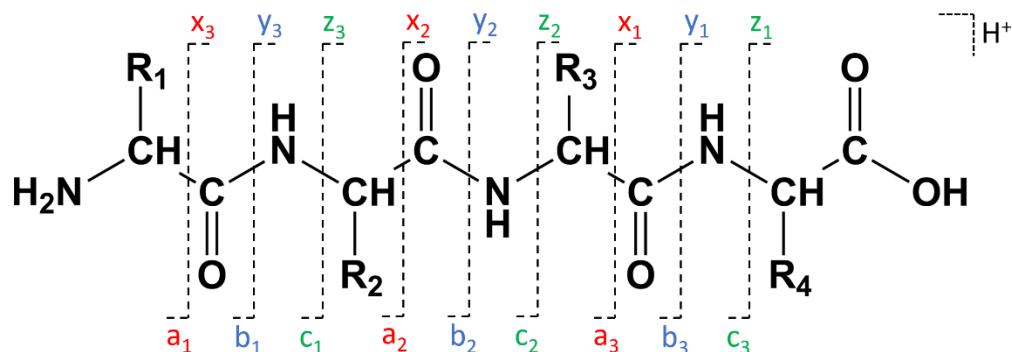
Figure 1.2: Nomenclature for peptide backbone fragments, a-, b-, c-, x-, y-, and z-type ions, produced in the positive ion mode. The numerical subscript refers to the number of amino acids included in the fragment ion.

Tandem-MS techniques are generally robust for characterization of the primary

sequence of peptides and proteins. Collision-based methods, like CID, are the most widely

used technique for such analyses. Such MS/MS experiments are performed in a variety

of mass spectrometry instruments due to its speed and sensitivity. In CID, protonated

peptides and proteins are subjected to collisions with a neutral gas, most commonly He or N_2. These molecules are heated through these collisions, which results in the breaking of the lowest energy bonds. The formation of b/y-ions is most commonly observed for CID providing information about the primary structure of the peptide or protein. Higher-energy collision dissociation (HCD) is a recent alternative to CID. In HCD, ions are accelerated into a high-pressure collision cell (HCD cell) where more energy per impact of the peptides with the neutral gas is observed. By using orbitrap detection, HCD can resolve ions of lower masses by avoiding the low mass cutoff. This allows for the identification of the presence of protein modifications such as phosphorylation. HCD offers a higher energy alternative to CID, producing a wider range of fragment ions. Although, they primarily remain b/y-type fragments. A common drawback to CID and HCD is that these methods often lead to dissociation of common post-translational modifications (PTMs) (e.g. phosphorylation and glycosylation) which limits their utility for the site localization of such modifications.

ECD and ETD are two electron-based activation methods which have gained widespread use in the characterization of peptides and proteins.[22] ECD relies on the capture of an electron by a positively-multicharged analyte, which leads to charge reduction and fragmentation. An alternative to ECD is ETD, which relies on a positively charged analyte reacting with a radical anion, leading to transfer of an electron and subsequent fragmentation. The resulting charge reduced ions produced from either ECD or ETD yield dissociation of the $N-C_\alpha$ bond, generating c- and z-type fragment ions. Notably, these two techniques can retain PTMs on the peptide or protein allowing for identification of the modified site.

1.5 Spectroscopy-Coupled Mass Spectrometry

Spectroscopy offers a different avenue for the characterization of peptides and proteins. The coupling of a laser to a mass spectrometer first occurred over forty years ago.[23] With the advancement of technology in mass spectrometers and increased laser availability, the number of applications has significantly increased. For instance, the Orbitrap Fusion Lumos is a commercially available instrument, which can be outfitted with a 213 nm ultraviolet (UV) laser to perform photodissociation experiments. Infrared multiphoton dissociation (IRMPD) and ultraviolet photodissociation (UVPD) induce fragmentation of ions through the absorption of photons. Photons in IRMPD are produced from a continuous wave CO_2 laser at a wavelength of 10.6 um, which corresponds to 0.1 eV per photon. As such, a single absorbed photon in IRMPD is often not enough to induce fragmentation of a molecule due to its low energy. This necessitates that multiple photons must be absorbed by the molecule in order to induce fragmentation.[24] In contrast to low energy photons produced in IRMPD, higher energy photons are produced in UVPD. UV photoactivation can be accomplished with a variety of wavelengths, including but not limited to 157, 193, 213, 266 and 355 nm.[25,26] Often, a single absorbed photon is enough to induce fragmentation of the molecule. For instance, 7.9, 6.4, and 4.7 eV per photon corresponds to the UV wavelengths of 157, 193, and 266 nm respectively. Each of these wavelengths can provide unique structural information about the biomolecule, provided a suitable chromophore is present for photon absorption.

The application of 157 and 193 nm has been well-documented in the study of biological and organic molecules, such as nucleic acids, proteins, lipids and glycopeptides, due to their high photoabsorptivity.[25,27] Much of the mechanistic framework for 157 nm UVPD was established by Reilly and coworkers.[28-32] The use of 157 nm UVPD as a proteomic tool is

has seen limited use due to the high energy photons, costly F_2 laser set-up, and complicated data analysis (due to obliteration of the biological molecules of interest), which makes 193 nm UVPD the more commonly utilized wavelength for proteomics. Photons are generated through the use of an ArF excimer laser which allows for higher pulse energies than the 157 nm laser. The amide bonds of peptides and proteins act as chromophores for the 193 nm photons, allowing for abundant backbone fragmentation throughout the peptides and proteins.[33] Due to the high energy photons and abundance of chromophores, 193 nm UVPD has mostly been utilized for top-down characterization of peptides and moderately sized proteins.[34-36] To a lesser degree, 193 nm UVPD has also been utilized for the characterization of PTMs in peptides and proteins, including glycosylation, phosphorylation, and sulfation.[37-39] The utility of the high energy lasers to cleave the backbone of proteins is great for top-down proteomics, enabling full characterization of proteins of various sizes.[40]

266 nm photons are produced by the fourth harmonic of a solid state Nd:YAG laser and are lower in energy than those produced at 193 nm. These lower energy photons do not often induce backbone fragmentation as they are not absorbed well by the amide bonds. Native aromatic residues like tyrosine, tryptophan, and phenylalanine show significant absorption at 266 nm, and the presence of these chromophores allow photon absorption and subsequent fragmentation of the biomolecule. With the requirement of specific aromatic residues, 266 nm can offer more specific fragmentation pathways than that of 157 or 193 nm. For instance, modification of tyrosine with iodine creates photocleavable C-I bonds. Dissociation of these C-I bonds gives rise to a radical, which can migrate through space to induce peptide fragmentation. This technique, termed radical directed dissociation (RDD), is structurally sensitive and can offer information about the local 3-D

10

structure of a protein.[41-46] With the variety of wavelengths available, many opportunities exist to develop new and exciting ways for the characterization of peptides and proteins.

1.6 Energy Transfer in the Gas Phase: FRET and Dexter

The coupling of spectroscopy with mass spectrometry has led to interest in using energy transfer systems for gas phase protein structure characterization. Energy transfer experiments rely on the presence of a donor, excited at a specific wavelength, followed by energy transfer to an acceptor, whose absorption band overlaps the emission band of the donor. The requirement of specific distance constraints for energy transfer to occur makes these systems useful for the determination of three-dimensional protein structure. Many mechanisms of energy transfer exist, the most common of which is Förster resonance energy transfer (FRET), which has been extensively used in protein conformational studies, single molecule structural dynamics, enzymatic pathways, and many other biological applications.[47-50] FRET is a radiationless process in which energy is transferred from the donor to the acceptor through dipole-dipole coupling between the chromophores. This process does not require any molecular collision, and energy is not converted to thermal energy prior to energy transfer occurring. While many factors influence the efficiency of FRET, a few of these factors standout above the rest: 1) the donor and acceptor must be between 10-100 Å apart, 2) the absorption spectrum of the acceptor must overlap the emission spectrum of the donor, and 3) the dipole-dipole orientation of the chromophores must be close to parallel. The most important element of the three listed is that the chromophores must be close in space. The efficiency of energy transfer is dependent on one-sixth the inverse distance between the donor and acceptor (see Equation 1.1). This is to say that small changes in distance between the chromophores will have a large impact on the energy transfer efficiency.

$$E = \frac{1}{1 + \dfrac{r^6}{R_o^6}}$$

Equation 1.1: The FRET equation, where E is a measure of the energy transfer efficiency, r is the distance between the donor and acceptor, and R_o is the Förster distance for the donor-acceptor pair where FRET occurs with 50% energy transfer efficiency.

In contrast to FRET, Dexter energy transfer is an energy transfer mechanism which relies on significantly tighter distance constraints. In Dexter energy transfer, an electron exchange occurs between the donor and acceptor through the overlap of molecular orbitals. Equation 1.2 highlights the key variables of Dexter energy transfer. Much like FRET, Dexter energy transfer is distance sensitive. However, exponential decay is observed as the distance between the chromophores increases. Thus, Dexter energy transfer typically occurs between the donor and acceptor at distances <20 Å. While these two energy transfer mechanisms do differ, there are fundamental difficulties in determining which mechanism energy transfer has occurred. In order to be inclusive to all energy transfer mechanisms, excitation energy transfer (EET) will be used to refer to a combination of energy transfer mechanisms likely at play.

$$k_{dexter} = KJe^{\frac{-2R_{DA}}{L}}$$

Equation 1.2: The Dexter energy transfer rate equation, where K_{dexter} is the rate constant for the energy transfer efficiency. J is the normalized spectral overlap integral, K is an experimental determined factor, R_{DA} is the distance in space between the donor and acceptor, and L is the sum of the van der Waals radius.

With the spectroscopic successes for biomolecular characterization in the condensed phase, it is only natural for the desire to couple these methods with mass spectrometry. In order to monitor FRET, there are two ways to measure fluorescence within a mass

spectrometer. The first is to draw parallels to the liquid phase and measure direct photon fluorescence from the acceptor chromophore. However, direct measurement of the photon fluorescence inside a mass spectrometer is challenging due to a few compounding reasons: 1) the implementation of fluorescence detectors inside mass spectrometry instrumentation is challenging, 2) the direction of photon emission inside the mass spectrometer is impossible to control, resulting in sensitivity losses that scale exponentially with the distance from the emitted radiation and the photon collector, and 3) due to the low availability of ions in the ion trap, it is often difficult to obtain significant optical signal for detection. Despite these obstacles, there are reports of FRET measurements in the gas phase. The Parks group has performed gas-phase FRET experiments in examining oligonucleotides and peptides derivatized with common liquid phase FRET dyes.[51,52] The Jockush and Zenobi groups have performed similar experiments using a modified quadrupole ion trap mass spectrometer to allow for the direct measurement of photons inside the mass spectrometer.[53-55] In their experiments, peptides were modified with FRET chromophores and subsequently sprayed into a mass spectrometer and photoactivated. An emission spectrum was obtained for the FRET acceptor, providing spatial information about the distance between the donor and acceptor on the peptide in the gas phase.

Rather than direct measurement of photon absorption or emission, it is also possible to examine how photons affect a molecule by monitoring changes in their m/z following photon absorption. This technique, termed action spectroscopy, has been around for some time.[56-59] Following photoexcitation, mass losses are detected by mass spectrometry and correlated to the excitation wavelength. By coupling action spectroscopy with EET, it is possible to have a probe that takes advantage of the unparalleled ability of mass spectrometry to discriminate mass changes. Therefore, the probe is sensitive to

13

changes in the biomolecule's structure. The Dugourd and Zenobi groups have examined such systems through the modifications of small peptides and proteins with large liquid phase FRET chromophores (action-FRET).[60-63] Our group has also examined such action energy transfer systems (action-EET) using different donor and acceptor systems for energy transfer. However, rather than modification of peptides with large FRET chromophores, we have relied on native aromatic donors: tyrosine, tryptophan, and phenylalanine, in conjunction with a disulfide bond as the acceptor. In such systems, photoexcitation at 266 nm of peptides containing tyrosine or tryptophan leads to energy transfer to a disulfide bond and enhanced S-S dissociation. Tyrosine is able to participate as an energy donor at distances <6 Å from the acceptor while tryptophan shows energy transfer efficiency up to 15 Å from the acceptor.[64,65] Phenylalanine does not show any single-step energy transfer to the disulfide bond. However, a two-step system showed that phenylalanine could transfer energy to tyrosine, and this energy could be transferred to the disulfide bond, yielding enhanced bond-specific dissociation.[66]

1.7 Education and Research: Tools for Improved Student Learning Outcomes

The process of thinking as a scientist is not limited to the lab experiments with chemicals and complex instrumentation.[67] Chemical education research (CER) is focused on teaching and learning in chemistry. Rather than a focus on molecules, CER investigates people and how the implementation of both quantitative and qualitative methods will improve students' understanding and application of the material. This field is rather broad, with research ranging from the teaching and learning of thermodynamics to spatial reasoning within chemistry.[68,69] As such, there are many possible ways in which one can implement CER.

Early chemistry courses were not taught through a research-based approach. In fact, it wasn't until the early 1800's that it was believed that a better understanding of chemistry could lead to improved productivity in the workplace.[70] Many early theories of learning revolved around treating the learner as a "blank slate", devoid of any pre-existing knowledge or conceptions.[71] However, this has since been shown to be false. In chemistry, students often enter with many misconceptions about the way in which things work. For instance, Osborn and Cosgrove identified that students believed that when water boils, the bubbles were composed of hydrogen and oxygen gases.[72] This certainly confirms that students do not enter the classroom as a "blank slate". By building on the groundwork laid by John Dewey, Jean Piaget was key in laying the empirical foundations for constructivism, the notion that pre-existing knowledge plays a key role in the development of an individual in science.[73,74] In 1968, American psychologist David Ausubel played a key role in refining constructivism further and defined "meaningful learning" in the classroom.[75] Ausubel believed that three criteria must be met for learning to occur. These criteria are: 1) students must have appropriate prior knowledge that they can connect to new ideas, making the act of removing any misconceptions key, 2) in order to build the new knowledge into existing knowledge, it must be perceived as relevant, and 3) students must choose to make these connections, i.e. the student must perform the act of learning for it to be meaningful. This indicates that it is the choice of the student to want to learn and how they choose to learn. These three ideas have since become known as Ausubel's assimilation theory. Without meeting these conditions, the learner may lean on rote learning, or rote memorization, in which the student memorizes a fact without connecting it to any pre-existing ideas or cognitive structures.

Joseph Novak has elaborated on what Ausubel called "meaningful learning" by proposing that thinking, feeling, and acting were necessary for meaningful learning to occur.[76,77] Thus, it is important that methodology is developed to engage students in the material for the best results. As such, considerable effort has been undertaken to apply metacognitive interventions to improve student learning. These metacognitive interventions include, but are not limited to, paraphrasing and rewriting, working on homework problems, previewing material, and pretending to teach information.[78] While these strategies have been well-reported in the literature, there remains a need for the development and implementation of strategies which can assess student learning and improve student conceptual retention.[79-81]

1.8 Scope of the Dissertation

As previous sections in this chapter have demonstrated, chemistry exists in the classroom and the laboratory. In order to demonstrate this, the work herein will focus on two chemistry domains: 1) the development of new methodology for the characterization of biological molecules with mass spectrometry and 2) the implementation of a conceptual learning tool, concept maps, to improve student conceptual retention of general chemistry material.

Chapter 2 explores the utility of 213 nm photons for initiating bond-selective fragmentation. It is found that bonds previously determined to be labile at 266 nm, including carbon-iodine and sulfur-sulfur bonds, can also be cleaved with high selectivity at 213 nm. In addition, many carbon-sulfur bonds that are not subject to direct dissociation at 266 nm can be selectively fragmented at 213 nm. This capability can be used to site-specifically create alaninyl radicals that direct backbone dissociation at the radical site, creating diagnostic d-ions. Furthermore, the additional carbon-sulfur bond fragmentation

capability leads to signature triplets for fragmentation of disulfide bonds. Absorption of amide bonds can enhance dissociation of nearby labile carbon-sulfur bonds and can be used for stochastic backbone fragmentation typical of UVPD experiments at shorter wavelengths. Several potential applications of the bond-selective fragmentation chemistry observed at 213 nm are discussed.

In chapter 3, two new energy acceptors are identified that can be utilized for action-excitation energy transfer experiments. In the first system, C-S bonds in methionine act as energy acceptors from native chromophores, including tyrosine, tryptophan, and phenylalanine. Comparison among chromophores reveals that tyrosine transfers energy most efficiently at 266 nm, but phenylalanine and tryptophan also transfer energy with comparable efficiencies. Overall, the C-S bond dissociation yields following energy transfer are low for methionine, which led to investigation of selenomethionine, a common analogue that is found in many naturally occurring proteins. Sulfur and selenium are chemically similar, but C-Se bonds are weaker than C-S bonds and have lower lying σ* anti-bonding orbitals. Excitation of peptides containing tyrosine and tryptophan results in efficient energy transfer to selenomethionine and abundant C-Se bond dissociation. A series of helical peptides were examined where the positions of the donor or acceptor were systematically scanned to explore the influence of distance and helix orientation on energy transfer. Distance was found to be the primary factor affecting energy transfer efficiency.

In chapter 4, radical chemistry is used for bond synthesis in the gas phase rather than traditional bond breaking. Single peptides containing two cysteine residues capped with propylmercaptan (PM) often form disulfide bonds following ultraviolet excitation at 266 nm and loss of both PM groups. Similarly, noncovalently-bound peptide pairs, where each

17

peptide contains a single cysteine residue, can be induced to form disulfide bonds. Comparison with disulfide-bound species sampled directly from solution yields identical collisional activation spectra, suggesting that native disulfide bonds have been recapitulated in the gas phase syntheses. Another approach utilizing radical chemistry for covalent bond synthesis involves creation of a reactive diradical that can first abstract hydrogen from a target peptide, creating a new radical site, and then recombine the second radical with the new radical to form a covalent bond. This chemistry is illustrated with 2-(hydroxymethyl-3,5-diiodobenzoate)-18-crown-6 ether, which attaches noncovalently to protonated primary amines in peptides and proteins. Following photoactivation and crosslinking, the site of noncovalent adduct attachment can frequently be determined. The ramifications of these observations on peptide structure and noncovalent attachment of 18-crown-6-based molecules is discussed.

In chapter 5, a concept map assignment was implemented in a first-year college-level general chemistry course in an effort to improve student conceptual understanding and help students better connect pre-existing knowledge to new ideas. This implementation included a quasi-experiment that was carried out in two separate courses in the year-long general chemistry sequence. Students enrolled in a single section of the course were divided into two groups in which a concept map treatment was compared to a control group that completed short journal entries. Students in the treatment group were instructed on a weekly basis to develop increasingly complicated maps that added all new concepts and connecting ideas, whereas the journal control group was simply instructed to summarize the activities completed each week in lecture. The Fall 2017 implementation did not show any significant difference between the treatment and control group in exam scores, self-reported gains in conceptual understanding, or self-reported gains in

connecting chemistry to other disciplines. In the Spring 2018 implementation, a multi-variable regression analysis of concept test scores revealed the top 33% of students in the concept map treatment scored on average 7% higher than the top 33% of students in the journal control group (unstandardized B = 1.165, p = 0.021). The quality of the students' concept maps was also evaluated and correlated to student performance on the concept inventory, and it appears students who were better at concept mapping made greater gains in conceptual understanding (Pearson's r = 0.295, p = 0.003). An examination of SALG survey questions revealed higher self-reported gains for the treatment group. This study does not provide unequivocal evidence that a concept map treatment leads to greater gains in conceptual understanding compared to a control population or that students with better concept mapping skills performed better on the concept inventory instrument. However, the results presented herein might prompt chemistry instructors to consider including concept map assignments as part of their instructional arsenal.

1.9 References

[1] Y. Shi, Cell. 159 (2014) 995–1014.

[2] J.D. Watson, F.H.C. Crick, Nature. 171 (1953) 737–738.

[3] M.F. Perutz, H. Muirhead, J.M. Cox, L.C.G. Goaman, F.S. Mathwes, E.L. McGandy, L.E. Webb, Nature. 219 (1968) 29–32.

[4] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, Nature. 181 (1958) 662–666.

[5] C.C.F. Blake, D.F. Koenig, G.A. Mair, A.C.T. North, D.C. Phillips, V.R. Sarma, Nature. 206 (1965) 757–761.

[6] G. Kartha, J. Bello, D. Harker, Nature. 213 (1967) 862–865.

[7] B.W. Matthews, P.B. Sigler, R. Henderson, D.M. Blow, Nature. 214 (1967) 652–656.

[8] P.R.L. Markwick, T. Malliavin, M. Nilges, PLOS Comput. Biol. 4 (2008) e1000168.

[9] A. Cavalli, X. Salvatella, C.M. Dobson, M. Vendruscolo, Proc. Natl. Acad. Sci. 104 (2007) 9615–9620.

[10] K. Murata, M. Wolf, Biochim. Biophys. Acta - Gen. Subj. 1862 (2018) 324–334.

[11] Y. Liu, D.T. Huynh, T.O. Yeates, Nat. Commun., 10 (2019) 1864.

[12] L. Konermann, J. Pan, Y.-H. Liu, Chem. Soc. Rev., 40 (2011) 1224–1234.

[13] G.R. Masson, M.L. Jenkins, J.E. Burke, Expert Opin. Drug Discov. 12 (2017) 981–994.

[14] J. Claesen, T. Burzykowski, Mass Spectrom. Rev. 36 (2017) 649–667.

[15] J.W. Back, L. de Jong, A.O. Muijsers, C.G. de Koster, J. Mol. Biol., 331 (2003) 303–313.

[16] A. Sinz, Mass Spectrom. Rev., 25 (2006) 663–682.

[17] Z. Hall, A. Politis, M.F. Bush, L.J. Smith, C. V Robinson, J. Am. Chem. Soc., 134 (2012) 3429–3438.

[18] M.F. Bush, Z. Hall, K. Giles, J. Hoyes, C. V Robinson, B.T. Ruotolo, Anal. Chem., 82 (2010) 9557–9565.

[19] C. Uetrecht, R.J. Rose, E. van Duijn, K. Lorenzen, A.J.R. Heck, Chem. Soc. Rev., 39 (2010) 1633–1655.

[20] C.D. Chouinard, M.S. Wei, C.R. Beekman, R.H.J. Kemperman, R.A. Yost, Clin. Chem., 62 (2016) 124 LP – 133.

[21] Y. Zhong, L. Han, B.T. Ruotolo, Angew. Chemie Int. Ed., 53 (2014) 9209–9212.

[22] Y. Qi, D.A. Volmer, Mass Spectrometry Reviews, 36(1), (2017) 4–15.

[23] R.R. Woodin, D.S. Bomse, J.L. Beauchamp, Academic Press; New York: 1979. pp. 355–388.

[24] J.S. Brodbelt, J.J. Wilson, Mass Spectrom. Rev., 28 (2009) 390–424.

[25] J.S. Brodbelt, Chem. Soc. Rev., 43 (2014) 2757–2783.

[26] T. Ly, R.R. Julian, Angew. Chemie-International Ed., 48 (2009) 7130–7137.

[27] A. Devakumar, Y. Mechref, P. Kang, M. V Novotny, J.P. Reilly, Rapid Commun. Mass Spectrom., 21 (2007) 1452–1460.

[28] T.-Y. Kim, M.S. Thompson, J.P. Reilly, Rapid Commun. Mass Spectrom., 19 (2005) 1657–1665.

[29] L. Zhang, W. Cui, M.S. Thompson, J.P. Reilly, J. Am. Soc. Mass Spectrom., 17 (2006) 1315–1321.

[30] M.S. Thompson, W. Cui, J.P. Reilly, J. Am. Soc. Mass Spectrom., 18 (2007) 1439–1452.

[31] L. Zhang, J. Reilly, J. Am. Soc. Mass Spectrom., 19 (2008) 695–702.

[32] J.P. Reilly, Mass Spectrom. Rev., 28 (2009) 425–447.

[33] J.A. Madsen, D.R. Boutz, J.S. Brodbelt, J. Proteome Res., 9 (2010) 4205–4214.

[34] L. Vasicek, J.S. Brodbelt, Anal. Chem., 82 (2010) 9441–9446.

[35] J.B. Shaw, W. Li, D.D. Holden, Y. Zhang, J. Griep-Raming, R.T. Fellers, B.P. Early, P.M. Thomas, N.L. Kelleher, J.S. Brodbelt, J. Am. Chem. Soc., 135 (2013) 12646–12651.

[36] J.P. O'Brien, W. Li, Y. Zhang, J.S. Brodbelt, J. Am. Chem. Soc., 136 (2014) 12920–12928.

[37] J.A. Madsen, T.S. Kaoud, K.N. Dalby, J.S. Brodbelt, Proteomics, 11 (2011) 1329–1334.

[38] B.J. Ko, J.S. Brodbelt, Int. J. Mass Spectrom., (2015) 385–392.

[39] M.R. Robinson, K.L. Moore, J.S. Brodbelt, J Am Soc Mass Spectrom, 25 (2014) 1461–1471.

[40] J.R. Aponte, L. Vasicek, J. Swaminathan, H. Xu, M.C. Koag, S. Lee, J.S. Brodbelt, Anal. Chem., 86 (2014) 6237–6244.

[41] X. Zhang, R.R. Julian, Int. J. Mass Spectrom., 308 (2011) 225–231.

[42] Q. Sun, H. Nelson, T. Ly, B.M. Stoltz, R.R. Julian, J. Proteome Res., 8 (2009) 958–966.

[43] T. Ly, R.R. Julian, J. Am. Chem. Soc., 132 (2010) 8602–8609.

[44] Q. Sun, S. Yin, J.A. Loo, R.R. Julian, Anal. Chem., 82 (2010) 3826–3833.

[45] H.T. Pham, R.R. Julian, Analyst, 141 (2016) 1273–1278.

[46] Y.A. Lyon, M.P. Collier, D.L. Riggs, M.T. Degiacomi, J.L.P. Benesch, R.R. Julian, J. Biol. Chem., (2019) jbc.RA118.007052.

[47] D. Kajihara, R. Abe, I. Iijima, C. Komiyama, M. Sisido, T. Hohsaka, Nat. Methods, 3 (2006) 923.

[48] M. Sugawa, Y. Arai, A.H. Iwane, Y. Ishii, T. Yanagida, Biosystems, 88 (2007) 243–250.

[49] R.T. Cummings, S.P. Salowe, B.R. Cunningham, J. Wiltsie, Y.W. Park, L.M. Sonatore, D. Wisniewski, C.M. Douglas, J.D. Hermes, E.M. Scolnick, Proc. Natl. Acad. Sci., 99 (2002) 6603 LP-6606.

[50] Y. Arai, T. Nagai, Microscopy, 62 (2013) 419–428.

[51] A.S. Danell, J.H. Parks, J. Am. Soc. Mass Spectrom., 14 (2003) 1330–1339.

[52] X. Shi, J.H. Parks, J. Am. Soc. Mass Spectrom., 21 (2010) 707–718.

[53] M.F. Czar, R.A. Jockusch, Curr. Opin. Struct. Biol., 34 (2015) 123–134.

[54] M.F. Czar, F. Zosel, I. König, D. Nettels, B. Wunderlich, B. Schuler, A. Zarrine-Afsar, R.A. Jockusch, Anal. Chem., 87 (2015) 7559–7565.

[55] M. Dashtiev, V. Azov, V. Frankevich, L. Scharfenberg, R. Zenobi, J. Am. Soc. Mass Spectrom., 16 (2005) 1481–1487.

[56] N.C. Polfer, J. Oomens, Mass Spectrom. Rev., 28 (2009) 468–494.

[57] T.D. Fridgen, Mass Spectrom. Rev., 28 (2009) 586–607.

[58] L. Feketeová, G.N. Khairallah, B. Chan, V. Steinmetz, P. Maître, L. Radom, R.A.J. O'Hair, Chem. Commun., 49 (2013) 7343–7345.

[59] Y. Kim, K. Motobayashi, T. Frederiksen, H. Ueba, M. Kawai, Prog. Surf. Sci., 90 (2015) 85–143.

[60] A. Kulesza, S. Daly, C.M. Choi, A.-L. Simon, F. Chirot, L. MacAleese, R. Antoine, P. Dugourd, Phys. Chem. Chem. Phys., 18 (2016) 9061–9069.

[61] S. Daly, G. Knight, M.A. Halim, A. Kulesza, C.M. Choi, F. Chirot, L. MacAleese, R. Antoine, P. Dugourd, J. Am. Soc. Mass Spectrom., 28 (2017) 38–49.

[62] S. Daly, F. Poussigue, A.-L. Simon, L. MacAleese, F. Bertorelle, F. Chirot, R. Antoine, P. Dugourd, Anal. Chem., 86 (2014) 8798–8804.

[63] V. Frankevich, V. Chagovets, F. Widjaja, K. Barylyuk, Z. Yang, R. Zenobi, Phys. Chem. Chem. Phys., 16 (2014) 8911–8920.

[64] N.G. Hendricks, R.R. Julian, Analyst, 141 (2016) 4534–4540.

[65] N.G. Hendricks, N.M. Lareau, S.M. Stow, J.A. McLean, R.R. Julian, J. Am. Chem. Soc., 136 (2014) 13363–13370.

[66] N.G. Hendricks, R.R. Julian, Chem. Commun., 51 (2015) 12720–12723.

[67] M.M. Cooper, R.L. Stowe, Chem. Rev., 118 (2018) 6053–6087.

[68] K. Bain, A. Moon, M.R. Mack, M.H. Towns, Chem. Educ. Res. Pract., 15 (2014) 320–335.

[69] M. Harle, M. Towns, J. Chem. Educ., 88 (2011) 351–360.

[70] H. Hale, J. Chem. Educ., 9 (1932) 729.

[71] B.S. Craig, Journal of Chemical Education, 49(12), (1972) 807. 7

[72] R. Osborne, M.J. Cosgrove, Res. Sci. Teach. 1983, 20, 825.

[73] J. Piaget, T. Brown, K.J. Thampy, University of Chicago Press: Chicago, 1985.

[74] J. Dewey, Macmillan: New York,1916.

[75] D.P. Ausubel, (1968). New York: Holt, Rinehart, & Winston.

[76] J.D. Novak, J. Chem. Educ., 61 (1984) 607.

[77] J.D. Novak, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, 1998.

[78] D. Rickey, A.M. Stacy, J. Chem. Educ., 77 (2000) 915.

[79] A.H. Johnstone, J. Chem. Educ., 70 (1993) 701.

[80] E. Cook, E. Kennedy, S.Y. McGuire, J. Chem. Educ., 90 (2013) 961–967.

[81] K.R. Galloway, S.L. Bretz, J. Chem. Educ., 92 (2015) 2019–2030.

Chapter 2: Directed-Backbone Dissociation following Bond-Specific Carbon-Sulfur

UVPD at 213 nm

2.1 Introduction

Ultraviolet Photodissociation (UVPD) is becoming an increasingly popular choice for fragmenting ions in mass spectrometers.[1] UVPD differs from other dissociation techniques such as collision-induced dissociation (CID)[2,3] or electron-transfer dissociation (ETD)[4] in several important ways. For example, UVPD causes dissociation through nearly simultaneous operation of two mechanisms, vibrational heating and excited-state dissociation.[5] The relevant timescale for UVPD experiments can also be significantly shorter than CID or ETD, occurring within nanoseconds for experiments utilizing high-powered lasers.[6] Furthermore, the requirement of absorption by a chromophore offers an orthogonal parameter, independent of mass or charge, for controlling which ions will be excited. UVPD experiments have been conducted at many wavelengths including 157, 193, 213, 266, and 355 nm.[7-13] Wavelength selection significantly impacts fragmentation by modulating the ratio of fragmentation mechanisms available. For example, at 266 nm, very few direct dissociation pathways exist, and the abundance of natural chromophores in analytes is small. At 157 nm, most bonds act as chromophores, and direct dissociation pathways are expected to be more plentiful.

Previous efforts to favor bond-selective direct dissociation, which occurs due to excitation to a dissociative excited state, have primarily utilized 266 nm photons.[14] Early experiments demonstrated that carbon-iodine bonds could be selectively and homolytically cleaved in peptides or proteins, leaving all other bonds intact.[15] Carbon-bromine bonds can also be cleaved in this fashion, though the yield is reduced.[16] Sulfur-sulfur bonds are the only native bonds in peptides or proteins that undergo direct

24

dissociation at 266 nm.[17] Carbon-sulfur bonds can be cleaved with 266 nm photons, but only if the sulfur atom is directly attached to a suitable chromophore, such as naphthalene.[18,19] Importantly, all of these labile bonds can be cleaved with a high degree of selectivity even in large molecules, meaning that in many cases, a single bond is the only site of dissociation. This degree of control over fragmentation enables a variety of unique applications.

For example, disulfide bond partners from a protein can be easily identified.[17] In this experiment, the protein is digested with disulfide bonds intact and then subjected to chromatographic separation and MS/MS analysis with UVPD at 266 nm. Peptide pairs bound by a disulfide bond cleave into the constituent peptides while other monomeric peptides remain intact. Bond-selective UVPD can also be used to identify UV labile post-translational modifications such as iodination of tyrosine or quinone reactions with cysteine.[16] When coupled with targeted wet-chemistry modifications, bond-selective UVPD can be used to identify sites of phosphorylation by directing backbone fragmentation at the site of modification.[18] All of these previous examples utilized 266 nm photons, which favor direct dissociation but also suffer from weak absorption by many chromophores.

Recently, interest in UVPD at 213 nm has increased because this wavelength is afforded by a solid-state laser, and the photon energy resides at the threshold of absorption by small chromophores such as amide bonds. In theory, this could allow UPVD at 213 nm to favor either direct dissociation or nonspecific dissociation, depending on experimental parameters (such as laser power and excitation time) or molecular composition (most critically, the chromophores available for absorption). Herein, the photochemistry of UVPD at 213 nm is explored for a variety of modified and native

peptides. The propensity for bond-selective fragmentation versus undirected UVPD is examined as a function of peptide composition. The capacity for fragmenting both native and synthetically appended C-S bonds with 213 nm photons is investigated. The prospect for using bond-specific C-S bond cleavage to subsequently direct backbone fragmentation at targeted residues is evaluated. Applications of 213 nm light for bond specific cleavage of S-S bonds, which simultaneously yields C-S bond fragmentation, are explored. Advantages and disadvantages relative to similar experiments conducted at 266 nm are discussed.

2.2 Experimental Methods

2.2.1 Materials

Peptides RKRRQtSM, RQSVELHsPQSLPR, RGDC, RPHERNGFTVLCPKN, HCLGKWLGHPDKF, NTWTTCQSIAFPSK, and SHLVEALYLVCGERG were purchased from Anaspec (San Jose, CA). SLRRSSCFGGR and CQDSETRTFY were purchased from Abbiotec (San Diego, CA). CDPGYIGSR was purchased from Apexbio, and CGYGPKKKRKVGG was purchased from American Peptide Company (Sunnyvale, CA). GSNKGAIIGLM (Piscataway, NJ) was purchased from GenScript. DRVYIHPF was purchased from Sigma-Aldrich (St. Louis, MO). Naphthoquinone (NQ), iodoacetamide and dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich (St. Louis, MO). Benzoquinone (BQ), benzyl mercaptan (BM), and trifluoroacetic acid (TFA) were purchased from Alfa Aesar (Haverhill, MA). Naphthalenethiol (NT) was purchased from Fluka Analytical (Mexico City, Mexico). Iodomethane was purchased from Arcos Organics (Geel, Belgium). Chloramine-T, sodium metabisulfite, and sodium iodide were purchased from Fisher Chemical (Fairlawn, NJ). Acetonitrile (ACN) and methanol were purchased from Fisher Scientific (Waltham, MA). Water was purified by Millipore Direct-Q (Millipore,

Billerica, MA). A Macrotrap holder and Macrotrap consisting of polymeric reversed-phase packing material were purchased from Michrom Bioresources, Inc (Auburn, CA).

2.2.2 Peptide Iodination

Peptides were iodinated using a previously published method.[20] Briefly, equimolar peptide and sodium iodide were mixed with a two-fold molar excess of chloramine-T for five minutes at room temperature in water. The reaction was then quenched by the addition of 4x molar excess sodium metabisulfite. The products were then purified with a peptide microtrap, rinsed with 0.1% TFA in 90:10 H_2O:ACN and eluted with 0.1%TFA in 2%:98% H2O:ACN.

2.2.3 Quinone Modification

Modification of peptides was carried out with either benzoquinone or naphthoquinone based on a previous procedure.[16] Stocks of each quinone were prepared fresh prior to each use and stored in the dark to reduce degradation. Quinone stocks were added to peptide solutions in 0.5 to 4x excess of the peptide concentration. The reaction proceeded for 4 hours in the dark at room temperature. Following the reaction time, the solution was purified by microtrap rinsing with 0.1% TFA in 90:10 H_2O:ACN and eluted in 0.1%TFA in 2%:98% H2O:ACN. The resulting solution was lyophilized, and the powder was redissolved in 50:50 H_2O:MeOH in 0.1% formic acid for a final concentration of 4 µM or 10 µM for MS analysis.

2.2.4 Carbamidomethylation and Methylation of Free Thiols

Carbamidomethylation of cysteine was carried out by preparing a stock solution of 9 mg of iodoacetamide into 1 mL of water. To 25 µL of a 50 mM ammonium bicarbonate buffer, 10 µL of 1 mM peptide was added, followed by 6 µL of the stock iodoacetamide solution. The resulting solution was placed in the dark at room temperature for 20 minutes.

The reaction mixture was lyophilized and the powder dissolved in 50:50 H_2O:MeOH in 0.1% formic acid for a final concentration of 4 µM or 10 µM for MS analysis.

Methylation was carried out by adding 1µL of iodomethane to 25 µL of ammonium bicarbonate buffer and 10 µL of 1 mM peptide, followed by incubation in the dark at room temperature for 20 min. The modified peptide was purified by peptide trap as described above for MS analysis.

2.2.5 Preparation of Disulfide-Linked Peptides

Disulfide-linked peptides were prepared by adding 10 µL of a 1 mM peptide to 5 µL of DMSO. The reaction was mixed well and placed in a water bath at 37 ℃ for 18 hours. The resulting solution was lyophilized to remove DMSO and dissolved in 50:50 H_2O:MeOH in 0.1% formic acid for a final concentration of 4 µM or 10 µM for MS analysis.

2.2.6 Dephosphorylation and Derivatization of Phosphopeptides

Peptides were dephosphorylated based on previous procedures.[19] Briefly, 10 µL of 1 mM peptide stock was combined with 3 µL of saturated barium hydroxide and incubated at 60 ºC for 30 min. The dephosphorylated peptides were subsequently modified through the addition of 5 µL of 50 mM naphthalenethiol (2 mg dissolved in 200 µL dioxane), or 2 µL of benzyl mercaptan at 60 ℃ over 4 hours. Each reagent was 0.5 to 4x excess of the peptide concentration. Purification of the peptides was carried out with a peptide trap. The resulting solution was lyophilized, and the powder was dissolved in 50:50 H_2O:MeOH in 0.1% formic acid for a final concentration of 4 or 10 µM for MS analysis. Deamidation of glutamine was observed for only a single residue of the RQSVQLHsAPQSLPR peptide, and it has been previously shown that high temperature can induce deamidation.[21]

2.2.7 Photodissociation of Derivatized and Non-Derivatized Peptides

UVPD experiments were performed on an Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific, Waltham, MA) with a HESI II electrospray source. The HCD vacuum housing was directly modified with a quartz window to transmit the fifth harmonic (213 nm) of a diode-pumped FQSS 213-Q4 laser (Crylas, Berlin, Germany). The pulse energy was 2.5 μJ @ 1000 Hz. Photodissociation was performed by trapping ions without activation in the HCD cell for either 50, 100, or 200 ms (every millisecond equates to approximately a single pulse from the laser system), followed by mass analysis in the Orbitrap. The resolution was set to 30000. Peptides were sprayed at 4 μM concentrations at 3 μL/min with electrospray voltages set between 3 and 4 kV and the capillary inlet temperature set to 300 ºC.

An LTQ linear ion trap mass spectrometer (Fisher Scientific, Waltham, MA) with a standard ESI source was utilized for fourth harmonic (266 nm) UVPD experiments. A quartz window was installed on the back plate of the LTQ for transmission of laser pulses from a flash-lamp pumped Nd:YAG Minilite laser (Continuum, Santa Clara, CA). A single pulse of the 4 mJ laser was synchronized to occur at activation step of an MS^2 experiment which was triggered by the same external delay generator state above. Peptides were sprayed at 10 μM concentrations at 3 μL/min with electrospray voltages set between 3 and 4 kV and the capillary inlet temperature set to 275 ºC.

An LTQ linear ion trap mass spectrometer (Fisher Scientific, Waltham, MA) with a standard ESI source with an OPO tunable laser was utilized for 213 nm UVPD experiments of disulfide-linked peptides. A quartz window was installed on the back plate of the LTQ for transmission of laser pulses from a flash-lamp pumped Nd:YAG Minilite laser (Continuum, Santa Clara, CA). Three pulses of the 0.7 mJ laser were synchronized

to occur at activation step of an MS^2 experiment. Peptides were sprayed at 10 µM concentrations at 3 µL/min with electrospray voltages set between 3 and 4 kV and the capillary inlet temperature set to 275 ºC.

2.3 Results/Discussion

Photoactivation of DRVY^{I}IHPF (where Y^I = iodotyrosine) at 213 nm produces the spectrum shown in Figure 2.1a. Loss of iodine yields the most abundant product ion, suggesting that bond-selective fragmentation of the C-I bond is a favorable process. In addition, a subsequent loss of CO_2 is noted, as are numerous other fragments at lower intensity. These results bear similarity to those obtained previously[15,22] at 266 nm, but loss of iodine is not accompanied by backbone fragmentation at the longer wavelength. Closer examination of the minor products from Figure 2.1a is shown in the zoomed-in view in Figure 2.1b. A series of a-type ions and a few other backbone fragments consistent with nonspecific UVPD are noted. Excitation of the same peptide without iodotyrosine is shown in Figure 2.1c, revealing a similar dissociation pattern to that found in Figure 2.1b with the exception of a large -107 Da peak. This loss results from direct dissociation and subsequent loss of the tyrosine side chain. This pathway is much less favorable than C-I bond dissociation, rendering it uncompetitive in the presence of iodotyrosine. These results illustrate that both bond-specific dissociation and backbone fragmentation can be produced by 213 nm light. If a favorable direct dissociation pathway is available, such as cleavage of the C-I bond in iodotyrosine, then bond-selective fragmentation will likely dominate. However, in the absence of such favorable pathways, nonspecific backbone fragmentation and side chain losses will yield the most abundant fragments.
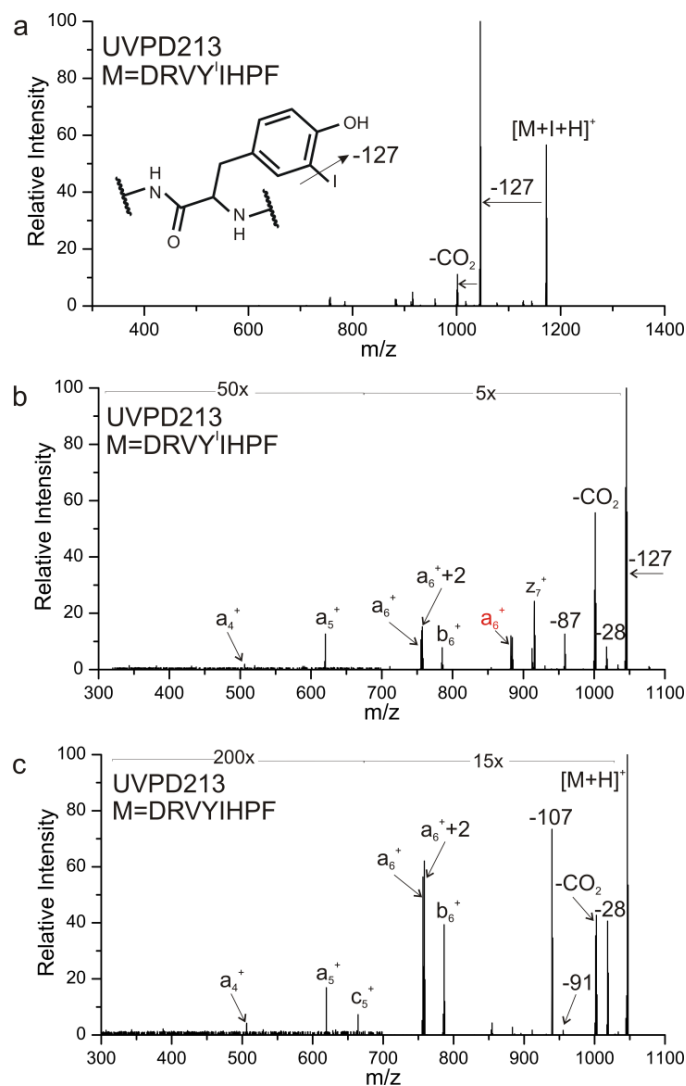
Figure 2.1: a) Photoactivation of DRVY^{I}IHPF at 213 nm, yielding dominant loss of iodine radical. b) zoom in of minor losses from a). c) Photoactivation of unmodified DRVYIHPF at 213 nm, yielding similar nonspecific fragments.

Photodissociation of native C-S bonds. The canonical amino acids cysteine and methionine contain native C-S bonds, which are candidates for bond-selective fragmentation at 213 nm. Excitation of the methionine containing peptide [GSNKGAIIGLM+H]$^+$ with 213 nm photons is illustrated in Figure 2.2a and leads to cleavage of several side chain bonds, including losses of 15 and 47 Da, corresponding to loss of methyl and thiomethyl radical, respectively. These C-S bond cleavages yield the

31

most abundant products in Figure 2.2a, suggesting that bond-selective direct dissociation is occurring. In contrast, photoexcitation of the same ion with 266 nm light yields almost no fragmentation and certainly no observable C-S bond cleavage, as shown in Figure 2.2b. Similar experiments were conducted with $[SHLVEALYLVCGERG+2H]^{2+}$, which contains a free cysteine residue. Excitation at 213 nm yields a variety of side chain losses and some backbone fragments. Importantly, the loss of 33 Da corresponds to loss of SH and fragmentation of a C-S bond. In this case, the loss of 33 Da generates an alaninyl beta radical that is poised to cleave the backbone and yield a C-terminal d-ion, as shown in Scheme 2.1. Indeed, a d_{11}^{+} ion C-terminal to the cysteine residue is the only d-ion observed in the spectrum. Excitation of $[SHLVEALYLVCGERG+2H]^{2+}$ with 266nm photons produces some dissociation, attributable primarily to absorption of the native chromophore in the side chain of tyrosine, but the loss of 33 Da is barely detectable. In addition, no d-ions are observed. These results reveal that 213 nm photons are more efficient at cleaving C-S bonds than 266 nm photons. It is likely that the dissociative excited state responsible for C-S bond cleavage is not directly accessible at 266 nm.
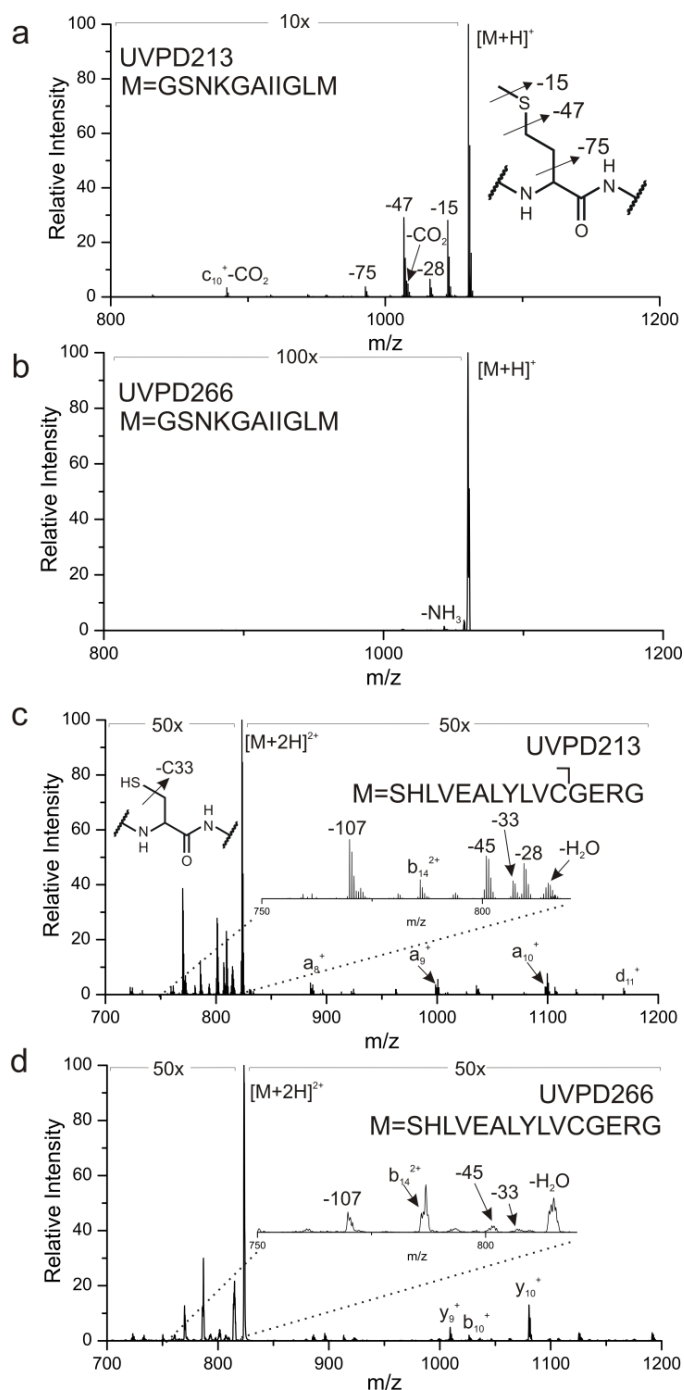
Figure 2.2: UVPD of unmodified [GSNKGAIIGLM+H]$^+$ at either a) 213 nm or b) 266 nm. Activation at 213 nm leads to dissociation of both C-S bonds of methionine. c) 213 nm UVPD of [SHLVEALYLVCGERG+2H]$^{2+}$ leads to dissociation of the single C-S bond at cysteine. d) 266 nm activation of the same peptide leads to significantly less C-S dissociation. Relevant side chain mass losses are illustrated in the insets.

33

Scheme 2.1: Pathway for d-ion formation.

d_n fragment x_n fragment
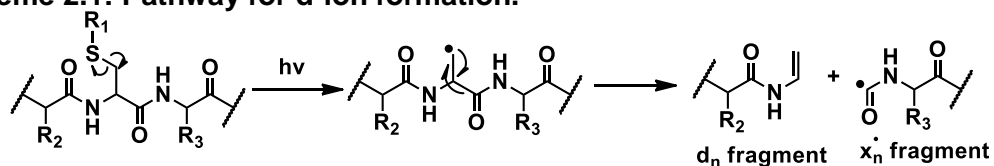
Modification of Free Thiols. Reaction of cysteine with quinones leads to nearly quantitative coupling via Michael Addition[23] chemistry, allowing for facile and specific cysteine residue labeling in peptides. If the quinone is chromophoric, then modulation of the peptide photochemistry is also possible. The results for several UVPD experiments with modified peptides are shown in Figure 2.3. Excitation of [RPHERNGFTVLC^{bq}PKN + 3H]$^{3+}$ (where C^{bq} = benzoquinone modified cysteine) with 213 nm light yields several side chains losses and backbone fragments, as illustrated in Figure 2.3a. Cleavage of both C-S bonds is observed, although the abundance does not appear to be enhanced significantly by the benzoquinone chromophore. Backbone fragmentation to generate the d_{12} ion yields similar abundance to other sites of backbone dissociation, including those at proline that are responsible for the unusual y_{14}-2 and b_{12}+2 ions.[24] In contrast, the naphthoquinone modification leads to significant enhancement of C-S bond cleavage, as shown in Figure 2.3b. In this case, the majority of ions undergo rearrangement subsequent to C-S bond cleavage, leading to the abstraction of additional hydrogens from the peptide prior to quinone loss. In other words, the primary loss is observed at 192 Da instead of the expected 189 Da, but observation of the d_{12} ion confirms creation of the expected alaninyl radical as a transient intermediate.
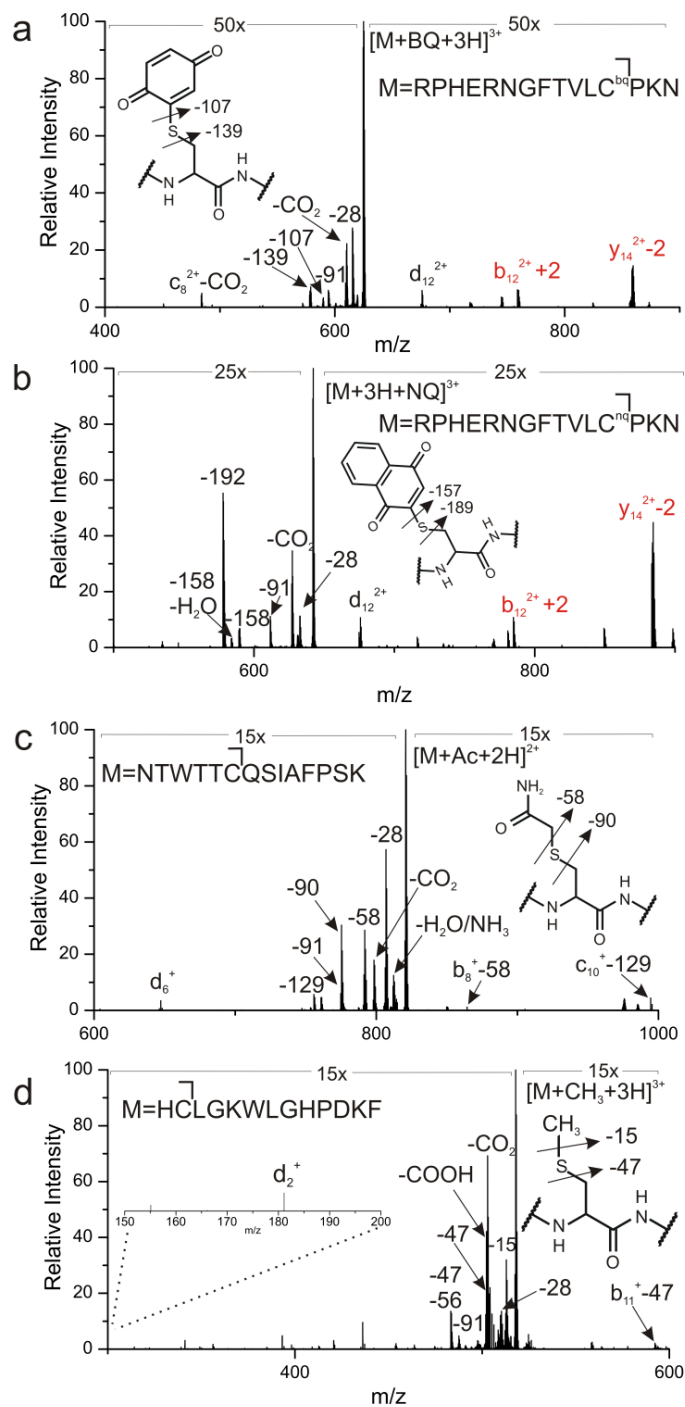
Figure 2.3: UVPD at 213 nm of [RPHERNGFTVLCPKN]³⁺ modified with either a) BQ and b) NQ at the cysteine residue. Red fragments retain the quinone modification. c) UVPD at 213 nm of [NTWTTCQSIAFPSK]²⁺ following modification with iodoacetamide. d) UVPD at 213 nm of [HCLGKWLGHPDKF]³⁺ following modification with iodomethane. Dissociation of C-S bonds is noted in the skeletal structure for each modification. Relevant side chain mass losses are illustrated in the insets.

In addition to quinone chemistry, cysteine residues are often capped with acetamide following thiol reduction in proteomics protocols.[25] Photoactivation of acetamide capped NTWTTCQSIAFPSK is shown in Figure 2.3c, revealing surprisingly abundant C-S bond cleavage. Amide functional groups are weak chromophores at 213 nm, which may account for the enhanced fragmentation. Again, the characteristic d_6 ion is generated and identifies the sequence position of the cysteine residue. Although not a commonly encountered cysteine modification, we also examined methylated cysteine in the peptide HCLGKWLGHPDKF. Methylated cysteine behaves similarly to methionine, with the exception that a modest d-ion is generated. This result is not unexpected because the cysteine side chain is one methylene unit shorter than methionine and will therefore yield an alaninyl radical following C-S bond cleavage. Comparable experiments at 266 nm yield little to no dissociation of the C-S bond.

2.3.1 C-S bonds formed during Identification of Phosphorylation Site Locations

Phosphorylated serine is easily converted into dehydroalanine under basic conditions.[26] Michael addition chemistry can then be used to install C-S bonds that will fragment and yield a beta radical, directing backbone dissociation and identifying the original site of phosphorylation with a characteristic d-ion. Excitation at 213 nm of RQSVQLHsAPQSLPR modified with naphthalene thiol at Ser8 (denoted with small cap s) yields cleavage of both C-S bonds and an abundant d_8 ion, see Figure 2.4a. Unfortunately, the placement of sulfur directly adjacent to the aromatic ring makes naphthalenethiol a poor nucleophile, which leads to significantly decreased reactivity and modification yield. In contrast, spacing the sulfur out from the aromatic ring by one carbon serves to enhance nucleophilicity and reactivity. Therefore, the same modification with benzyl mercaptan proceeds much more efficiently, with yields increasing from ~20% to ~100%. However,

the chromophore is also spaced away from the C-S bond that must be cleaved to yield the alaninyl radical. At 213 nm, this is not a problem, as shown in Figure 2.4b. A reasonable amount of the correct C-S bond is cleaved, and the d_8 ion is easily detected. Similar results are obtained for the benzyl mercaptan modification of a phosphorylated threonine containing peptide, RKRRQtSM, as shown in Figure 2.4c. Importantly, for RKRRQtSM, the modification yield increases from ~0% to ~100% for naphthalene thiol and benzyl mercaptan, respectively. The ability to use the more reactive thiol is therefore a significant advantage for UVPD at 213 nm. This point is illustrated further in Figure 2.4d where photoactivation of the same peptide at 266 nm fails to produce significant yield of the -123 Da loss or a detectable d-ion.
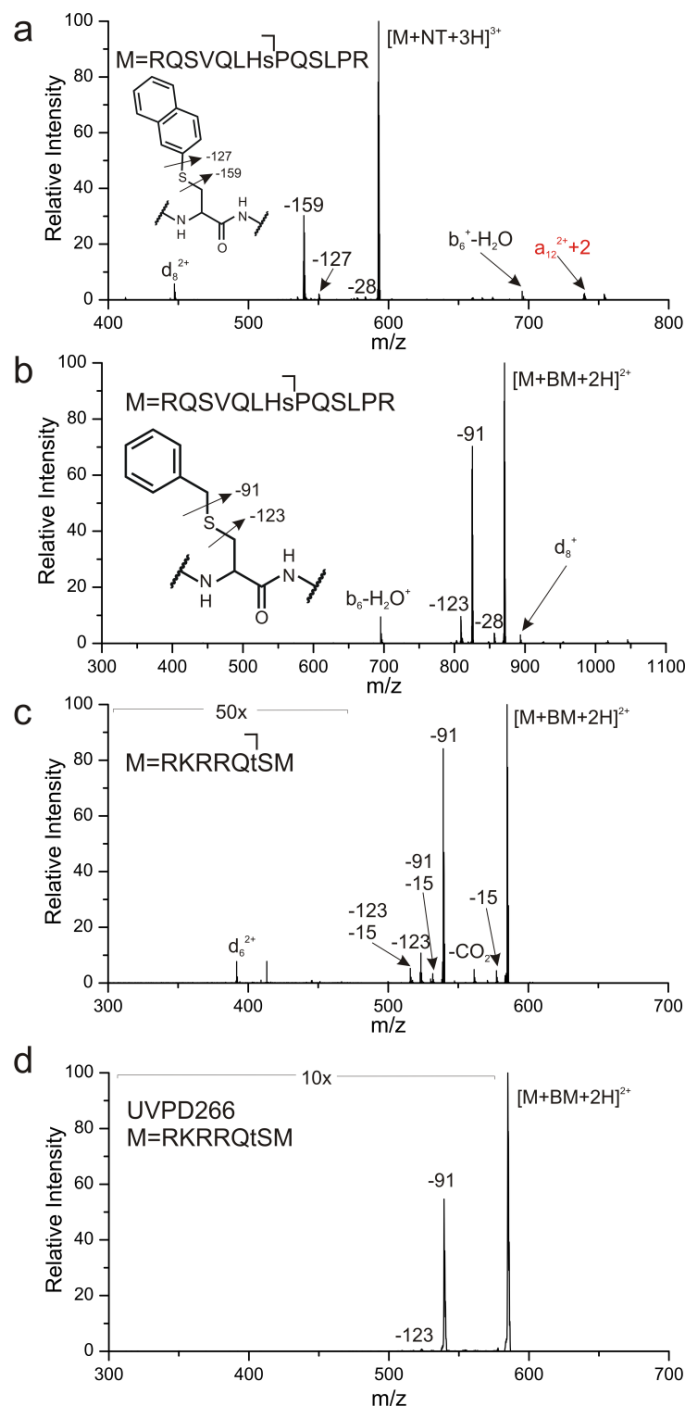
Figure 2.4: a) UVPD at 213 nm of naphthalene thiol modified [RQSVQLHsPQSLPR]$^{3+}$ leads to C-S bond dissociation. b) UVPD at 213 nm of benzyl mercaptan modified [RQSVQLHsPQSLPR]$^{2+}$ and c) [RKRRQtSM]$^{2+}$ yields loss C-S bond cleavage and d-ion formation. d) 266 nm activation of BM modified [RQSVQLHsPQSLPR]$^{2+}$.

2.3.2 Gas Phase Disulfide Identification

Identification of disulfide bond partners in proteins is an active area of interest. Previous results have demonstrated that 266 nm light effectively cleaves S-S bonds, enabling facile identification of peptides linked by disulfide bonds.[17] Similar experiments were conducted at 213 nm and the results are shown in Figure 2.5. Photoactivation of disulfide-bound peptides RGDC and CDPGYIGSR yields the spectrum in Figure 2.5a. Similar to results obtained at 266 nm, S-S bond fracture is favorable. In addition, 213 nm light also cleaves the adjacent C-S bonds, yielding peaks at ±32 Da relative to the S-S bond dissociation products. The end result is a signature 'triplet' that can be used to easily identify the presence of a disulfide bond and the masses of the constituent peptides that were connected by it. In addition, several other UVPD type fragments are observed, such as the y_7+2 ion, characteristic of proline. More importantly, the C-S bond fragmentations again yield beta radicals that can produce signature d-ions, identifying the location of the linking cysteine residues. Indeed, a d_4 ion correctly identifies the location of cysteine in the RGDC peptide. Location of cysteine at the N-terminus prevents observation of the corresponding d-ion for CDPGYIGSR. Examination of disulfide-bound CGYGPKKKRKVGG and SLRRSSCFGGR yields similar results, as shown in Figure 2.5b. The abundance of S-S bond cleavage is higher for this peptide, which may be related to energy transfer from tyrosine.[27] Nevertheless, significant cleavage of C-S bonds is also observed, and an easily detectable d_7 ion is produced, revealing the location of the cysteine residue in SLRRSSCFGGR. Interestingly, the results shown in Figure 2.5 were obtained with a relatively high energy pulse (~0.4mJ, one shot), but repeated excitation at lower pulse energy (~1.5µJ/pulse) yields less C-S bond dissociation (see supporting information). We suspect that the β-carbon and perthiyl radicals generated by C-S bond

cleavage are less stable than the corresponding thiyl radical and undergo additional losses upon repeated exposure to 213 nm light.
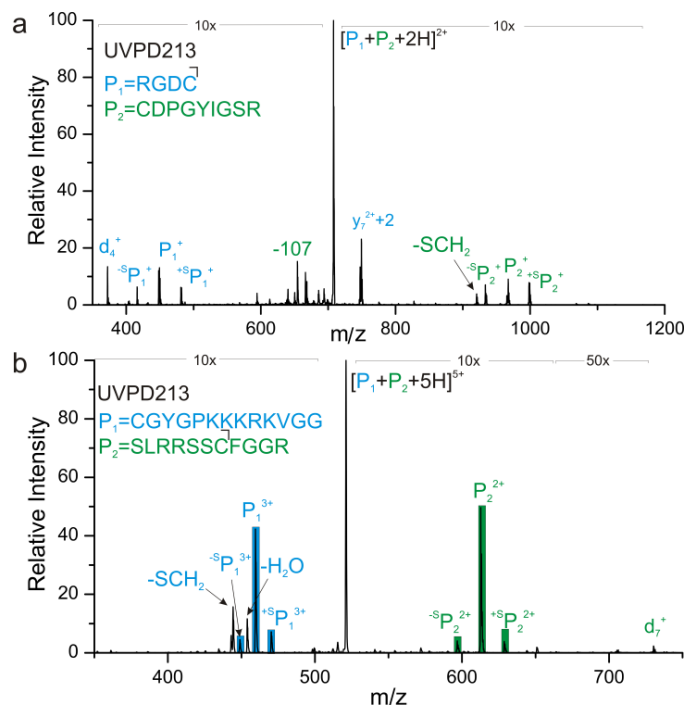
Figure 2.5: UVPD of a) RGDC and CDPGYIGSR b) CGYGPKKKRKVGG and SLRRSSCFGGR each disulfide-linked. Dissociation of the disulfide bond as well as each adjacent C-S bond is identified for each peptide pair.

2.4 Conclusion

Photoactivation of peptides with 213 nm photons can be used to drive both bond-specific dissociation and traditional nonspecific UVPD. 213 nm works well for bond-selective photodissociation of labile bonds previously found to undergo direct dissociation at 266 nm. In addition, many native and non-native carbon-sulfur bonds are labile at 213 nm, including several that generate alaninyl radicals that initiate beta dissociation of the peptide backbone at the radical site. The signature d-ions generated by this mechanism can be used to site-specifically identify the location of cysteine or phosphorylated

40

serine/threonine residues. The expanded list of labile carbon-sulfur bonds also allows for more reactive groups to be used in the derivatization of dehydroalanine, enabling quantitative conversion to a photoactive species. For disulfide bond characterization, signature triplets are generated at 213 nm due to dissociation of both the sulfur-sulfur and carbon-sulfur bonds, facilitating identification. 213 nm is a promising wavelength for the further expansion of bond-specific photodissociation and its applications for biomolecular characterization.

2.5 Acknowledgements

2.6 References

[1] J.S. Brodbelt, Chem. Soc. Rev., 43 (2014) 2757–2783.

[2] J. Mitchell Wells, S.A. McLuckey, A.L. Burlingame, Methods in Enzymology, 2005.

[3] V.H. Wysocki, G. Tsaprailis, L.L. Smith, L.A. Breci, J. Mass Spectrom., 35 (2000) 1399–1406.

[4] F. Tureček, R.R. Julian, Chem. Rev., 113 (2013) 6691–6733.

[5] R. R. Julian, J. Am. Soc. Mass Spectrom., (2017).

[6] T. Ly, R.R. Julian, Ultraviolet Photodissociation: Angew. Chemie-International Ed., 48 (2009) 7130–7137.

[7] T.-Y. Kim, M.S. Thompson, J.P. Reilly, Rapid Commun. Mass Spectrom., 19 (2005) 1657–1665.

[8] J.W. Morgan, J.M. Hettick, D.H.B.T.-M. in E. Russell, in: Biol. Mass Spectrom., Academic Press, 2005: pp. 186–209.

[9] K.M. Choi, S.H. Yoon, M. Sun, J.Y. Oh, J.H. Moon, M.S. Kim, J. Am. Soc. Mass Spectrom., 17 (2006) 1643–1653.

[10] L.J. Morrison, J.S. Brodbelt, J. Am. Chem. Soc., 138 (2016) 10849–10859.

[11] S. Park, W.-K. Ahn, S. Lee, S.Y. Han, B.K. Rhee, H. Bin Oh, Mass Spectrom., 23 (2009) 3609–3620.

[12] A. Theisen, B. Yan, J.M. Brown, M. Morris, B. Bellina, P.E. Barran, Anal. Chem., (2016).

[13] J.J. Wilson, J.S. Brodbelt, Anal. Chem., 79 (2007) 7883–7892.

[14] N.G. Hendricks, R.R. Julian, Analyst, 141 (2016) 4534–4540.

[15] T. Ly, R.R. Julian, J. Am. Chem. Soc., 130 (2008) 351–358.

[16] J.K. Diedrich, R.R. Julian, Anal. Bioanal. Chem., 403 (2012) 2269–2277.

[17] A. Agarwal, J.K. Diedrich, R.R. Julian, Anal. Chem., 83 (2011) 6455–6458.

[18] J.K. Diedrich, R.R. Julian, J. Am. Chem. Soc., 130 (2008) 12212–12213.

[19] J.K. Diedrich, R.R. Julian, Anal. Chem., 82 (2010) 4006–4014.

[20] E. Regoeczi, Iodine-Labeled Plasma Proteins, CRC Press, 1984.

[21] J.W. Scotchler, A.B. Robinson, Anal. Biochem., 59 (1974) 319–322.

[22] Q. Sun, H. Nelson, T. Ly, B.M. Stoltz, R.R. Julian, J. Proteome Res., 8 (2009) 958–966.

[23] D.P. Nair, M. Podgórski, S. Chatani, T. Gong, W. Xi, C.R. Fenoli, C.N. Bowman, Chem. Mater., 26 (2014) 724–744.

[24] M. Girod, Z. Sanader, M. Vojkovic, R. Antoine, L. MacAleese, J. Lemoine, V. Bonacic-Koutecky, P. Dugourd, J. Am. Soc. Mass Spectrom., 26 (2015) 432–443.

[25] R.L. Gundry, M.Y. White, C.I. Murray, L.A. Kane, Q. Fu, B.A. Stanley, J.E. Van Eyk, Curr. Protoc. Mol. Biol., CHAPTER (2009) Unit10.25-Unit10.25.

[26] C. Klemm, S. Schröder, M. Glückmann, M. Beyermann, E. Krause, Rapid Commun. Mass Spectrom., 18 (2004) 2697–2705.

[27] N.G. Hendricks, R.R. Julian, Phys. Chem. Chem. Phys., (2015).

Chapter 3: Novel Methionine and Selenomethionine Energy Transfer Systems for Biomolecular Structure Elucidation in the Gas Phase

3.1 Introduction

There is significant interest in the development of new methodology for the characterization of biological molecules in both a rapid and sensitive fashion. Mass spectrometry (MS) offers an avenue to accomplish these goals and, as such, has been used extensively for protein identification, sequencing, and quantitation.[1-5] MS technology is now widely available due to the popularity of proteomics, which has spurred efforts to develop methods capable of higher order structure determination. Ion mobility-MS is the most commonly used gas phase method capable of providing information about three-dimensional structure.[6-12] Ion mobility yields structural data in the form of an averaged collision-cross sections for each ion. Ion mobility is fairly easy to implement and can examine a wide range of molecular targets, but it is not always possible to completely constrain molecular structure with a single parameter. Recent developments have expanded the parameter space of ion mobility by examining cross sections as a function of activation, which can provide clues about substructural elements and organization.[13]

Spectroscopy can also be coupled with MS for structural characterization, including both ultraviolet (UV) and infrared (IR) based approaches.[14,15] Common condensed-phase techniques like Förster Resonance Energy Transfer (FRET) can be employed in order to obtain structural information about the peptide or protein of interest. Traditional FRET reveals the distance between a donor and acceptor, typically between 10-100 Å, by exciting the donor and monitoring emission of the acceptor.[16] Work by the Jockusch and Zenobi groups has demonstrated that the FRET-based approach with fluorescence detection can be successfully implemented in the gas phase within the confines of a mass

spectrometer.[17-19] It is also possible to utilize action spectroscopy-based experiments in a mass spectrometer where absorption of photons is assumed to correlate with ion dissociation. Action spectra are generated by scanning the excitation wavelength while monitoring fragment intensity. By coupling energy transfer with action spectroscopy, donor/acceptor distance constraints can be determined without detecting photons directly.[20] Dugourd and coworkers have used this approach to examine small peptides and ubiquitin, modified with FRET chromophores.[21,22]

Our lab has developed an action spectroscopy system that relies on native chromophoric amino acids as donors and propylmercaptan (PM) modified cysteine as the acceptor. Energy transfer in this system occurs over short distances, smaller than the typical range for FRET, so the more general term excitation energy transfer (EET) has been used to describe the experiment. Excitation of peptides containing either tyrosine or tryptophan yields minimal nonspecific fragmentation at 266 nm. However, if either residue is spatially located near a PM-modified cysteine, energy transfer to the disulfide results in homolytic cleavage in a distance dependent fashion.[23,24] Experiments with molecules of known structure revealed that energy transfer can occur at <15Å for tryptophan and <6Å for tyrosine. Phenylalanine, while an aromatic residue, does not directly transfer energy to disulfide bonds efficiently. However, phenylalanine can participate in a two-step process where energy transfer passes through an intermediary tyrosine residue.[25] Recently Rizzo and coworkers also examined energy transfer from electronically excited phenylalanine to tyrosine in a cryogenic trap.[26] This system represents a close analog of the action-EET experiments we have conducted, and it was determined that energy transfer is best represented by the full Coulomb interaction between donor and acceptor. The dipole-

based FRET approximation was insufficient to describe the results, and Dexter exchange pathways were found to be negligible.

The addition of propyl mercaptan to a free thiol on cysteine is a mild and relatively easy-to-carry-out modification, but one that becomes more challenging with increasing protein size and complexity. The presence of native disulfides can introduce the potential for disulfide scrambling, particularly if the protein does not have a cysteine in the free thiol state natively, which is not particularly common. This prompted an investigation into other native amino acids that could act as acceptors in an entirely native action-EET system. Recently, we reported that 213 nm photons homolytically cleave C-S bonds in both cysteine and methionine.[27] Although S-S bonds are inherently cleavable, photolysis of C-S bonds is not typically observed at 266 nm unless the C-S bond comprises part of a more suitable chromophore.[28,29] However, the 213 nm experiments revealed that C-S bonds harbor dissociative electronic excited states that might be accessible via energy transfer under appropriate conditions.

Methionine therefore warrants additional attention because it contains C-S bonds but not disulfide bonds and is not subject to disulfide scrambling. It is an essential amino acid in humans and the most common start codon for protein synthesis. Methionine can also behave as an antioxidant.[30,31] A common amino acid analogue of methionine, selenomethionine (SeMet), also exists in nature and only differs by a single atom, selenium in place of sulfur. SeMet is also biologically relevant and is the major dietary source of selenium for many living organisms, where selenium is both essential and toxic. Uptake of SeMet is typically followed by conversion to selenocysteine, the 21st amino acid, and incorporation into proteins.[32] SeMet is also routinely incorporated abiologically into proteins for phasing in X-ray crystallography.[33,34] Sulfur and selenium share many physical

and chemical properties including similar electronegativity, ionic radii, and bond length.[35] Important chemical differences include: carbon-selenium bonds are weaker than carbon-sulfur bonds, and they have lower lying σ* anti-bonding orbitals.[36] These differences in bonding could lead to significantly altered energy-transfer efficiency.

Herein we examine the potential for Met and SeMet to act as native amino acid energy-transfer acceptors. Data on model peptides shows that both C-S bonds in Met dissociate following energy transfer from all three aromatic amino acid donors: phenylalanine, tyrosine, and tryptophan. Energy transfer is most efficient for tyrosine, followed by phenylalanine and then tryptophan. While energy transfer is observed, the C-S bonds are generally poor energy acceptors and can only accommodate low PD yields. In contrast, C-Se bond dissociation following energy transfer to SeMet is significantly more abundant. Donor efficiency for SeMet follows the order Trp>Tyr>>Phe. To further explore the structure probing capabilities of SeMet, a series of helical peptides were synthesized where the relative donor/acceptor positions were systematically varied. The trends in bond dissociation followed the characteristic pattern expected for a helical peptide, confirming that SeMet is a viable acceptor for probing peptide structure in the gas phase.

3.2 Experimental Methods

3.2.1 Materials and Peptide Synthesis

Methanol was purchased from Fisher Scientific (Waltham, MA). Water was purified by Millipore Direct-Q (Millipore, Billerica, MA). Fmoc-SeMet was purchased from Matrix scientific (Columbia, Sc). Polyalanine peptides were synthesized with the following sequences: Ac-A_5XA_3MK and Ac-$A_5XA_3M^{Se}K$ (where X = Y, F or W), A_3MK, and $A_3M^{Se}K$. Additionally, a series of SeMet helical peptides were synthesized containing either tyrosine or tryptophan with the following sequences: Ac-$XA_8M^{Se}K$, Ac-$AXA_7M^{Se}K$, Ac-$A_2XA_6M^{Se}K$,

Ac-$A_3XA_5M^{Se}K$, Ac-$A_4XA_4M^{Se}K$, Ac-$A_5XA_3M^{Se}K$, (X = Y or W). A series of peptides with tryptophan near the C-terminus were also made with the following sequences: Ac-$M^{Se}A_8WK$, Ac-$AM^{Se}A_7WK$, Ac-$A_2M^{Se}A_6WK$, Ac-$A_3M^{Se}A_5WK$, Ac-$A_4M^{Se}A_4WK$, and Ac-$A_5M^{Se}A_3WK$. Peptides were synthesized following standard solid phase peptide synthesis procedures for most amino acid additions.[37] The coupling and deprotection steps were each given six minutes for each amino acid addition. The SeMet coupling step was extended to an hour and a half and deprotection was completed in six minutes prior to the addition of the next amino acid in order to improve peptide synthesis yields.

3.2.2 Photodissociation of Peptides

An LTQ linear ion trap mass spectrometer (Fisher Scientific, Waltham, MA) with a standard ESI source was utilized for photodissociation experiments. A quartz window was installed on the back plate of the LTQ for transmission of laser pulses from a flash-lamp pumped Nd:YAG Minilite laser (Continuum, Santa Clara, CA). A single 266 nm pulse from the 4 mJ laser was synchronized to occur at the activation step of the MS^2 experiment, triggered by the instrument. Peptides were dissolved in 50:50 water:MeOH in 0.1% TFA at 10 μM concentrations and electrosprayed at 3 μL/min with a voltage of 3.25 kV and a capillary inlet temperature set to 215 ºC.

For action-EET experiments, an optical parametric oscillator (OPO) laser (Opotek Inc., Carlsbad, CA) was utilized. Ions were isolated and subjected to a single laser pulse at wavelengths varying from 250-300 nm. Laser power levels for each wavelength were recorded and normalized. The OPO laser power is typically one-fourth of the Nd:YAG at 266 nm, which is reflected when comparing the PD yields between the two setups. Action spectra were compiled using the PD yield at each wavelength. Instrument and peptide conditions were the same as the 266 nm photodissociation experiments described above.

3.2.3. Simulated Annealing and Molecular Dynamics Simulations

Simulated annealing and molecular dynamics (MD) simulations were performed with MacroModel (Schrödinger Inc., Portland, Oregon) using the OPLS2005 atomic force fields. Simulated annealing was first performed by heating the starting structure up to 1000 K and then slowly cooled to 200 K. This process was repeated until 1000 structures were obtained yielding a population of potential structures. The lowest energy structure was selected, and a subsequent MD simulation was performed. For the simulation, 1000 structures were obtained under the following conditions; 300K with a 1.0 ps equilibration time, 1.5 fs step time and a 5 ns simulation time. To estimate the distance between the donor and acceptor, distances were measured from all atoms located on the aromatic rings, of either tyrosine or tryptophan, to one of the two carbon atoms adjacent to selenium in the SeMet side chain (CH_3-Se/Se-CH_2) as well as to the selenium atom. The distances were sorted and the shortest distances were averaged to represent the closest point of contact for each 1000 structure MD simulation.

3.3 Results and Discussion

To test the feasibility of energy transfer to the C-S bonds of methionine, helical peptides with the sequence Ac-A$_5$XA$_3$MK were synthesized where X was either tyrosine, tryptophan, or phenylalanine. These peptides have been well characterized in previous studies to maintain their alpha-helical structure during transition from liquid to the gas phase and place the chromophore in close proximity to the position occupied by methionine.[23,38] Figure 3.1a and 3.1b show excitation at 266 nm of the tyrosine and tryptophan containing helical peptides, respectively. Ac-A$_5$YA$_3$MK primarily dissociates at the methionine side chain C-S bonds, yielding mass losses of -15 and -47 Da. The loss of -75 Da represents a secondary product ion formed following loss of ethylene from the

original -47 Da product ion. The loss of -107 Da is not related to energy transfer and results from direct dissociation of the bond connecting the tyrosine side chain. Analogous results (where -130 Da corresponds to the tryptophan side chain loss) are obtained for Ac-A$_5$WA$_3$MK with additional fragments including loss of CO_2 and a series of b- and y-ions. The sequence ions result from internal conversion of the photon energy into heat. The CO_2 (-44 Da) loss likely results hydrogen atom abstraction from the C-terminal -COOH.[39] Figure 3.1c shows results for 266 nm excitation of Ac-A$_5$FA$_3$MK. Interestingly, dissociation of the methionine side chain is observed. While phenylalanine is a weak chromophore at 266 nm, there is sufficient absorption and energy transfer to the C-S bonds in methionine to clearly generate the -15 and -47 Da losses.[25] The secondary loss of -75 Da is not observed, which may suggest any conversion of additional photons to heat with phenylalanine is insufficient to cause fragmentation on the timescale of these experiments.. Photoactivation of a peptide containing no aromatic residues, A$_3$MK, is shown in Figure 3.1d. No dissociation of the methionine side chain is observed, indicating that the C-S bonds do not absorb 266 nm photons on their own.
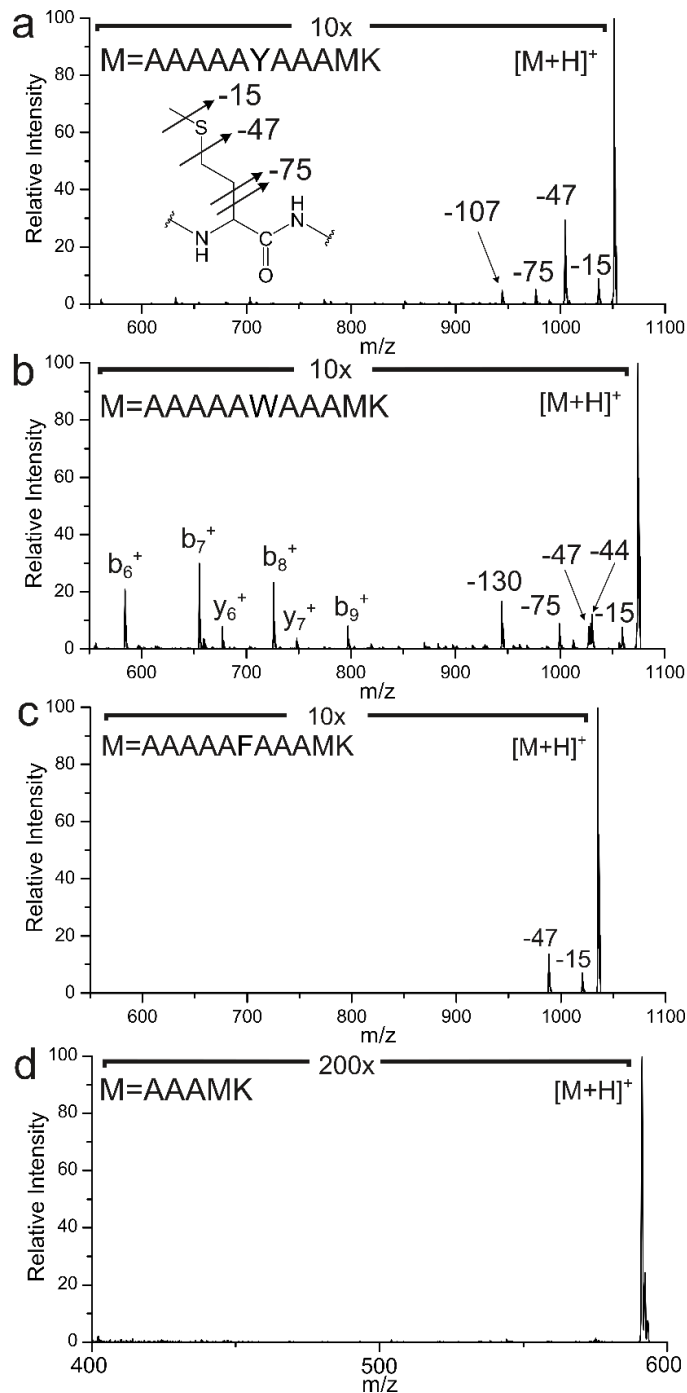
Figure 3.1: Photoactivation of A_5XA_3MK peptides at 266 nm, where X is a) tyrosine, b) tryptophan or c) phenylalanine. The fragmentation pathways for the methionine side chain are shown in the inset graphic, where a double arrow indicates a sequential double loss. d) PD of A_3MK with no observable methionine side chain dissociation.

50

To explore whether the dissociation observed in Figure 3.1 was due to energy transfer, action spectra were collected and are shown in Figure 3.2. The PD yields are a summation of the two primary losses (-15 and -47 Da) that result from excited-state C-S bond cleavage. The spectra clearly reflect the characteristic features for absorption of phenylalanine, tyrosine, and tryptophan in the range of 250-300 nm.[40] The pink trace exhibits the two characteristic peaks for tyrosine around 275 and 285 nm, while the blue trace for tryptophan is broader with two distinctive features at 283 and 287 nm. The red trace shows a triplet of peaks between 250-265 nm that correspond well with the known absorption profile for phenylalanine. In contrast, the methionine-only peptide (green trace) does not yield any appreciable amount of C-S bond cleavage. The action spectra in Figure 3.2 provide strong evidence that dissociation of C-S bonds in these peptides is due to energy transfer. From the perspective of PD yield, tyrosine is the most efficient energy donor at 266 nm, while Trp and Phe only net approximately half the energy transfer of Tyr. Overall, energy transfer to methionine is not particularly efficient and the PD yields for all chromophores are small. Although this does not present a problem for examination of small peptide systems, it is not optimal as molecular scale increases to larger proteins where many other competing fragmentation pathways, though they may also be minor, are likely to compete.
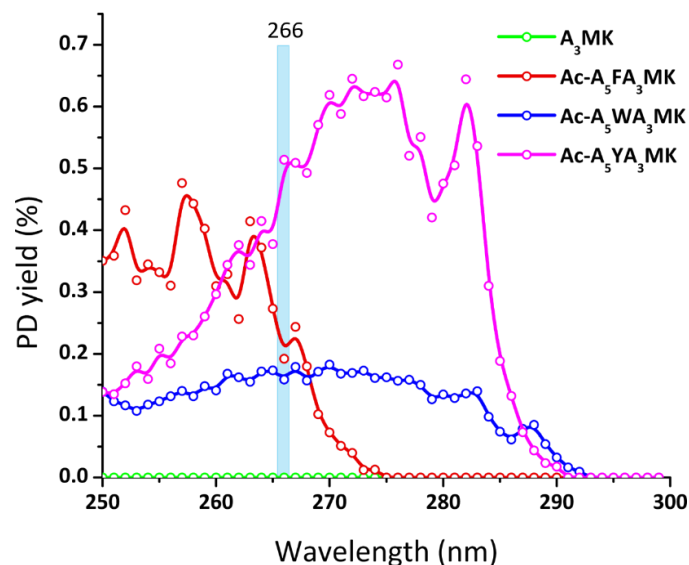
Figure 3.2: Action spectra representing the sum of -15 Da and -47 Da losses for each respective peptide.

Given the limited success of C-S bonds as energy acceptors for action-EET, we explored alternative forms of methionine to see if more abundant energy transfer and PD yield could be achieved. Figure 3.3 shows the results for photoexcitation of a series of synthetic helical peptides incorporating SeMet into the generic sequence Ac-$A_5XA_3M^{Se}K$, where X = tyrosine, tryptophan or phenylalanine. Fragments analogous to those observed for methionine are also detected for SeMet, but the PD yields are much higher for Tyr and Trp (~10x for Tyr, ~80x for Trp). However, energy transfer from Phe to SeMet does not appear to be a favorable process and yields minimal C-Se bond dissociation. The relative amount of backbone fragmentation to side chain fracture is significantly less for Ac-$A_5WA_3M^{Se}K$ compared to its methionine counterpart in Figure 3.1b. This also suggests more efficient energy transfer to C-Se bonds, resulting in less vibrational excitation following internal conversion. Examination of the $A_3M^{Se}K$ peptide reveals that, in the absence of aromatic residues, only slight dissociation of C-Se bonds is observed. Overall,

energy transfer to C-Se bonds appears to be much more favorable than what is observed in methionine C-S bonds.
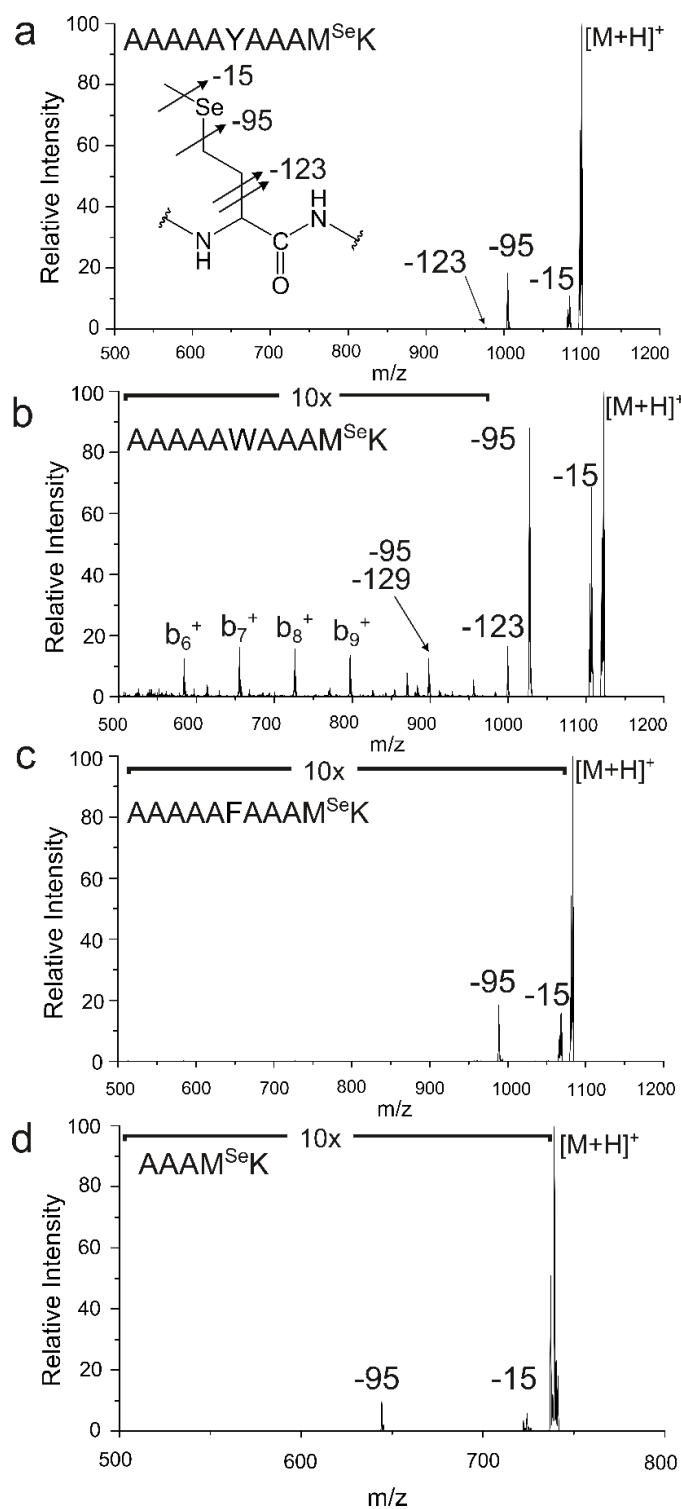
Figure 3.3: Photoactivation of Ac-A$_5$XA$_3$M^{Se}K peptides at 266 nm, where X is a) tyrosine, b) tryptophan or c) phenylalanine. d) Same experiment with A$_3$M^{Se}K.

Figure 3.4 shows the action spectra collected between 250-300 nm for the peptides examined in Figure 3.3. Again, the action spectrum for each chromophore correlates strongly with the known absorption profile. Even the phenylalanine-containing peptide, which despite exhibiting very weak energy transfer, yields a spectrum with the characteristic triplet of peaks (see inset). In contrast, the action spectrum for SeMet itself, as observed in the peptide $A_3M^{Se}K$, reveals no characteristic features and simply decays with increasing wavelength. This is similar to what is observed for the action spectrum of disulfide bonds in the absence of other chromophores. It is likely that the lower-lying C-Se σ^* anti-bonding orbital facilitates energy transfer from the excited states accessible in this wavelength region, particularly in the case of tryptophan. The greatly increased PD yields for SeMet suggest that it could serve as a viable energy acceptor for structure exploration in proteins if the energy transfer correlates with distance.
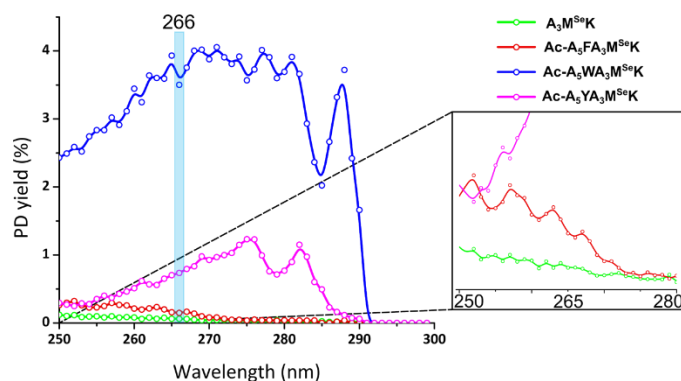
Figure 3.4: Action spectra showing the sum of -15 Da and -95 Da losses for each respective peptide.

To investigate the relationship between C-Se bond energy transfer efficiency and distance, a series of helical peptides containing SeMet and tyrosine were synthesized. The generic sequence corresponds to Ac-$A_xYA_yM^{Se}K$, where x and y vary from 0-5 and 3-8 in concert, respectively. Individual PD yields for the -15, -95, and -123 Da losses are shown in Figure 3.5a for the entire peptide series. The abundance of the two-step -123

Da loss is low for all peptides. The two losses corresponding to C-Se bond cleavage (-15 and -95 Da) are more abundant and generally comparable in intensity, although some variation is noted. However, the overall intensity of these losses changes significantly as the position of the tyrosine chromophore is scanned through the sequence. Importantly, the intensity drops initially, but then increases again for Ac-$A_2YA_6M^{Se}K$, which is consistent with the expected distances between Tyr and SeMet for a helical structure. This is illustrated visually for an idealized alpha-helix in Figure 3.5b from the end-on perspective. To provide quantitative values, distances were measured during molecular dynamics (MD) simulations. Simulated annealing was first used to examine a broad population of potential structures and identify low-energy candidates. The lowest energy structures exhibited alpha helical motifs in agreement with the data in Figure 3.5 and previous studies.[23,41] Average distances extracted from subsequent MD simulations at 300K are shown in inverse form in Figure 3.5a. Although the relative differences for each peptide do not match exactly, similar trends are observed for the experimentally derived PD yields and the computationally derived inverse distances between residues. These results confirm that the peptide adopts a helical form in the gas phase and that probability for energy transfer to C-Se bonds scales with distance.
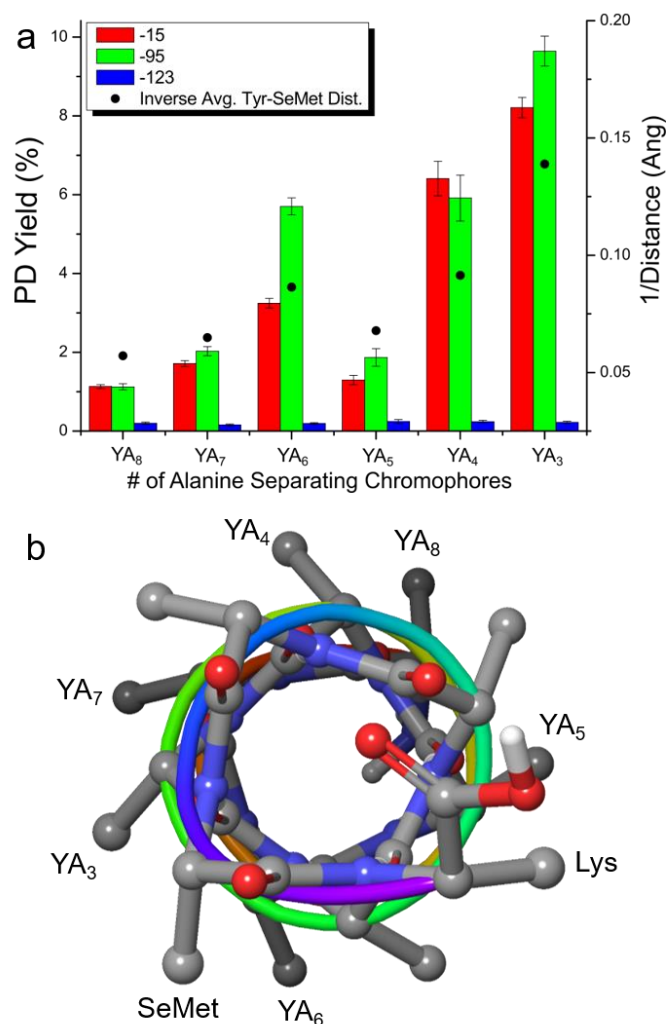
Figure 3.5: a) PD yields for Ac-A_xYA$_y$M^{Se}K peptides, where x and y range from 0-5 and 3-8, respectively. The black dots refer to the inverse of the average distance measured between SeMet and tyrosine obtained from MD simulations. b) Structure of an alpha helix, with C-terminus projecting toward viewer. The positions of Lys and SeMet are labeled and remain fixed for all peptides. The relative position of the tyrosine residue changes for each peptide and is shown with the corresponding YA$_y$ label. It is clear that YA$_3$ and YA$_6$ locate SeMet and Tyr in the closest proximity.

The results in Figure 3.3 and Figure 3.4 reveal that Trp donates energy more efficiently to C-Se bonds than Tyr. The distance dependence for energy transfer from Trp was examined with an analogous series of peptides, Ac-A_xWA$_y$M^{Se}K, and the results are shown in Figure 3.6a. The overall trends in Figure 3.6a are similar to those obtained previously

with Tyr as the donor, but there are several differences as well. Most notably, the overall PD yield is higher for all peptides, suggesting again that energy transfer is more efficient with Trp. In addition, the abundance of the -123 Da sequential loss peak is higher, suggesting that some additional photons are converted into heat. Although the ratio of the -95 Da to -15 Da peaks occasionally inverts for Tyr, the -95 Da loss is always more abundant for Trp. The factors that account for this difference are not immediately clear. Previous work examining energy transfer between phenylalanine and tyrosine revealed that the relative chromophore orientation could influence energy transfer efficiency by roughly a factor of ~2.[26] To explored potential orientation effects, we examined the inverted peptide series, Ac-A_xM^{Se}A$_y$WK, where the position of SeMet was scanned through the helix. The results are shown in Figure 3.6b. The data in Figure 3.6a and 6b are similar, particularly beyond the M^{Se}A$_5$ position. The orientation of the donor and acceptor to the peptide is reversed between these series, making it likely that the relative orientation or orientations of the donor and acceptor are different as well. If present, these differences do not significantly change the degree of energy transfer as measured in PD yield. It is possible that this lack of sensitivity to orientation occurs because these ions are at or near room temperature where local dynamic motion may average out such effects. In any case, the data suggest that both Tyr and Trp can be used with SeMet for probing structure by distance-dependent energy transfer.
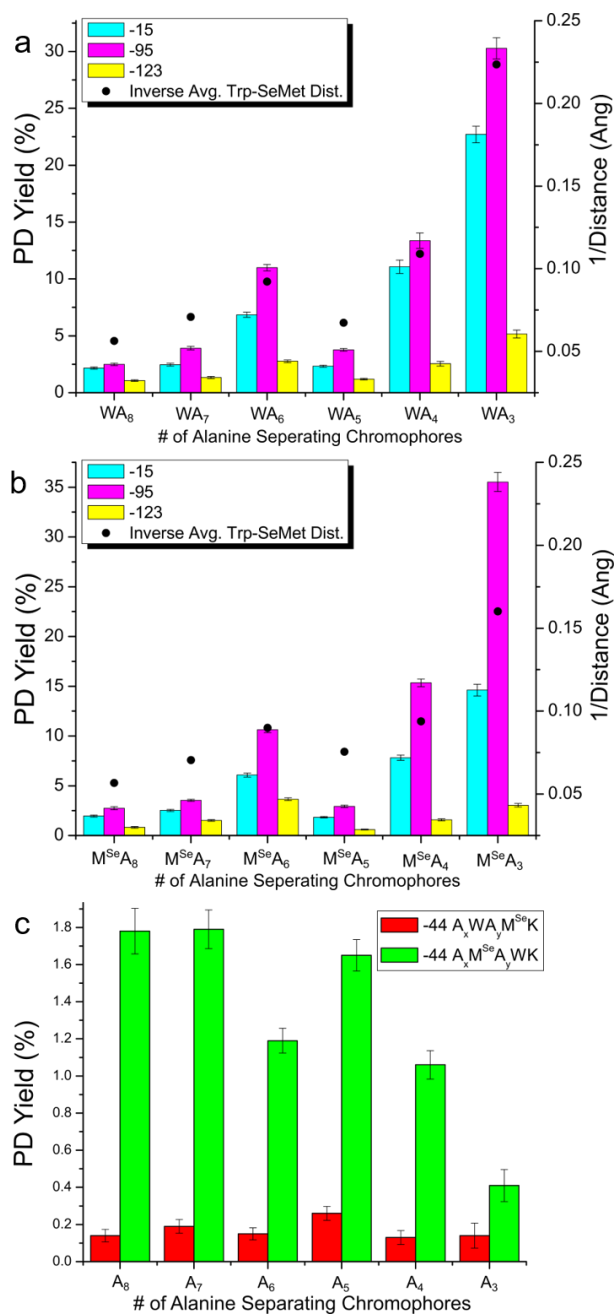
Figure 3.6: a) PD yields for tryptophan containing peptides with the sequence of Ac-$A_xWA_yM^{Se}K$ where x and y refer to the number of alanines ranging from 0-5 and 3-8 respectively. b) PD yields for the reversed Ac-$A_xM^{Se}A_yWK$ peptide series. MD and molecular dynamics simulations were used to obtain average distance between the tryptophan and SeMet side chains. Inverse of the average distances are plotted as black dots in both (a) and (b). c) PD yields for the loss of CO_2 observed for the two tryptophan containing peptide series Ac-$A_xWA_yM^{Se}K$ (red) and Ac-$A_xM^{Se}A_yWK$ (green). The x-axis refers to the number of alanines, separating SeMet and tryptophan.

Placement of Trp next to the C-terminus also increases the loss of -44 Da, corresponding to loss of CO_2. Furthermore, the abundance of the CO_2 loss correlates with donor/acceptor distance in a reciprocal fashion. Essentially, this suggests that excitation of the Trp is quenched by competing pathways, either energy transfer to the SeMet side chain or by reaction with the C-terminus. As the SeMet side chain is moved further away, the probability for reaction with the C-terminus increases. Given that the peptides are all singly charged at the lysine residue with the N-terminus acetylated, the C-terminus is most likely in the –COOH state, necessitating loss of hydrogen prior to loss of CO_2. The excited state Trp may abstract a hydrogen atom, similar to previous observations with electronically excited benzophenone.[39] Since this is a chemical reaction that may require proper orientation of the Trp sidechain and C-terminus, it is less likely to compete with fast energy transfer when the SeMet sidechain is also in close proximity. However, if Trp is not adjacent to the C-terminus, the probability for CO_2 loss is low and fairly uniform, indicating a pathway independent of direct interaction with the chromophore. Although the data in Figure 3.6c make it tempting to consider CO_2 loss as an additional potential reporter for structural information, the competing background channel would make it difficult to implement in more complicated systems.

3.4 Conclusion

We have examined methionine and SeMet as potential energy acceptors in action-EET experiments. When studying simple helical peptide models, the C-S bonds of methionine can be coupled with tryptophan, tyrosine, and phenylalanine as energy donors. As the complexity and size of the target molecule increases, the low observed yields of C-S bond dissociation may become difficult to observe against the background of more common losses like NH_3 (-17 Da), H_2O (-18 Da), CO_2 (-44 Da), and CO_2H (-45 Da). In contrast, the

C-Se bonds from SeMet dissociate with significantly higher yields when adjacent to either tyrosine or tryptophan. A series of helical peptides containing either tyrosine or tryptophan revealed that energy transfer to SeMet is efficient within ~15 Å from the acceptor. Swapping the positions of the donor and acceptor resulted in small changes to dissociation yields, suggesting that orientational effects may not be significant in these room-temperature experiments. Incorporation of SeMet into proteins is not difficult because it is essentially recognized as methionine during protein synthesis in cells.[42] These experiments pave a path forward for future examination of protein structure in the gas phase by action-EET relying on essentially native amino acid donor/acceptor pairs.

3.5 Acknowledgements

3.6 References

[1] X. Han, A. Aslanian, J.R. Yates, Curr. Opin. Chem. Biol. 12 (2008) 483–490.

[2] P. Picotti, M. Clément-Ziza, H. Lam, D.S. Campbell, A. Schmidt, E.W. Deutsch, H. Röst, Z. Sun, O. Rinner, L. Reiter, Q. Shen, J.J. Michaelson, A. Frei, S. Alberti, U. Kusebauch, B. Wollscheid, R.L. Moritz, A. Beyer, R. Aebersold, Nature, 494 (2013) 266.

[3] P.J. Boersema, R. Raijmakers, S. Lemeer, S. Mohammed, A.J.R. Heck, Nat. Protoc., 4 (2009) 484.

[4] S.D.H. Shi, M.E. Hemling, S.A. Carr, D.M. Horn, I. Lindh, F.W. McLafferty, Anal. Chem., 73 (2001) 19–22.

[5] T.P. Cleland, C.J. DeHart, R.T. Fellers, A.J. VanNispen, J.B. Greer, R.D. LeDuc, W.R. Parker, P.M. Thomas, N.L. Kelleher, J.S. Brodbelt, J. Proteome Res., 16 (2017) 2072–2079.

[6] B.C. Bohrer, S.I. Merenbloom, S.L. Koeniger, A.E. Hilderbrand, D.E. Clemmer, Annu. Rev. Anal. Chem. (Palo Alto. Calif)., 1 (2008) 293–327.

[7] S.R. Harvey, C.E. MacPhee, P.E. Barran, Methods, 54 (2011) 454–461.

[8] M.F. Bush, Z. Hall, K. Giles, J. Hoyes, C. V Robinson, B.T. Ruotolo, Anal. Chem., 82 (2010) 9557–9565.

[9] J.A. McLean, B.T. Ruotolo, K.J. Gillig, D.H. Russell, Int. J. Mass Spectrom., 240 (2005) 301–315.

[10] N.A. Pierson, L. Chen, S.J. Valentine, D.H. Russell, D.E. Clemmer, J. Am. Chem. Soc., 133 (2011) 13810–13813.

[11] T. Wyttenbach, M.T. Bowers, J. Phys. Chem. B. 115 (2011) 12266–12275.

[12] J.C. May, C.R. Goodwin, N.M. Lareau, K.L. Leaptrot, C.B. Morris, R.T. Kurulugama, A. Mordehai, C. Klein, W. Barry, E. Darland, G. Overney, K. Imatani, G.C. Stafford, J.C. Fjeldsted, J.A. McLean, Anal. Chem., 86 (2014) 2107–2116.

[13] Y. Zhong, L. Han, B.T. Ruotolo, Angew. Chemie Int. Ed., 53 (2014) 9209–9212.

[14] S. Warnke, C. Baldauf, M.T. Bowers, K. Pagel, G. von Helden, J. Am. Chem. Soc., 136 (2014) 10308–10314.

[15] N.C. Polfer, J. Oomens, Mass Spectrom. Rev. 28 (2009) 468–494.

[16] A. Periasamy, Fluorescence resonance energy transfer microscopy: a mini review, J. Biomed. Opt., 6 (2001) 287–291.

[17] M.W. Forbes, R.A. Jockusch, J. Am. Soc. Mass Spectrom., 22 (2011) 93–109.

[18] M.F. Czar, F. Zosel, I. König, D. Nettels, B. Wunderlich, B. Schuler, A. Zarrine-Afsar, R.A. Jockusch, Anal. Chem., 87 (2015) 7559–7565.

[19] M. Dashtiev, V. Azov, V. Frankevich, L. Scharfenberg, R. Zenobi, J. Am. Soc. Mass Spectrom., 16 (2005) 1481–1487.

[20] N.G. Hendricks, R.R. Julian, Analyst, 141 (2016) 4534–4540.

[21] S. Daly, F. Poussigue, A.-L. Simon, L. MacAleese, F. Bertorelle, F. Chirot, R. Antoine, P. Dugourd, Action-FRET: Probing the Molecular Conformation of Mass-Selected Gas-Phase Peptides with Förster Resonance Energy Transfer Detected by Acceptor-Specific Fragmentation, Anal. Chem., 86 (2014) 8798–8804.

[22] S. Daly, G. Knight, M.A. Halim, A. Kulesza, C.M. Choi, F. Chirot, L. MacAleese, R. Antoine, P. Dugourd, J. Am. Soc. Mass Spectrom., 28 (2017) 38–49.

[23] N.G. Hendricks, N.M. Lareau, S.M. Stow, J.A. McLean, R.R. Julian, J. Am. Chem. Soc., 136 (2014) 13363–13370.

[24] N.G. Hendricks, R.R. Julian, Phys. Chem. Chem. Phys. (2015).

[25] N.G. Hendricks, R.R. Julian, Chem. Commun., 51 (2015) 12720–12723.

[26] V. Scutelnic, A. Prlj, A. Zabuga, C. Corminboeuf, T.R. Rizzo, J. Phys. Chem. Lett. 9 (2018) 3217–3223.

[27] L.E. Talbert, R.R. Julian, J. Am. Soc. Mass Spectrom., (2018).

[28] J.K. Diedrich, R.R. Julian, Anal. Chem. 83 (2011) 6818–6826.

[29] J.K. Diedrich, R.R. Julian, Anal. Chem. 82 (2010) 4006–4014.

[30] M.S.P. Soares, P.S. Oliveira, G.N. Debom, B. da Silveira Mattos, C.R. Polachini, J. Baldissarelli, V.M. Morsch, M.R.C. Schetinger, R.G. Tavares, F.M. Stefanello, R.M. Spanevello, Amino Acids, 49 (2017) 129–138.

[31] O.M. Hamdy, A. Alizadeh, R.R. Julian, Analyst, 140 (2015) 5023–5028.

[32] D. Hatfield, A. Diamond, Trends Genet., 9 (1993) 69–70.

[33] R.P.-A. Berntsson, N. Alia Oktaviani, F. Fusetti, A.-M.W.H. Thunnissen, B. Poolman, D.-J. Slotboom, Protein Sci., 18 (2009) 1121–1127.

[34] W.A. Barton, D. Tzvetkova-Robev, H. Erdjument-Bromage, P. Tempst, D.B. Nikolov, Protein Sci., 15 (2006) 2008–2013.

[35] L.A. Wessjohann, A. Schneider, M. Abbas, W. Brandt, Biol. Chem., 388 (2007) 997.

[36] H.J. Reich, R.J. Hondal, ACS Chem. Biol., 11 (2016) 821–841.

[37] W.C. Chan, P.D. White, 2004.

[38] M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, M. Scheffler, J. Phys. Chem. Lett., 1 (2010) 3465–3470.

[39] R.E. Bossio, R.R. Hudgins, A.G. Marshall, J. Phys. Chem. B, 107 (2003) 3284–3289.

[40] H. Mach, G. Sanyal, D.B. Volkin, C.R. Middaugh, in: Ther. Protein Pept. Formul. Deliv., American Chemical Society, 1997: pp. 11–186.

[41] R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc., 121 (1999) 3494–3501.

[42] L. Ouerdane, Z. Mester, J. Agric. Food Chem. 56 (2008) 11792–11799.

Chapter 4: Synthesis of New S-S and C-C Bonds by Photoinitiated Radical Recombination Reactions in the Gas Phase

4.1 Introduction

Manipulation of ions within mass spectrometers is more often used to break bonds than to make them. However, the synthesis of new bonds affords extraction of novel information such as structural relationships or relative reactivity, which has motivated a limited number of studies. For example, irradiation of photoleucine can lead to covalent coupling of peptides by insertion of reactive carbenes.[1] Carbenes can also be generated by collisional activation of diazo groups, even when molecules to be coupled are only held together by noncovalent bonds.[2] Carbene chemistry is advantageous in that covalent links are easily formed by insertion to a large variety of bonds, but at the same time such crosslinks are dictated primarily by proximity and afford poor chemical selectivity. Covalent bonds have also been synthesized by a variety of ion/ion reactions pioneered by the McLuckey lab that rely on reactive functional groups and do afford selectivity.[3-5] New bonds can also be created by accessing excited state reactivity, as demonstrated in experiments where vacuum ultraviolet light led to condensation reactions between peptide fragments.[6] Taken together, these studies have revealed that covalent bond formation is possible in the gas phase if reactive species can be brought into sufficiently close proximity for enough time to allow the coupling to proceed.

Radicals are another highly reactive group with the potential to form covalent bonds. Both the recombination of two radicals or insertion of a radical into a suitable bond followed by elimination of a different radical can lead to the formation of new covalent linkages. Indeed, radical recombination of separately generated radical species was previously used to probe the conformational states of polyproline peptides.[7] Similar to photoleucine,

chemistries are available for the highly specific creation of radicals by photoactivation of ultraviolet (UV) labile groups.[8,9] Both open shell carbon and sulfur atoms can be created at atomically precise locations in this fashion with little additional heating of the remaining molecule.[10] Radicals can also be generated by collisional activation, which can lead to intramolecular bond formation reactions useful for examining covalent architecture such as disulfide bonds.[11] However, once again in the majority of cases, radical chemistry is utilized to modulate dissociation when applied to biomolecular study rather than to initiate bond formation.[12-]

   Herein we demonstrate that radicals can be used to efficiently create covalent bonds. New S-S bonds that recapitulate native disulfide links can be generated in highly efficient gas phase reactions in both intramolecular and intermolecular systems. New C-C bonds can be created from diradical, noncovalently bound species, allowing for the location of noncovalent adducts to be investigated. The details and potential applications of these reactions are discussed.

4.2 Experimental Methods

4.2.1 Materials

   Peptides RGDC, VTCG, PHCKRM, YGLSKGCFGLKLDRIGSMSGLGC, and SPKTMRDSGCFGRRLDRIGSLSGLGCNVLRRY were purchased from American Peptide Company (Sunnyvale, CA). SLRRSSCFGGR and CQDSETRTFY were purchased from Abbiotec (San Diego, CA) and CDPGYIGSR was purchased from Apexbio. Ammonium bicarbonate was purchased from Avantor (Center Valley, PA). Dithiothreitol was purchased from Arcos Organics (Geel, Belgium). Dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich (St. Louis, MO). Trifluoroacetic acid (TFA) and N,N'-Dicyclohexylcarbodiimide was purchased from Alfa Aesar (Haverhill,

MA). Acetonitrile (ACN), methanol and 1,4 dioxane were purchased from Fisher Scientific (Waltham, MA). 2-hydroxymethyl-18-crown-6 ether was purchased from TCI America (Portland, OR) and 3,5-diiodobenzoic acid was purchased commercially. Water was purified by Millipore Direct-Q (Millipore, Billerica, MA). A MacroTrap holder and MacroTrap consisting of polymeric reversed-phase packing material were purchased from Michrom Bioresources, Inc. (Auburn, CA).

4.2.2 Peptide and Protein Derivatization with Propyl mercaptan

Disulfide containing peptides were reduced as follows: 10 µL of a 1 mM peptide stock was diluted in 25 µL of 50 mM ammonium bicarbonate. To this solution, 1.5 µL of dithiothreitol (DTT) was added and the resulting solution was incubated in the dark at 37 ℃ for 1 hour. To remove salt from the mixture, the peptide was purified by MacroTrap and eluted in 0.1%TFA in 2%:98% H2O:ACN and lyophilized to a powder.

The lyophilized powder was dissolved in 10 µL of water. Modification of the free thiols was carried out by adding 10 µL of dimethyl sulfoxide and 1.5 µL of propyl mercaptan to the peptide solution. This mixture was placed in a 37 ℃ water bath overnight. The peptide was subsequently cleaned by Macrotrap and eluted in 0.1%TFA in 2%:98% H2O:ACN, lyophilized to a powder and dissolved in 50:50 H_2O:MeOH in 0.1% formic acid for a final concentration of 10 µM for MS analysis.

For the intermolecular disulfide formation experiments, peptides containing a single cysteine were modified with propyl mercaptan as described above. In order to encourage peptide interaction, peptides were mixed together to obtain final concentrations of roughly 100-150 µM for MS analysis.

4.2.3 Synthesis of 2-(hydroxymethyl-3,5-diiodobenzoate)-18-crown-6 ether (BC)

The synthesis of 2-(hydroxymethyl-3,5-diiodobenzoate)-18-crown-6 ether was carried out using a protocol previously described.[8] Briefly, 0.5 mmol DCC was dissolved in 5 mL of dioxane and added to a round bottom flask containing 0.50 mmol of 3,5-diiodobenzoic acid and 0.50 mmol of 2-hydroxymethyl-18-crown-6 ether. The mixture was left to react for 12 hours and gravity filtration was used to remove the crystalline hair-like precipitate.

4.2.4 Noncovalent Adduct of Synthetic Crown (BC)

Peptides and proteins were dissolved in a mixture of 50:50 H_2O:MeOH in 0.1% formic acid to obtain a final concentration of 10 μM. The synthetic crown BC was added to each of the solutions in a ratio of 1:4 (BC:tertiary amine). The number of BC noncovalent adducts observed varied for each sample.

4.2.5 Photodissociation of Peptides and Proteins

An LTQ linear ion trap mass spectrometer (Fisher Scientific, Waltham, MA) with a standard ESI source was utilized mass analysis. A quartz window was installed on the back plate of the LTQ for transmission of laser pulses from a flash-lamp pumped Nd:YAG Minilite laser (Continuum, Santa Clara, CA). For all experiments except those with ubiquitin, a single pulse of fourth harmonic (266 nm) 4 mJ light was triggered during the MS^2 activation step. For the ubiquitin+BC experiments, four pulses were triggered during the MS^2 activation step. Peptides were sprayed at the concentrations stated above at 3 μL/min with electrospray voltage of 3.25 kV and a capillary inlet temperature set to 215 ºC.

4.3 Results and Discussion

4.3.1 Intramolecular Disulfide Formation

Disulfide bonds are commonly found within peptides and proteins.[18] In proteins, such bonds help define tertiary and quaternary structure. In peptides, disulfide bonds are common in venoms and toxins where the crosslink not only serves to define structure and promote toxicity, but also helps prevent deactivation by proteolysis.[19,20] Peptides in aqueous solution that contain cysteine in the free thiol state can spontaneously form disulfides in the presence of oxygen, particularly if the pH is above 8.[21] Similar chemistry does not occur in the gas phase because solvent is not available to facilitate deprotonation of the thiol. Previous work has demonstrated that disulfide bonds can be selectively cleaved by UV photons, yielding homolytic cleavage of the S-S bond and two sulfur radicals.[9,22] In addition to native disulfide bond cleavage, cysteine residues modified with propyl mercaptan (PM) are also susceptible to homolytic fragmentation by either direct photon absorption or via energy transfer.[23,24] This type of fragmentation is demonstrated in Figure 4.1a for the 3+ charge state of porcine brain natriuretic peptide (BNP) 1-32 that had been reduced and then modified with PM at both cysteine residues (native sequence = SPKTMRDSGC FGRRLDRIGS LSGLGCNVLR RY, where Cys9 and Cys26 form a disulfide bond). The loss of one and two PMs yields the most abundant products in Figure 4.1a. The double PM loss creates a biradical species that could, in theory, recombine to form a disulfide by the pathway shown in Scheme 4.1a. Additional pathways are available following a single PM loss, which can yield a sulfur radical at either cysteine residue specifically or at a combination of both sites. Subsequent collisional activation of this product cannot distinguish between these options because loss of the second PM radical represents the primary product, as shown in Figure 4.1b. This suggests that regardless of

68

which sulfur radical is initially created, selective attack to displace the second PM and reform the native disulfide bond is the most favored outcome. This mechanism is shown in Scheme 4.1b.
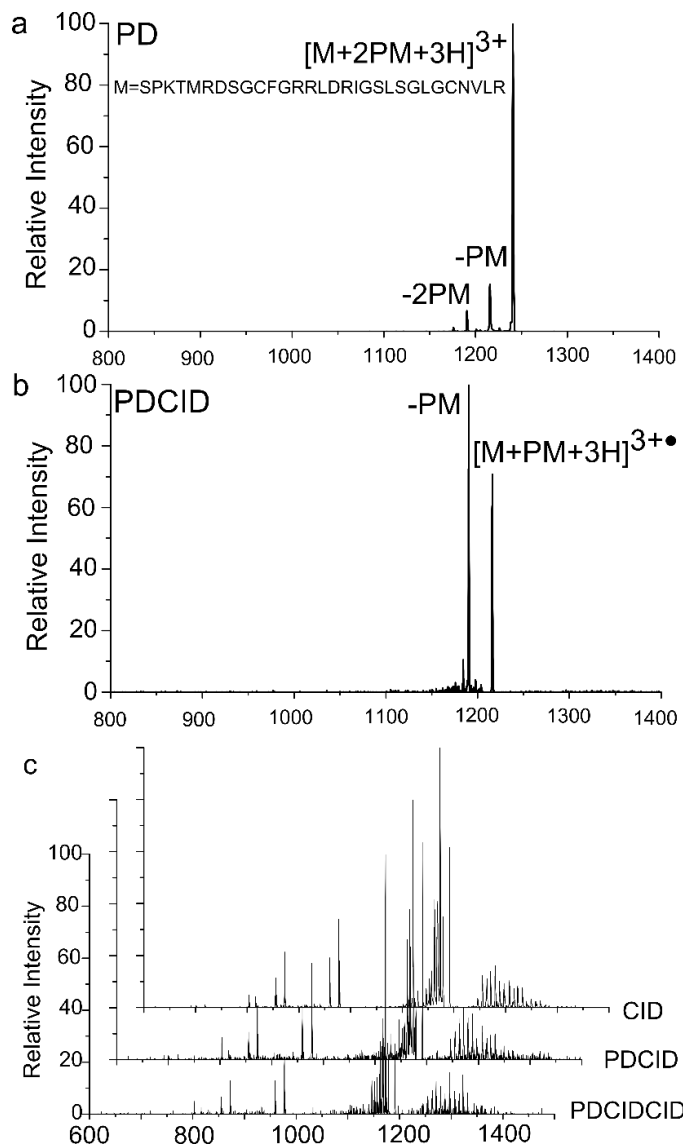
Figure 4.1: a) Photodissociation of the +3 charge state of BNP yields loss of one and two PM modifications. b) CID of the single PM loss leads to loss of the second PM. c) Stacked mass spectra comparing canonical peptide (CID) to disulfide generated by double loss of PM (PDCID) observed in (a) and sequential loss of two PM observed in (b). All three spectra are very similar, suggesting the same disulfide is activated in each case.

a) Biradical Recombination

PD 266 nm

Radical Recombination
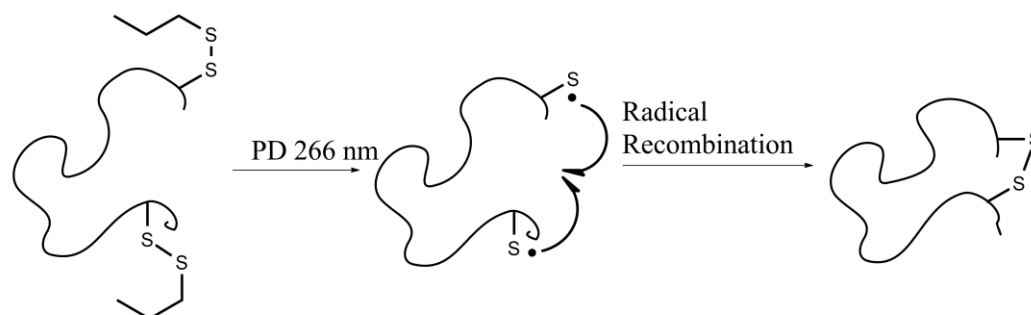
b) Sequential Radical Attack

PD 266 nm

CID

-PM
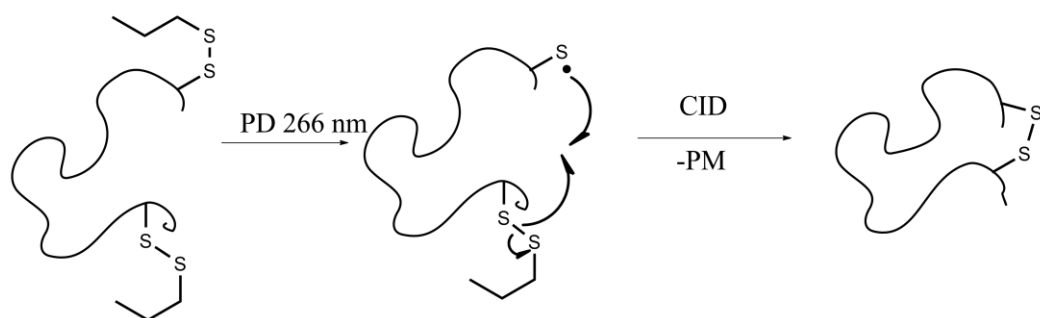
Scheme 4.1: a) Biradical species can recombine to directly form a disulfide bond. b) A single thiol radical can attack a nearby disulfide, leading exchange and loss of a different thiol radical.

To confirm disulfide bond formation, fragmentation of the product ions can be compared with that of the native disulfide electrosprayed directly from solution. Fragmentation data for both disulfides generated in the gas phase are compared with the native results in the stacked plot in Figure 4.1c. The bona fide disulfide bound peptide yields a complex spectrum with many product ions, consistent with its cyclic nature that requires fracture of multiple bonds to observe fragmentation within the ring (Figure 4.1c, top). The lower two spectra derived from gas phase syntheses are very similar. The probability for reproducing such complex spectra from different molecular structures is low, suggesting that disulfide bond formation occurred for both the simultaneous and sequential double PM losses. The selectivity of the sequential attack (Scheme 4.1b) can be rationalized in terms of the

relative reactivity of the sulfur radical initially generated by photodissociation. Sulfur radicals are stable relative to most positions within a peptide, making radical migration to most locations energetically uphill and unfavorable.[25,26] However, attack of another disulfide bond creates an equally stable radical, which is neutral from a thermodynamic perspective. Radical attack of disulfides in related systems has been shown to be energetically favorable and occur with relatively low activation energies.[11,27,28] The results illustrate that the two sulfur moieties encounter each other prior to the addition of sufficient energy to enable uphill radical migration or another dissociation pathway.

Similar experiments were conducted with Type C natriuretic peptide (CNP) with sequence YGLSKGCFGL KLDRIGSMSG LGC that contains a native disulfide bond between Cys7 and Cys23. After reduction and modification with two equivalents of PM, photoactivation of the 4+ ion at 266 nm yields the spectrum shown in Figure 4.2a. Interestingly, the loss of two PM is the most abundant loss, contrasting the results from BNP in Figure 4.1a. The double loss again results in formation of a disulfide bond and a subsequent activation spectrum indistinguishable from the native peptide, see insets in Figure 4.2a. Collisional activation of the single PM loss, however, does not yield primarily loss of a second PM as shown in Figure 4.2b. Loss of the second PM is the most abundant product, but many additional fragments are noted, including backbone dissociation where the PM modification has been retained as highlighted in the inset of Figure 4.2b. These fragments reveal that the initial PM was lost from both cysteine residues, though loss was more favorable from Cys7. This may be due to the closer proximity of Cys7 to Tyr1, which could enhance disulfide bond cleavage by energy transfer. In any case, there appear to be two structure populations for CNP, one where the two cysteine residues can easily interact and form a disulfide bond, and a second structure where this is prohibited by

71

significant energy barriers. The additional charge (4+ versus 3+) relative to the BNP should lead to greater Coulomb repulsion and may structural flexibility. Subsequent activation of the –PM product peak from Figure 4.2b leads to a fragmentation spectrum similar to that of the native disulfide peptide (see Figure 4.2c), although the presence of additional peaks may indicate a more complicated cross-linked product was created as well.
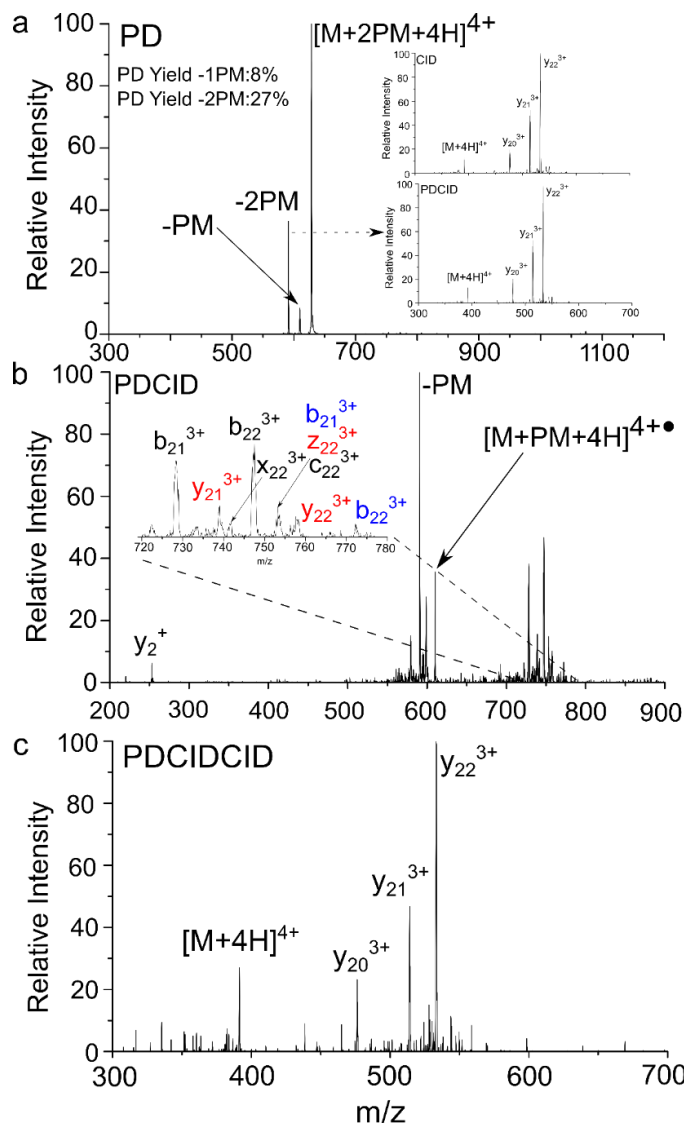
Figure 4.2: a) Photodissociation of the +4 charge state of YGLSKGC$_7$FGL KLDRIGSMSG LGC$_{23}$ yields primarily loss of both PM modifications. The insets compare collisional activation of solution (top) and gas phase (bottom) synthesized disulfides. b) CID activation of the single PM loss from the PD step. Blue labeled fragments indicate retention of PM at Cys23, black fragments indicate retention of PM at Cys7, and red fragments are ambiguous to the location of the PM modification. c) CID activation of the sequential PM loss observed in b.

4.3.2 Intermolecular Disulfide Formation

Disulfide bonds can also connect two different peptide strands, as occurs in the well-known case of insulin. To evaluate whether it is possible to synthesize such a bond in the

73

gas phase, we electrosprayed a solution containing two peptides, each with a single cysteine (RGDC and CQDSETRTFY). Under gentle electrospray, a noncovalent complex comprised of these two peptides can be observed and photoactivation yields the spectrum shown in Figure 4.3a. Interestingly, the weak noncovalent bonds holding the complex together are not broken, but losses of one and two PMs are noted with the double loss being significantly favored. Subsequent activation of the -2PM product yields a spectrum nearly identical to collisional activation of the disulfide bound pair directly electrosprayed into the mass spectrometer (Figure 4.3b and 4.3c). Given the low absorption of disulfide bonds, it is highly unlikely that the majority of the double –PM loss is generated by two separate photodissociation events. More likely, the loss of either PM individually enables attack at the remaining PM site and subsequent loss of the second PM to form the disulfide bond. Collisional activation of the single PM loss leads primarily to separation of the two peptides, suggesting the existence of a conformation that is unable to bring both cysteine residues into close proximity prior to reaching the noncovalent dissociation threshold.
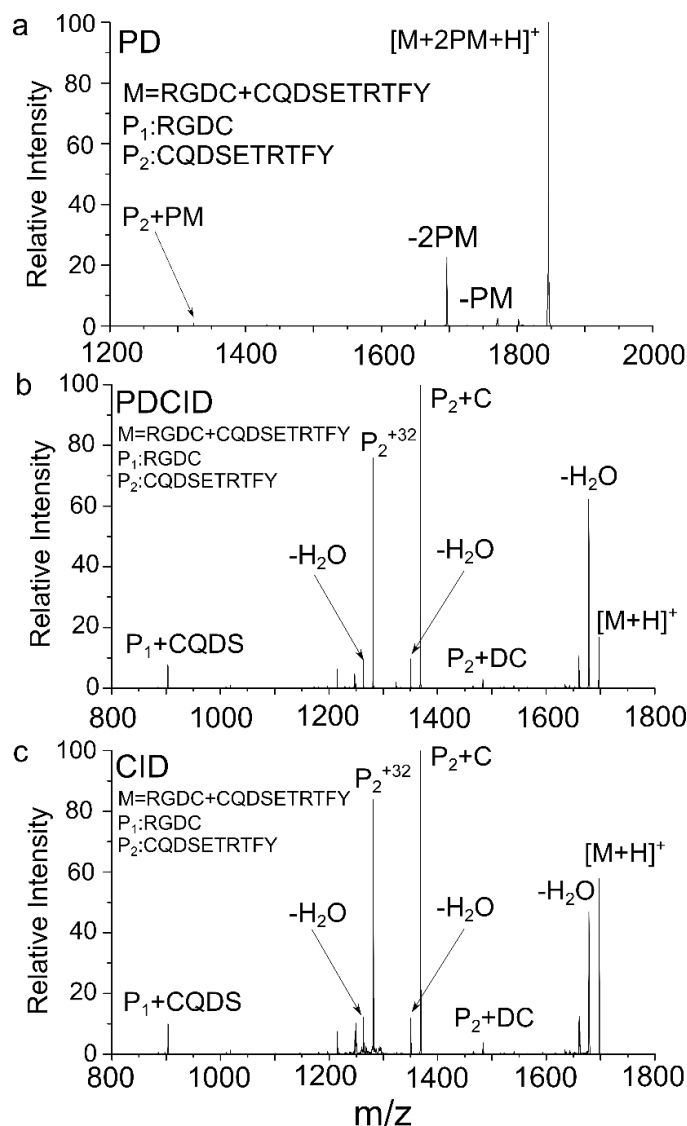
Figure 4.3: a) PD of the two noncovalent peptides RGDC and CQDSETRTFY each modified with PM. b) Subsequent CID activation of the loss of 2PM modifications from PD. c) CID fragmentation of the liquid phase formed disulfide linkage. P_1+CQDS refers to the peptide RGDC linked to the CQDS portion of P_2, while P_2^{+32} refers to CQDSETRTFY with an additional 32 Da mass corresponding to the sulfur from RGDC. –H2O refers to losses from the labeled peaks immediately to the right.

The results in Figure 4.3 clearly demonstrate it is possible to form disulfide bonds between noncovalently bound peptide pairs. However, to better evaluate the likelihood for such chemistry, a variety of peptide pairs were examined and the results are summarized in Table 1. The ratio of single to double PM loss varies by peptide pair, most likely due to

75

structural differences that influence the ability of the two PM groups to interact with each other. However, in every case the -2PM loss is observed, and the subsequent collisional activation spectra are close matches to CID spectra of the native disulfide. Although it is interesting that disulfide bonds can be synthesized in the gas phase, the implications of these results are also important in a larger context. The facility of so many peptides and peptide pairs to form disulfide bonds suggests that peptide structures are either 1) highly flexible in the gas phase, allowing cysteine residues to 'find' each other, or 2) composed of highly heterogeneous mixtures of structures where some fraction of conformations contain cysteine residues in close proximity, or 3) some combination of both possibilities. Activation of the single PM loss products offers some insight into the matter. The sequential PM losses listed in Table 1 are not uniformly abundant, implying that high flexibility is certainly not a uniform property of noncovalent peptide dimers. Inherent within the ability to explore structural space is the strength of the noncovalent interactions binding the two peptides together, which likely varies significantly depending on the nature of the peptides. The data are consistent with a heterogeneous population of many structures, some of which contain cysteine residues in close enough proximity to form a disulfide bond following minimal activation.

Table 1: Peptide pairs and modifiers examined for disulfide formation

Peptide Pair	Double PM Loss (%)	Single PM Loss (%)	Sequential PM Loss (%)
RGDC+CQDSETRTFY (2PM)	18	2	7
CQDSETRTFY+SLRRSSCFGGR (2PM)	19	4	16
RGDC+CDPGYIGSR (2PM)	24	5	17
RGDC+SLRRSSCFGGR (2PM)	6	10	32
RGDC+PHCKRM (2PM)	9	9	17

4.3.3 Diradical Crosslinking

Previous work has demonstrated that covalent bonds can be formed in the gas phase by recombination of biradical species (where the two radicals are independent of each other).[7] Another possibility can be envisioned starting with a diradical, where the two radicals are equivalent and contained on the same functional group. 2-(hydroxymethyl-3,5-diiodobenzoate)-18-crown-6 ether (BC) contains two C-I bonds which are photocleavable at 266 nm, affording a facile route to generate a diradical. Figure 4.4a shows the photoactivation of the peptide RGYALG noncovalently linked to BC. Dissociation of a single C-I bond is the major fragmentation pathway, yielding a single radical. Subsequent CID activation of the radical complex is shown in Figure 4.4b. As expected, collisional activation leads to migration of the radical to the peptide, loss of the noncovalent complex, and radical-directed dissociation of the peptide yielding backbone and side chain fragmentation.[29]
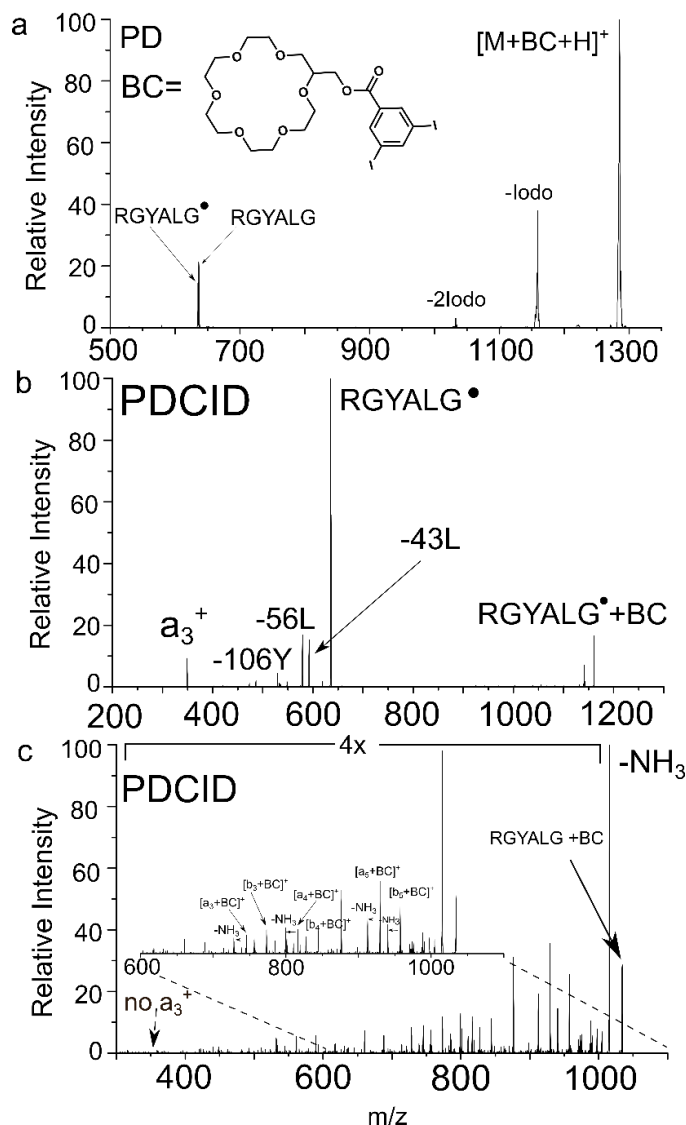
Figure 4.4: a) PD of the peptide RGYALG complexed with BC. Subsequent CID on either the b) single iodine loss or c) double iodine loss is shown. Peptide fragments indicate BC is attached near the N-terminus of the peptide.

The second most abundant products in Figure 4.4a are protonated RGYALG in both radical and canonical form, but there is also a minor product corresponding to the loss of two iodine atoms. This product would correspond to the diradical shown in Scheme 4.2. Aromatic radicals are highly reactive and capable of abstracting hydrogen from most sites with a peptide.[30] Abstraction of hydrogen, as shown in Scheme 4.2, would result in net migration of the radical from the BC to the peptide, a process already confirmed to occur

in both Figure 4.4a and 4.4b. If the remaining radical then reorients via rotation, it could recombine with the radical site on the peptide and form a new covalent bond. Collisional activation of the two iodine loss peak yields the spectrum shown in Figure 4.4c, which is dominated by various backbone fragments. Importantly, the loss of BC is not observed, consistent with covalent bond formation. Furthermore, the observation of a_3 ions with BC attached, and absence of a_3 ions without BC suggests attachment near the N-terminus. Given that the BC must be attached to either the side chain of arginine or the N-terminal amine,[31-33] this result suggests that localized cross-linking was favored.
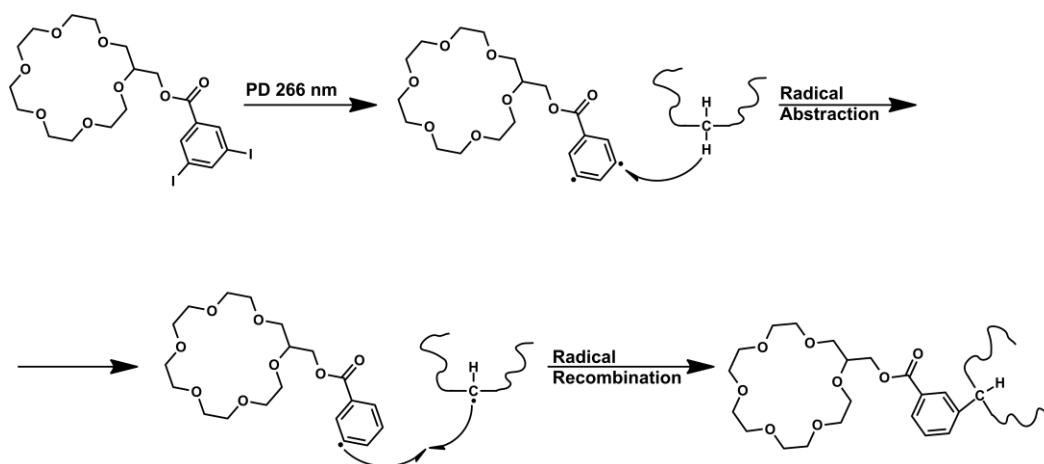
Scheme 4.2: Radical recombination of BC with peptide following photoactivation at 266 nm.

To explore the potential for using BC to identify crown ether binding sites, experiments were conducted with melittin (GIGAVLKVLT TGLPALISWI KRKRQQ-NH$_2$), which contains lysine residues near the N- and C-termini. Figure 4.5a and 4.5b show results from collisional activation of the double iodine losses for the 3+ and 4+ charge states of melittin+BC noncovalent complexes. The most probable sites of crown ether attachment are highlighted in red on the peptide sequence in Figure 4.5b. The most abundant fragment for both the 3+ and 4+ charge states is the loss of the BC without recombination,

but a significant number of backbone fragments with BC attached are also observed. Analysis of the fragment ions from the 3+ charge state indicates the site of modification is near the C-terminus, i.e. y_{13}^{2+} y_{17}^{2+} y_{18}^{2+} y_{19}^{2+} and y_{20}^{2+} ions all have BC attached while the a_9^{2+} does not. A similar conclusion is drawn from the series of b and y-ions identified for the +4-charge state (although the exact fragment ions differ). These results suggest that the binding site rich KRKR sequence near the C-terminus of melittin is the preferred binding site.
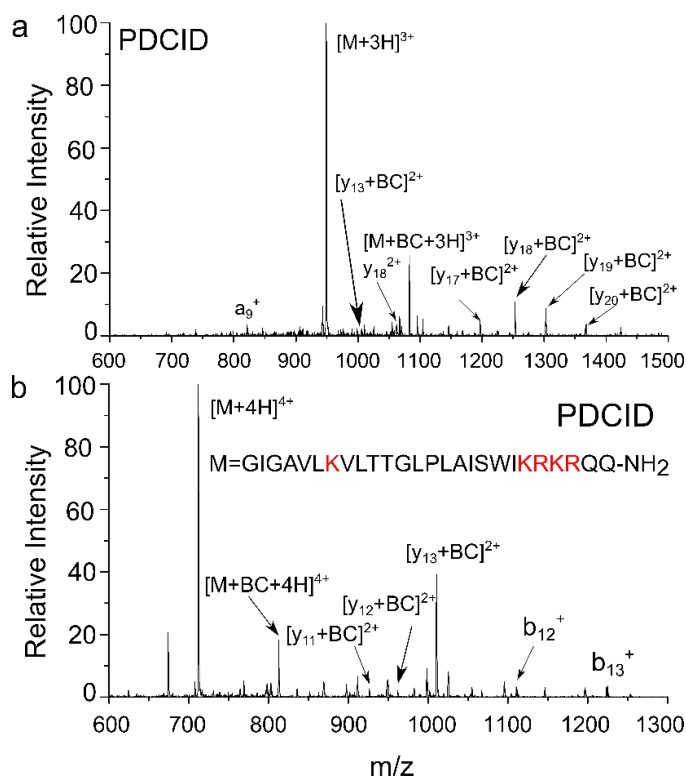
Figure 4.5: PDCID of the a) +3 and b) +4 charge states of the small protein melittin (GIGAVLKVLTTGLPALISWIKRKRQQ-NH2). Possible sites for the noncovalent attachment of BC are labeled in red.

The small protein ubiquitin has been studied extensively by mass spectrometry.[34-40] Investigation of the preferred 18-crown-6 binding sites has been evaluated by site directed mutagenesis where each lysine residue was sequentially removed in a difficult and time-

consuming process.[41] For comparison, the preferred binding site was examined herein by electrospraying a mixture of ubiquitin and BC, followed by photoactivation of a single-adduct noncovalent complex, see Figure 4.6a. Collisional activation of double iodine-loss peak yields the spectrum shown in Figure 4.6b. The most abundant product is loss of the BC adduct, but many b and y ions are also detected. A series of y-ions out to y_{60}^{5+} is observed with no BC attached, consistent with BC being attached near the N-terminus. A single b-ion with BC attached is also detected, although unfortunately, the ion corresponds to b_{58}^{5+} and does not significantly constrain the location of the BC adduct. However, when taken as a whole, the fragments in Figure 4.6b suggest that the BC attaches near the N-terminus, in agreement with previous results.[41] These results were obtained in relatively simple experiments and reveal that diradical attack to create covalent crosslinks is a promising way to extract structural information about noncovalent assemblies. To verify the location of noncovalent binding, additional experiments could be conducted where potential sites identified rapidly by BC binding would be mutated to non-interacting residues.
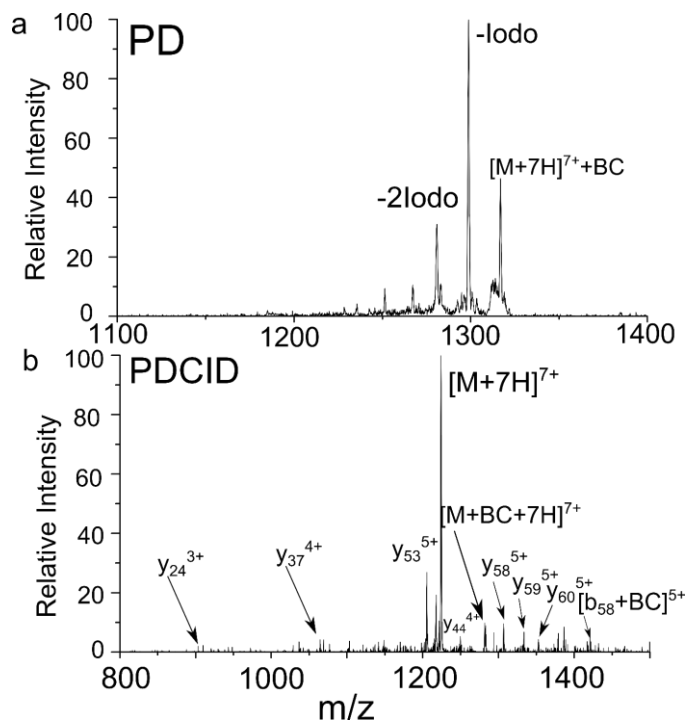
Figure 4.6: a) PD of the +7 charge state of ubiquitin and BC complex. b) CID on the loss of 2 iodine produces fragment ions which localize modification to the N-terminal side of ubiquitin

4.4 Conclusion

The present results indicate that radicals can be utilized for the formation of new S-S and C-C bonds in the gas phase. Interestingly, in one case unreactive sulfur radicals are used, while in the other highly reactive aromatic phenyl radicals form new bonds. In both situations, the relative reactivity is key for successful creation of a new bond. Disulfide bonds can be formed through the photoactivation of peptide species containing two PM modified cysteines. Simultaneous or sequential loss of both PM groups leads to the formation of a new disulfide bond. Formation of these new disulfide bonds occurs both intermolecularly and intramolecularly. These observations suggest that thiol radicals are quite selective, preferring to attack other thiol radicals or disulfide bonds while leaving most sites in peptides untouched. In UVPD and ECD experiments where sulfur radicals could be formed, caution should be taken when examining proteins containing multiple

disulfide bonds where gas phase disulfide scrambling could occur even if only a single disulfide bond were homolytically cleaved.

Diradical initiated bond formation is also shown to be useful for identifying sites of interaction for crown ether based noncovalent complexes. The sites of adduction do not appear to be random, but occur at preferred locations. Crown ethers have also been shown to replace solvent in peptides and proteins sprayed into a mass spectrometer,[42-44] helping maintain their liquid phase structure. Diradical covalent coupling may provide useful insights into the sites of solvation that help retain tertiary structure as proteins are transferred from liquid to the gas phase.

4.5 Acknowledgements

4.6 References

[1] C.J. Shaffer, P.C. Andrikopoulos, J. Řezáč, L. Rulíšek, F. Tureček, J. Am. Soc. Mass Spectrom., 27 (2016) 633–645.

[2] R.R. Julian, J.A. May, B.M. Stoltz, J.L. Beauchamp, Angew. Chemie Int. Ed., 42 (2003) 1012–1015.

[3] H. Han, S.A. McLuckey, J. Am. Chem. Soc., 131 (2009) 12884–12885.

[4] J.R. Stutzman, K.M. Hassell, S.A. McLuckey, Int. J. Mass Spectrom., 312 (2012) 195–200.

[5] W.M. McGee, M. Mentinova, S.A. McLuckey, J. Am. Chem. Soc., 134 (2012) 11412–11414.

[6] S. Lee, R.R. Julian, S.J. Valentine, J.P. Reilly, D.E. Clemmer, Biomolecular condensation via ultraviolet excitation in vacuo, Int. J. Mass Spectrom., 316–318 (2012) 6–11.

[7] X. Zhang, R.R. Julian, Phys. Chem. Chem. Phys., 14 (2012) 16243–16249.

[8] T. Ly, X. Zhang, Q. Sun, B. Moore, Y. Tao, R.R. Julian, Chem. Commun., 47 (2011) 2835–2837.

[9] L.E. Talbert, R.R. Julian, J. Am. Soc. Mass Spectrom., (2018).

[10] R. R. Julian, J. Am. Soc. Mass Spectrom., (2017).

[11] C.H. Sohn, J. Gao, D.A. Thomas, T.-Y. Kim, W.A. Goddard Iii, J.L. Beauchamp, Chem. Sci., 6 (2015) 4550–4560.

[12] B.N. Moore, S.J. Blanksby, R.R. Julian, Chem. Commun., (2009) 5015–5017.

[13] T. Ly, R.R. Julian, J. Am. Chem. Soc., 132 (2010) 8602–8609.

[14] J.K. Diedrich, R.R. Julian, Anal. Chem., 83 (2011) 6818–6826.

[15] I.K. Chu, J. Laskin, Eur. J. Mass Spectrom. 17 (2011) 543–556.

[16] C.K. Barlow, W.D. McFadyen, R.A.J. O'Hair, J. Am. Chem. Soc. 127 (2005) 6109–6115.

[17] M. Lee, M. Kang, B. Moon, H. Bin Oh, Analyst, 134 (2009) 1706–1712.

[18] M. V Trivedi, J.S. Laurence, T.J. Siahaan, Curr. Protein Pept. Sci. 10 (2009) 614–625.

[19] S.Y. Huang, S.F. Chen, C.H. Chen, H.W. Huang, W.G. Wu, W.C. Sung, Global Anal. Chem. 86 (2014) 8742–8750.

[20] P. Anand, A. Grigoryan, M.H. Bhuiyan, B. Ueberheide, V. Russell, J. Quinoñez, P. Moy, B.T. Chait, S.F. Poget, M. Holford, PLoS One. 9 (2014) e94122.

[21] G. Rajpal, P. Arvan, Elsevier, 2013; pp 1721–1729.

[22] A. Agarwal, J.K. Diedrich, R.R. Julian, Anal. Chem., 83 (2011) 6455–6458.

[23] N.G. Hendricks, N.M. Lareau, S.M. Stow, J.A. McLean, R.R. Julian, J. Am. Chem. Soc., 136 (2014) 13363–13370.

[24] N.G. Hendricks, R.R. Julian, Analyst, 141 (2016) 4534–4540.

[25] B.N. Moore, R.R. Julian, Phys. Chem. Chem. Phys., 14 (2012) 3148–3154.

[26] M. Lesslie, S. Osburn, M.J. van Stipdonk, V. Ryzhov, Eur. J. Mass Spectrom. 21 (2015) 589–597.

[27] F. Tureček, M. Polášek, A.J. Frank, M. Sadílek, J. Am. Chem. Soc. 122 (2000) 2361–2370.

[28] E.H. Krenske, W.A. Pryor, K.N. Houk, J. Org. Chem. 74 (2009) 5356–5360.

[29] Q. Sun, H. Nelson, T. Ly, B.M. Stoltz, R.R. Julian, J. Proteome Res. 8 (2009) 958–966.

[30] Blanksby, S. J.; Ellison, G. B. *Acc. Chem. Res.* **2003**, *36* (4), 255–263.

[31] R.R. Julian, J.L. Beauchamp, Int. J. Mass Spectrom. 210–211 (2001) 613–623.

[32] R.R. Julian, M. Akin, J.A. May, B.M. Stoltz, J.L. Beauchamp, Int. J. Mass Spectrom. 220 (2002) 87–96.

[33] R.R. Julian, J.L. Beauchamp, J. Am. Soc. Mass Spectrom. 13 (2002) 493–498.

[34] K. Breuker, H. Oh, D.M. Horn, B.A. Cerda, F.W. McLafferty, J. Am. Chem. Soc. 124 (2002) 6407–6420.

[35] E.W. Robinson, E.R. Williams, J. Am. Soc. Mass Spectrom. 16 (2005) 1427–1437.

[36] J.R. Cannon, K. Martinez-Fonts, S.A. Robotham, A. Matouschek, J.S. Brodbelt, Anal. Chem. 87 (2015) 1812–1820.

[37] S. Warnke, C. Baldauf, M.T. Bowers, K. Pagel, G. von Helden, J. Am. Chem. Soc. 136 (2014).

[38] G. Wang, R.R. Abzalimov, C.E. Bobst, I.A. Kaltashov, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 20087–20092.

[39] H. Shi, D.E. Clemmer, J. Phys. Chem. B. 118 (2014) 3498–3506.

[40] G.E. Reid, J. Wu, P.A. Chrisman, J.M. Wells, S.A. McLuckey, Anal. Chem. 73 (2001) 3274–3281.

[41] Z. Liu, S. Cheng, D.R. Gallie, R.R. Julian, Anal. Chem., 80 (2008) 3846–3852.

[42] J.G. Bonner, N.G. Hendricks, R.R. Julian, J. Am. Soc. Mass Spectrom. 27 (2016) 1661–1669.

[43] Y. Tao, R.R. Julian, Int. J. Mass Spectrom., 409 (2016) 81–86.

[44] S. Warnke, G. von Helden, K. Pagel, J. Am. Chem. Soc., 135 (2013) 1177–1180.

Chapter 5: Revisiting the Use of Concept Maps in a Large Enrollment General Chemistry Course: Implementation and Assessment

5.1 Introduction and theoretical framework of concept mapping

General Chemistry is often the first science class taken by students in their undergraduate science curriculum. Unfortunately, historically high failure rates have given this class "gatekeeper" status for students who wish to major in STEM fields. One of the factors contributing to the lack of student success is the fact students often enter their first undergraduate chemistry course possessing misconceived mental models that thereby become a barrier to learning the foundational concepts covered in the general chemistry curriculum.[1-3] Even students who have performed well on typical classroom assessments in their general chemistry courses have been found to struggle when asked to provide conceptual explanations for questions related to core learning objectives in the general chemistry curriculum.[4]

To overcome the limitations typical classroom assessments possess in regards to helping students develop clear and cogent mental models and conceptual understanding, considerable effort has been given to develop and employ metacognitive interventions.[5,6] Metacognition is the process utilized to evaluate and monitor one's understanding and performance of the material. There are currently several strategies which are used to enhance metacognition, including paraphrasing and rewriting, working on homework problems, previewing material, and pretending to teach information.[6] While the use of these strategies has been reported in the educational research literature,[7] there remains a need to develop more widely applicable implementation strategies that can both help develop and assess student understanding of conceptual ideas.[8910]

Anecdotal evidence has broadly confirmed that students in our general chemistry program also continue to struggle to provide scientifically acceptable conceptual explanations for many fundamental learning objectives. We therefore decided to revisit the use of concept maps as both an intervention to help students better develop conceptual models and provide a means to measure this type of learning outcome. One of the first reports on the use of concept maps and their impact on student learning was published by Novak,[11] and concept maps have been subsequently proposed numerous times as a strategy to increase student retention and learning.[12-14] Generally speaking, a concept map requires students to reflect on their learning and the specific learning objectives by having them identify key concepts that have been covered in the course, and using words or short phrases to link multiple concepts where appropriate. It is important for students to draw links between the concepts to establish the relationship between them and solidify their deeper understanding of the foundational course concepts. The process of making concept maps engages students in a form of active learning, and the continual cycle of reflection required to update the conceptual links makes concept mapping a potentially effective form of metacognitive engagement.[15]

The theoretical framework underpinning the impact of concept mapping on student conceptual understanding can be traced to Ausubel's assimilation theory of learning.[16] Though this is reviewed in detail by Novak,[11] we highlight here that Ausubel proposes that meaningful learning takes place when the learner relates new knowledge to concepts he/she already knows, and when the learner can identify the key concepts in the new knowledge and how to relate these to other concepts in other contexts. Ausubel points out if this connection of concepts from new knowledge to other concepts does not occur, verbatim/non-substantive learning can still occur, but this type of learning has less value

87

and can actually interfere with subsequent learning. It is quite clear concept mapping can play an important role in promoting the type of meaningful learning described by Nesbit and Adesope and Turan-Oluk and Ekmekci,[17,18] and should certainly be more strongly considered by STEM instructors to be part of their instructional arsenal.

5.1.1 Barriers to implementing concept map assignments

As alluded to above, not only does a concept map assignment provide an opportunity for students to develop more completely developed mental models, it can be a valuable tool for assessing the students' conceptual understanding.[19] Assessment of the effectiveness of student conceptual understanding might be a daunting task for many instructors, but fortunately a variety of rubrics have been developed for carrying out the evaluation of student conceptual thinking. Typical rubrics focus on the scoring of hierarchies presented by the students and/or assess at the validity of the proposed links in traditional concept map assignments,[13,19] whereas more recent studies have proposed creating new types of concept linking activities that can be more easily evaluated.[20]

Though concept maps are clearly a valuable intervention in developing and assessing student conceptual understanding, their use in first-year college level chemistry courses has not been broadly demonstrated. Previous studies on the use of concept maps in general chemistry courses are generally limited to small enrollment implementations (n < 100),[21-24] and are often used only to map concepts within specific learning units of the course. For studies that have been carried out in large enrollment courses, the analysis of student data has generally been limited to only a subset of the class population.[25] Perhaps the most noteworthy implementation of a concept map intervention in a large enrollment general chemistry course comes from Francisco and co-workers.[13] In this study, the implementation focused not only on the student aspect of completing the concept map

assignments, but also incorporated graduate student teaching assistants (TAs) into the grading of concept maps and delivery of feedback to the students. It was shown concept maps in large enrollment classes do positively impact the development of student conceptual knowledge, in particular demonstrating the concept map intervention appeared to improve student performance on complex multi-step algorithmic problems. However, Francisco and coworkers also report significant resistance from the students to include the concept maps as graded activities and the implementation required significant training of the TAs in order to achieve an observable positive impact.

5.1.2 Motivation for the study and experimental hypotheses

Hence, the goal was to create a concept map implementation for a large enrollment general chemistry course that achieved the following objectives: 1) balance the need to create a significant incentive for students to complete the concept maps with the desire to avoid student resistance to being graded on a subjective measure; 2) provide a template for a concept map assignment that can be more easily adopted by other instructors of large enrollment courses; and 3) allow for the design of a quasi-experiment that adds to the limited pool of data describing the impact of concept map assignments on student learning outcomes in large enrollment general chemistry courses. The experimental hypotheses that drove the research design were: 1) students who completed a quarter-long concept map intervention would achieve more significant learning gains compared to a control group of students who did not use concept mapping; and 2) higher proficiency in creating valid and well-developed concept maps would correlate to gains in conceptual understanding. Herein, we will describe how a concept map assignment was administered in a streamlined fashion using TA-led discussion group recitation sessions and how the concept map intervention was improved in a second phase of implementation. We will

also describe results from a quasi-experiment in which student performance on a concept inventory and student responses from the Student Assessment of Learning Gains (SALG)[a] survey instrument were compared between a concept map treatment group, and a control group in which students completed weekly journal entries.

5.2 Quasi-experimental design and research methods - F17

5.2.1 Course description - F17

The first concept map implementation was carried out in a first quarter general chemistry course (CHEM 001A). This is the first course in a three-quarter general chemistry sequence that is required for both chemistry majors and all other science majors on the University of California-Riverside (UCR) campus. The topics covered in this course include: the scientific method and measurement; atomic structure; compounds and the mole concept; chemical reactions and stoichiometry; thermochemistry and energy changes in reactions; electronic structure of the atom; and chemical bonding and molecular structure. This course was taught for an "on sequence" cohort of first-year students in the fall of 2017 (F17), but also included second year students and upper-class students who may have needed the course as a general college science requirement or a pre-requisite for health professional schools.

5.2.2 Course structure and experimental study groups - F17

The F17 course consisted of a large enrollment lecture, which met two times per week for 80 minutes, and associated recitation sections that met once per week for 50 minutes (30-40 students each). The quasi-experimental design assigned recitation sections taught by a TA who had prepared in advance to implement the concept map intervention as the treatment group. The remaining recitation sections, taught by a different TA, were assigned as the control group and required the students to complete a weekly journal entry

in lieu of a concept map. This design also allowed the instructors to assign similar workloads and use the same grading scheme both groups. To minimize potential "treatment" effects in the journal entry control group, these students wrote weekly journals about what they had learned in each week's lectures and no guidance regarding how to structure the journal entries was provided. Students were simply instructed to summarize what they learned in lecture each week. Conversely, the concept map treatment group was instructed to build on the map from the previous week, and to emphasize important concepts from each chapter and show the relationship between various chapters.

Students were able to choose the discussion group sections in which they enrolled. Students generally choose sections which best fit within their class schedule, but the time of day in which the sections are scheduled certainly impacts the student choices (i.e., sections scheduled early in the morning are generally less desirable). Though students self-selected into the discussion group sections, they did not know which TAs would be teaching the sections at the time of enrollment or that the different sections might do different activities throughout the term.

5.2.3 Concept map and journal entry implementation - F17

Each concept map or journal was awarded a single point for nine of the ten weeks, and four points for the final week. Though all assignments were graded simply for completion and these points were awarded as extra credit toward the overall lecture grade, students were informed that no points would be awarded if they simply submitted their previous concept map. During the course of the term the TA found no instances in which students handed in unchanged concept maps from previous weeks. This grading design was chosen in an effort to promote student compliance in completing the weekly concept map

91

assignment without creating the type of resistance one might expect to arise when subjective assignments such as this are graded more rigorously.

To increase the convenience of collecting and evaluating the concept maps, students were required to use a concept mapping program. Cmap is a freely available software program that students were able to download to their computers for the development of the concept maps,[b] and students were able to save and submit their maps as PDF files.[26,27] The TA was familiarized in the use of the Cmap program and was able to answer any questions students had with the program. To help ensure students were able to navigate the Cmap program and properly save their concept maps, the TA created a video tutorial that the students could view anytime during the term on the online course management system.

The first implementation of the concept map intervention was carried out in the F17 CHEM 001A class. In addition to encouraging students to view the online Cmap tutorial, the TA also facilitated a 1-hour discussion demonstrating how to build a concept map using the Cmap program using key concepts from the first week of class. Following a completely open level of inquiry for the concept maps, students were then instructed to build their own concept maps using that material that they learned each week.[28] No outline or guiding key terms were provided to the students. This enabled students to have full control over how they would build the concept map and in what way they would connect key concepts to one another. The TA briefly examined the concept maps each week but rarely provided feedback to the students unless a question about a particular concept was arose during the recitation section.

5.2.4 Grading rubric for concept maps and journals - F17

Concept maps and journals were scored using an adapted version of the concept map rubric developed by Besterfield-Sacre and co-workers.[23] Concept maps and journals were scored on comprehensiveness of the covered material, organization and linking between chapters, the correctness of the material in each chapter, and the correctness of the links between concepts (see Table 5.1). Students in the journal control group were not instructed as to what should be added other than discussion of the weekly material, and the journal entries were generally organized in chronological order in regards to the topics covered throughout the quarter. Therefore, the organization/links category focused on whether students took it upon themselves to discuss how a previous chapter impacted material in subsequent chapters. To ensure there was consistency in how the scoring rubric was used to allocate points, two graders discussed how points were to be awarded and each independently scored a random sample of 10 concept maps and 10 journal entries for comprehensiveness and organization/links (the graders consisted of a graduate student teaching assistant and an undergraduate research assistant). The course instructor then met with the two graders to discuss how the rubrics were applied in the grading and clarified how the student responses should be evaluated. The two graders then independently evaluated another random sample of 10 concept maps and journal entries and it was found the two graders agreed on the scoring on greater than 80% of the items. The remaining concept maps and journal entries were scored for the comprehensiveness and organization/links categories by the two graders. To ensure consistency and accuracy for scoring in the correctness category two experienced chemistry instructors evaluated a random sample of concept maps and journals, and scores were compared in the same manner as described above to calibrate the application

of the rubric. An additional random sample of concept maps and journal entries were evaluated by both instructors to ensure over 80% of the items were graded in an equivalent fashion, and the remaining concept maps and journal entries were scored for correctness.

Table 5.1: Scoring rubric for concept maps and journal entries. Each category is scored 1-3 with the total score for each concept map being scored out of 9.

Points	Comprehensiveness	Organization/Links [a]	Correctness
3	Contains 1 or more key concepts in each chapter.	Map is clear and easy to follow. Contains links between 4 or more discussed chapters.	Map contains 1 or less errors.
2	Contains 1 or more key concepts in 2-4 chapters.	Map is slightly cluttered. Contains links between 3 chapters.	Map contains 2-3 errors.
1	Contains one or more key concepts in 0-1 chapters.	Map is not easy to follow. Contains links between only 2 chapters.	Contains 4 or more errors.

a = Journal entries were evaluated through links between the chapters; links between 3 or more chapters received a score of 3; links between two chapters received a score of 2; no discussion of connection between the chapters earned a score of 1.

Figure 5.1 shows an example of a well-developed and underdeveloped concept maps for the F17 implementation. Figure 5.1a includes a color-coded is easy to follow, separating the content from each chapter into easily identifiable subsets. The concept map contains correct links between the different concepts and links together the chapters with previous material covered throughout the quarter. Figure 5.1b provides a map which is under-developed. The map includes a bare minimum of concepts selected and each chapter is isolated from one another leading to no information as to how the previous chapters may influence the current. The phrases connecting the concepts are also not explained fully leading to confusion as to what the student was trying to say about the concepts. However, in both maps, it is difficult to determine the starting location of how to read the map which may impact their effectiveness as a learning and study tool.
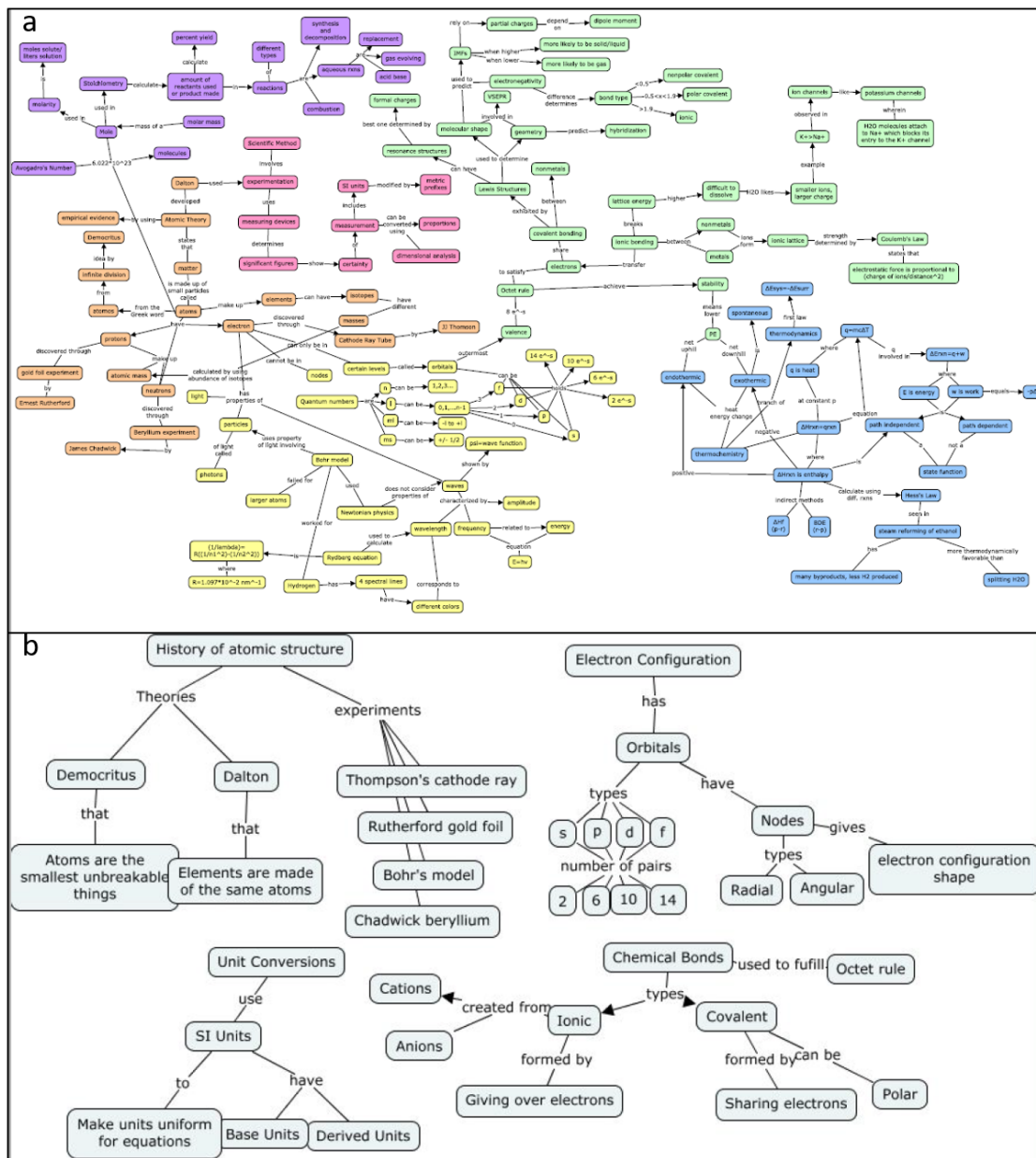
Figure 5.1: Example of a) a well-developed concept map with correct connections and connecting phrases (Comprehensiveness = 9, Organization/Links = 3, Correctness = 3, Total score = 9/9) and b) under-developed concept map with incorrect connections of missing links (Comprehensiveness = 1, Organization/Links = 1, Correctness = 1, Total score = 3/9)

5.2.5 Concept inventory - F17

In order to gauge the effectiveness of the treatment on the students' ability to engage in conceptual thinking, a concept inventory was utilized in addition to normal class exams. Because the traditional exams used in our general chemistry courses include a significant number of algorithmic problems that do not probe conceptual thinking, the questions included in the concept inventory placed an emphasis on conceptual understanding. A number of the items in the concept inventory also required applying multiple concepts in order to arrive at the correct answer. The F17 implementation consisted of a total of 10 questions (1 point for each item). These questions included 3 multiple choice questions and 8 free response questions. For the purposes of the concept inventory, these questions were graded as either incorrect (0 points), partially correct (0.5 points), or correct (1 point). The pre-test was administered in the second week of the recitation sections and the same questions were included in the final exam.

5.2.6 SALG survey and data collection - F17

To test the first quasi-experimental hypotheses a concept inventory and student self-assessment of leaning survey (SALG) was used in a pre-test/post-test alternative treatment/control group design as described by Barbara and co-workers.[29] Though to our knowledge the efficacy of a concept map intervention has not been previously evaluated using this type of quasi-experimental design in a chemistry course, we would like to highlight the report from Burdo and O'Dwyer that assessed the impact of a concept map treatment in an undergraduate physiology course.[30] In this prior study the concept map treatment group was compared to a negative control group and a second treatment group in which retrieval practice was chosen as the independent variable of interest. The experimental design described by Burdo and O'Dwyer acts as a good model for the quasi-

experimental reported herein, though it is noted this previous study assessed the impact of the concept map treatment on more traditional course exams as opposed to a measure of conceptual understanding.

The SALG survey was also administer in a pre/post format using a modified version of the freely available instrument. Questions that focused on conceptual learning gains, learning gains related to specific course topics, and learning gains related to applying course concepts in other contexts were chosen for the purpose of this study (see Tables 5.8-5.11). The pre-SALG survey was available to students during a seven-day period starting two days before the first class and the post-SALG was available to students during a seven-day period starting two days before the final exam, and both surveys were completed using the online SALG interface. Students were informed if they completed both the pre- and post-SALG they would receive a small amount of extra credit on their final course point total (5 points out of 1000 total course points).

5.2.7 Statistical analyses - F17

All statistical analyses were carried out using the SPSS Statistics 24 software package.[c] Analysis of covariance (ANCOVA) tests were carried out to compare the performance of the F17 treatment and control groups on the concept inventory post-test, while holding constant the concept inventory pre-test. Associated ANOVA analyses were used to confirm the pre-test baseline measures in the ANCOVA did not differ between treatment and control groups. In order to impart more rigorous statistical control of the student background traits, multiple regression analyses were carried out in which the concept inventory post-test scores and overall final exam scores were compared between treatment and control groups, while holding constant other independent variables related to student academic preparation and socio-economic status.[31] The ANCOVA and multiple

97

regression analyses were also carried out using only the top 33% of students from the concept map treatment and journal entry control groups, as determined by the scores on the concept map/journal rubric (this threshold was chosen as the cutoff to balance the desire to include higher performing students while maintaining an adequate sample size). These analyses were completed in order to determine if students with superior concept mapping skills performed better on the concept inventory and/or final exam compared to their counterparts in the journal control group (i.e., would students with good concept mapping skills achieve greater learning gains than students with good journal keeping skills).

Student responses to the previously validated SALG survey instrument were also compared between the treatment and control groups. In the F17 implementation, pre- and post-SALG data could not be matched between students, therefore chi-squared analyses were carried out to determine if the concept map treatment group had a significantly higher proportion of positive responses to the survey questions compared to the journal entry control group.

5.3 Results and Discussion - F17

An initial implementation of the treatment was performed during the F17 first-quarter general chemistry course. Table 5.2 summarizes the descriptive statistics for the control and treatment groups participating in the quasi-experiment. The academic preparation of students in the treatment and control groups appeared to be equivalent when considering the distribution of High School GPA, SAT math, and the pretest scores. These results would corroborate that there is no significant difference between the control and treatment groups for the F17 implementation.

Table 5.2: Descriptive Statistics for Fall 2017 treatment and control groups.

	Journal (n = 121)	Concept Map (n = 119)
High School GPA Avg.	3.76 +/- 0.28	3.75 +/- 0.29
SAT Math Avg.	612.29 +/- 87.29	617.22 +/- 80.42
Pre-Concept Inventory Avg. (out of 11)	1.42 +/-1.51	1.75 +/- 1.51
Post-Concept Inventory Avg. (out of 10)	5.75 +/- 2.45	5.90 +/- 2.36
Exam 1 Avg. (out of 100)	83.96 +/- 12.21	81.99 +/- 13.38
Exam 2 Avg. (out of 100)	70.91 +/- 17.90	73.20 +/- 20.39
Final Exam Total Avg. (out of 400)	267.35 +/- 57.32	270.00 +/- 58.07
Avg. rubric score on final concept/journal entry (out of 9)[a]	3.99 +/- 3.30 (n = 75)	4.04 +/- 2.85 (n = 86)
Asian	n = 53	n = 66
Black or African American	n = 5	n = 1
Hispanic or Latino	n = 45	n = 32
Multi-Race/Unknown/Non-Resident Alien	n = 11	n = 11
White	n = 7	n = 9
Male	n = 56	n = 61
Female	n = 62	n = 57
Gender Not Reported	n = 2	n = 0

[a]Not all students submitted a final concept map/journal. The n value designates the number of students that completed the final concept map/journal assignment.

The ANCOVA analysis in Table 5.3 indicates no significant difference between the concept map treatment and journal control groups in their performance on the concept inventory post-test or final exam, when holding constant the concept inventory pre-test. A multiple regression analysis was performed to hold constant the myriad of other possible factors that may have influenced student performance. The multiple regression analysis revealed no significant difference in concept inventory post-test or final exam scores between the concept map and journal groups, even while holding constant a variety of independent variables such as ethnicity, gender, high school GPA, and SAT scores

(Tables 5.4 and 5.6). Finally, the final exam and concept inventory post-test scores were compared between treatment and control groups, in which only the top 33% of students based on the concept map and journal rubric scores were included in the multiple regression analysis. These results also revealed no significant difference between the study groups (Tables 5.5 and 5.7).

Table 5.3: Control and Treatment ANCOVA Fall 2017 (Pre-concept Inventory = covariate; Final exam/Post-Concept Inventory = Dependent Variable).

Concept Map Treatment vs. Journal Control	F	p	partial η^2
Concept Inventory Pre-test Covariate/Concept Inventory Post-test Dependent Variable	0.301	0.584	0.001
Concept Inventory Pre-test Covariate/Final Exam Dependent Variable	0.496	0.482	0.002
Top 33% of Concept Map Students (n = 28) vs. Top 33% of Journal Students (n = 25)[b]	**F**	**p**	**partial η^2**
Concept Inventory Pre-test Covariate/Concept Inventory Post-test Dependent Variable	1.745	0.188	0.011
Concept Inventory Pre-test Covariate/Final Exam Dependent Variable	0.740	0.391	0.005

Table 5.4: Multiple Regression Fall 2017. (Full Class Dependent Variable: Final Exam). Group indicates coded treatment/control. Journal Group = 0; Concept Map = 1.

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	-54.126	40.172		-1.347	0.179
Group	-3.101	6.017	-0.026	-0.515	0.607
Asian	21.308	11.505	0.181	1.852	0.065
Black or African American	-7.888	20.903	-0.022	-0.377	0.706
Hispanic or Latino	22.704	12.025	0.184	1.888	0.060
White	15.173	15.904	0.062	0.954	0.341
Gender	0.001	0.004	0.010	0.186	0.853
High School GPA	18.121	7.897	0.116	2.295	**0.023**
SAT Math	0.351	0.040	0.498	8.672	**0.000**
Pre-Concept Inventory	12.292	2.167	0.312	5.673	**0.000**

Table 5.5: Multiple Regression Fall 2017. (Top 33% of Concept Map Treatment (n = 28) and Journal Control Group (n = 25). Dependent Variable: Final Exam). Group indicates coded treatment/control. Journal Group = 0; Concept Map = 1 (n = 53).

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	-16.215	103.404		-0.157	0.876
Group	0.656	12.846	0.006	0.051	0.960
Asian	15.612	34.756	0.127	0.449	0.656
Hispanic or Latino	-1.814	37.170	-0.014	-0.049	0.961
White	-50.530	56.499	-0.132	-0.894	0.377
Gender	-27.596	12.965	-0.245	-2.128	**0.040**
High School GPA	25.417	23.267	0.129	1.092	0.281
SAT Math	0.314	0.079	0.501	3.958	**0.000**
Pre-Concept Inventory	8.509	4.539	0.263	1.875	0.068

Table 5.6: Multiple Regression Fall 2017. (Full Class Dependent Variable: Post-Concept Inventory). Group indicates coded treatment/control. Journal Group = 0; Concept Map = 1.

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	-3.691	1.610		-2.292	0.023
Group	-0.006	0.241	-0.001	-0.024	0.981
Asian	0.199	0.461	0.041	0.431	0.667
Black or African American	0.516	0.838	0.036	0.616	0.538
Hispanic or Latino	0.667	0.482	0.132	1.384	0.168
White	0.157	0.637	0.016	0.246	0.806
Gender	-1.505E-05	0.000	-0.004	-0.085	0.932
High School GPA	-0.199	0.317	-0.031	-0.629	0.530
SAT Math	0.014	0.002	0.499	8.879	**0.000**
Pre-Concept Inventory	0.640	0.087	0.396	7.371	**0.000**

Table 5.7: Multiple Regression Fall 2017. (Top 33% of Concept Map Treatment (n = 28) and Journal Control Group (n = 25). Dependent Variable: Post-Concept Inventory). Group indicates coded treatment/control. Journal Group = 0; Concept Map = 1.

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	-2.775	3.905		-0.711	0.482
Group	0.380	0.485	0.084	0.783	0.438
Asian	-0.497	1.312	-0.098	-0.379	0.707
Hispanic or Latino	0.462	1.404	0.084	0.329	0.744
White	-2.882	2.134	-0.183	-1.351	0.184
Gender	-0.559	0.490	-0.121	-1.142	0.260
High School GPA	-0.497	0.879	-0.061	-0.565	0.575
SAT Math	0.017	0.003	0.652	5.610	**0.000**
Pre-Concept Inventory	0.420	0.171	0.315	2.448	**0.019**

Student self-reported learning gains were compared between the concept map and journal groups using the SALG survey. Though chi-squared analysis indicated a significantly higher proportion of positive responses was observed for the post-SALG compared to the pre-SALG within each group (Tables 5.8 and 5.9), a chi-squared analysis comparing the proportion of positive responses on the post-SALG found there was no significant differences between the concept map treatment and journal control groups (Tables 5.10 and 5.11). The fact the concept map intervention did not appear to impact student learning gains related to conceptual understanding or self-reported gains in

connecting the course concepts to other ideas is attributed to the lack of guidance provided to the students as they progressed through the course. This likely resulted in less well-developed concept maps, and also likely resulted in a general lack of engagement on the part of the students.

Table 5.8: Treatment Group (Concept Map) pre-SALG vs post-SALG: Chi-Square Fall 2017.

Survey Q	X^2	p	Cramer's-V
Pre-SALG:[a] Presently I understand… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in your UNDERSTANDING of each of the following…**			
1. Intermolecular Forces	55.405	**.000**	0.559
2. How ideas we will explore relate to ideas I have encountered in classes outside of this subject area	55.755	**.000**	0.561
3. How well do you understand the relevance of chemistry in biology	7.492	**0.024**	0.205
Pre-SALG:[a] Answer the following questions… **Post-SALG:[b] Answer the following questions…**			
4. How important is the study of chemistry in addressing the real world issues	6.768	**0.034**	0.195
5. How well do you understand how chemistry affects your everyday life	1.777	0.411	0.099
Pre-SALG:[a] Answer the following questions… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following…**			
6. Connecting key ideas I learn in my classes with other knowledge	21.362	**0.000**	0.344
Pre-SALG:[a] Presently I am in the habit of… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following?**			
7. Using systematic reasoning in my approach to problems	9.986	**0.007**	0.235

[a]Likert Scale: 1 = not at all, 2 = just a little, 3 = somewhat, 4 = a lot, 5 = a great deal.
[b]Likert Scale: 1 = no gains, 2 = a little gain, 3 = moderate gain, 4 = good gain, 5 = great gain.

Table 5.9: Control Group (Journal) pre-SALG vs post-SALG: Chi-Square Fall 2017.

Survey Q	χ^2	p	Cramer's-V
Pre-SALG:[a] Presently I understand… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in your UNDERSTANDING of each of the following…**			
1. Intermolecular Forces	60.015	**.000**	0.620
2. How ideas we will explore relate to ideas I have encountered in classes outside of this subject area	46.868	**0.000**	0.556
Pre-SALG:[a] Answer the following questions… **Post-SALG:[b] Answer the following questions…**			
3. How well do you understand the relevance of chemistry in biology	8.396	**0.015**	0.230
4. How important is the study of chemistry in addressing the real world issues	6.632	**0.036**	0.206
5. How well do you understand how chemistry affects your everyday life	1.300	0.522	0.090
Pre-SALG:[a] Answer the following questions… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following…**			
6. Connecting key ideas I learn in my classes with other knowledge	24.883	**0.000**	0.394
Pre-SALG:[a] Presently I am in the habit of… **Post-SALG:[b] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following?**			
7. Using systematic reasoning in my approach to problems	17.748	**0.000**	0.332

[a]Likert Scale: 1 = not at all, 2 = just a little, 3 = somewhat, 4 = a lot, 5 = a great deal.
[b]Likert Scale: 1 = no gains, 2 = a little gain, 3 = moderate gain, 4 = good gain, 5 = great gain.

Table 5.10: Control Group (Journal) vs Treatment Group (Concept Map) pre-SALG: Chi-Square Fall 2017.

Survey Q	X^2	p	Cramer's-V
Pre-SALG:[a] Presently I understand…			
1. Intermolecular Forces	1.267	0.531	0.081
2. How ideas from this class relate to ideas encountered in other classes within this subject area	0.153	0.926	0.028
Pre-SALG:[a] Answer the following questions…			
3. How well do you understand the relevance of chemistry in biology	5.221	0.073	0.162
4. How important is the study of chemistry for solving real world issues	2.564	0.278	0.114
5. Using a critical approach to information and arguments I encounter in daily life	4.502	0.105	0.150
Pre-SALG:[a] Answer the following questions…			
6. Connecting key class ideas with other knowledge	3.205	0.201	0.126
Pre-SALG:[a] Presently I am in the habit of…			
7. Using systematic reasoning in my approach to problems	3.209	0.201	0.126

[a]Likert Scale: 1 = not at all, 2 = just a little, 3 = somewhat, 4 = a lot, 5 = a great deal.

Table 5.11: Control Group (Journal) vs Treatment Group (Concept Map) post-SALG: Chi-Square Fall 2017.

Survey Q	X^2	p	Cramer's-V
Post-SALG:[a] As a result of your work in this class, what GAINS DID YOU MAKE in your UNDERSTANDING of each of the following…			
1. Intermolecular Forces	0.190	0.910	0.037
2. How ideas from this class relate to ideas encountered in other classes within this subject area	0.206	0.902	0.039
Post-SALG:[a] Answer the following questions…			
3. How well do you understand the relevance of chemistry in biology	1.141	0.565	0.091
4. How important is the study of chemistry for solving real world issues	2.074	0.355	0.122
5. Using a critical approach to information and arguments I encounter in daily life	1.644	0.440	0.109
Post-SALG:[a] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following…			
6. Connecting key class ideas with other knowledge	0.196	0.907	0.037
Post-SALG:[a] As a result of your work in this class, what GAINS DID YOU MAKE in INTEGRATING the following?			
7. Using systematic reasoning in my approach to problems	0.884	0.643	0.079

[a]Likert Scale: 1 = no gains, 2 = a little gain, 3 = moderate gain, 4 = good gain, 5 = great gain.

5.4 Conclusions - F17

In summary, the Fall 2017 implementation did not show any significant difference between the treatment and control group in exam scores, self-reported gains in conceptual understanding, or self-reported gains in connecting chemistry to other disciplines. Statistical analysis of either the pre/post-test or pre-test/final exam did not reveal any significant improvement in student learning gains. When controlling for factors outside of the control of this experiment, no additional learning gains were observed. The top 33% of students in each group as determined by their concept map or journal scores were selected for analysis to determine if performing well on the concept mapping activity leads to higher scores on the concept inventory post-test of final exam. No significant difference in student performance was observed within the top 33% population comparison. Based on these results it would appear that concept mapping is not a useful tool to improve student learning gains. Using a chi-squared analysis of the SALG survey questions revealed that students within either the control or treatment group were positively impacted by the course, reporting gains in questions 1-4, 6, and 7. However, when comparing responses between groups on the pre- and post-SALG survey responses, no higher gains were observed for either the control or treatment.

Anecdotal feedback from the end-of-course surveys indicated the concept maps were confusing for some students, and many students reported difficulty in deciding how to start constructing the maps. It is also noted many students reported being uncertain about whether their concept map was correctly structured or how to effectively utilize the concept map in their learning. In total, these limitations in the implementation appear to be reflected in the quasi-experimental data analysis, as student learning gains were observed only within the groups and it appears that the concept map treatment had no significant effect

compared to the journal control group. These results suggested further refinement of the concept map implementation was required in order to observe measurable gains in conceptual understanding.

5.5 Introduction and Changes to the Implementation - S18

Based on the student comments from the F17 implementation, the primary goals for improving the concept map intervention focused on providing a template to start the concept maps, providing periodic feedback on their maps, and emphasizing how the maps could be used as a study tool. It was thought these changes to the implementation would reduce student confusion and uncertainty about the purpose of the concept maps, and ultimately increase the level of engagement with the assignment. First, a pre-built "base" concept map (Figure 5.2) was provided at the beginning of the quarter to ensure the students included all of the major concepts to be covered in the course (i.e., the topic for each chapter). Second, the students were provided weekly feedback in their recitation sections to help identify incorrect connections observed on selected concept maps, and to discuss examples of well-developed connecting terms. Finally, the students were shown an example of well-developed concept map following the first exam in an effort to model for the students how the concepts maps could be used as an exam study tool. A more detailed description of the changes to the implementation is provided. Following consultation with an expert in the field, it was determined that the original statistical analyses performed were inadequate for this type of data. As such, changes to the types of statistical tests will be highlighted in the experimental design section below.

Figure 5.2: Concept Map provided as outline for Spring 2018 implementation.
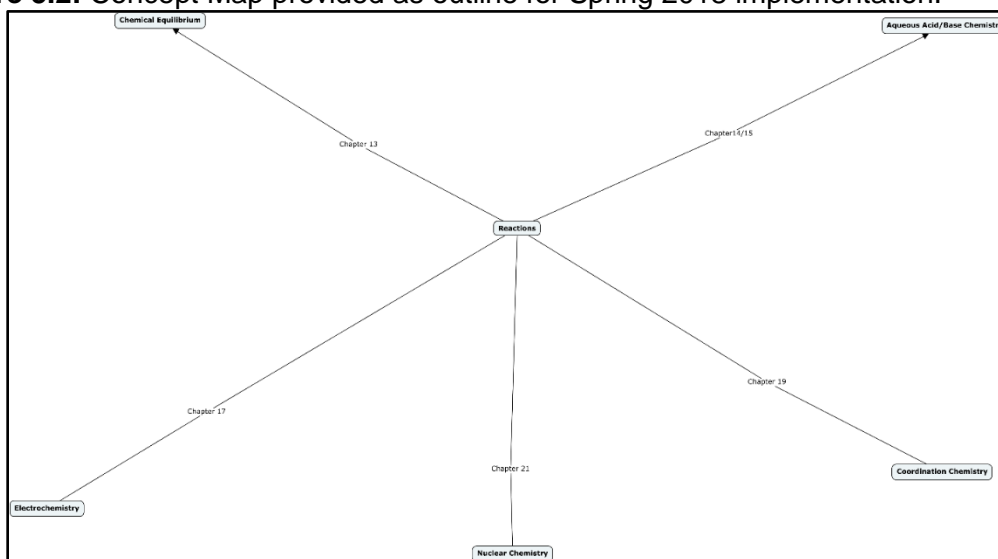
5.6 Experimental Design - S18

The second implementation of the concept map treatment was administered in the S18 CHEM 001C course. The concept map implementation was carried out in a third quarter general chemistry course (CHEM 001C). This is the third course in the three-quarter general chemistry sequence offered at UCR. The topics covered in this course include: chemical equilibrium; acid-base chemistry; buffers and titrations; electrochemistry; coordination chemistry; and nuclear chemistry. This course was taught for an "on sequence" cohort of first-year students in the spring of 2018 (S18), but also included second year students and upper-class students who may have needed the course as a general college science requirement or a pre-requisite for health professional schools.

The S18 course consisted of a large enrollment lecture, which met two times per week for 80 minutes, and associated recitation sections that met once per week for 50 minutes (30-40 students each). Approximately one-third of the lecture meetings used flipped classroom modules,[32] while the remaining class periods included a mixture of lecture, peer-to-peer discussion, and interactive clicker questions. The quasi-experimental design

assigned recitation sections taught by a TA who had prepared in advance to implement the concept map intervention as the treatment group (n = 115). The remaining recitation sections, taught by a different TA, were assigned as the control group and required the students to complete a weekly journal entry in lieu of a concept map (n = 123). The two TAs had a similar experience teaching the discussion group sections and had collaborated in preparing classroom activities for previous offerings of the discussion group sections attempting to minimize potential an effect on student learning outcomes. This design allowed the instructor to assign similar workloads and use the same grading scheme in both groups. Students in both the treatment and control were instructed to complete the weekly journal or concept maps as previously described in Section 5.2.4.

5.6.1 Concept Map Outline and Sample Maps - S18

Students were provided an outline of a concept map that they could use to build out the key concepts throughout the course (see Figure 5.2). Aside from being given the preliminary concept outline, students had complete control over their concept maps and were instructed to build their concept maps using the material they had learned each week. This enabled students to have decide how they would build the concept map and in what way they would connect key concepts to one another. Students were also provided weekly class-wide verbal feedback on any key misconceptions that were identified in the weekly evaluation of the concept maps. Following the first exam, an instructor-completed concept map was shown to the students in their discussion sections and the TA led a dialogue to help elucidate for the students how proper connecting terms should be constructed. This provided additional feedback and helped model for the students the types of conceptual thinking that should be used for the remainder of the quarter as they continued to build

113

their maps. The scoring of the concept maps and journals was held constant between the two implementations.

Figures 5.3 and 5.4 illustrate examples of a well-developed and poorly developed concept map, respectively. In Figure 5.3, the student used a color-coded system to connect subtopics to broader concepts and the concept map has several branches linking the broader concepts to each other. The connecting phrases between the various key concepts are also well developed and almost universally accurate. The concept map shown in Figure 5.4 lacks many of the key concepts covered in each chapter and does not include many of the sub-topics observed in the more developed concept maps. Additionally, many of the connections that should have been made between the broader key concepts were not observed (e.g., other than a single connection between ICE table and salt solutions, there are no connections drawn between any of the other chapters). It is worth noting, unlike the concept maps from the F17 implementation shown in Figure 5.1, these maps are much easier to follow due to the base map provided.
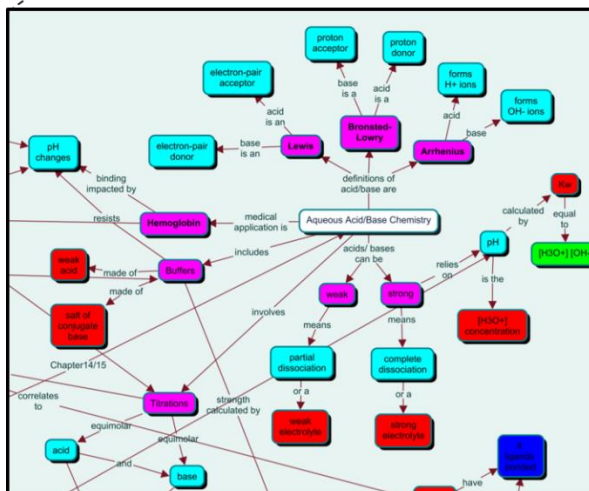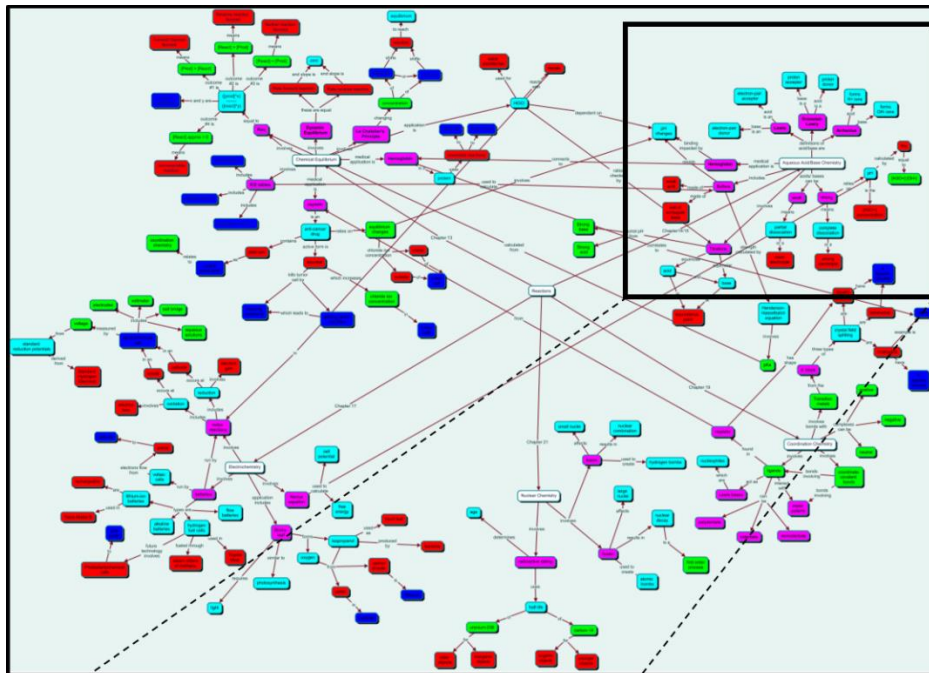
Figure 5.3: Example of a well-developed concept map with correct connections and connecting phrases (Comprehensiveness = 2, Organization/Links = 3, Correctness = 3, Total score = 8/9)
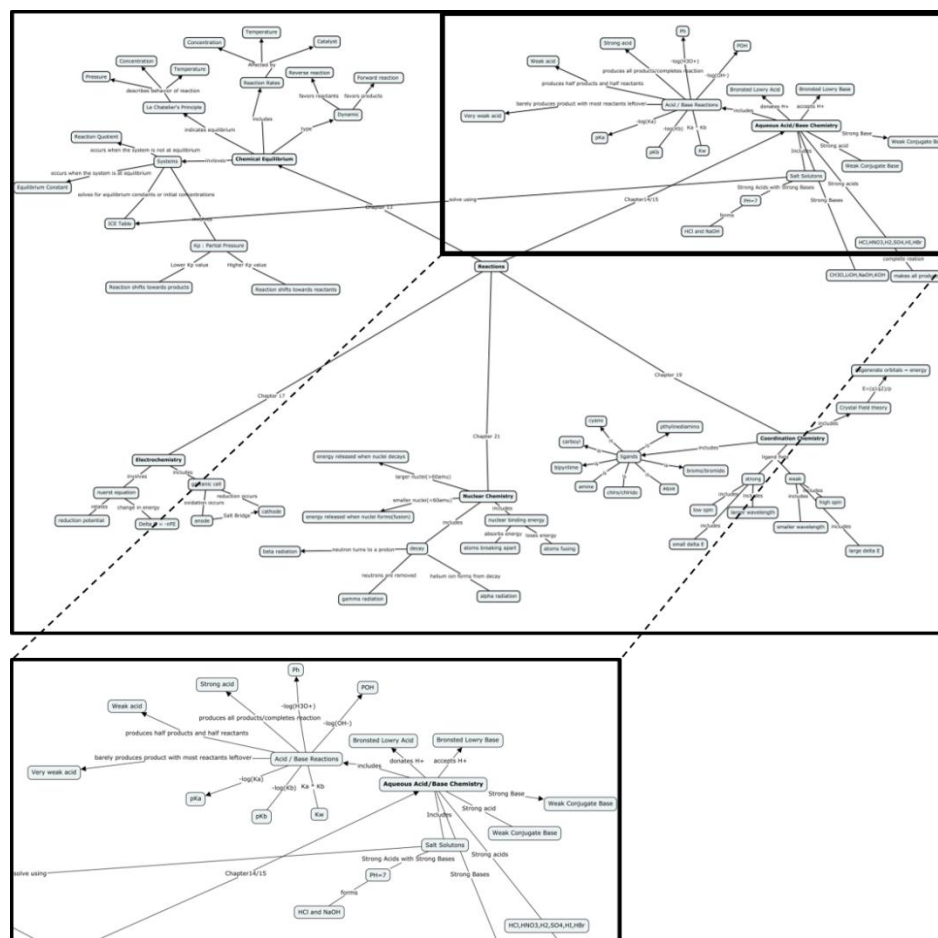
Figure 5.4: Example of a less well-developed concept map with significant numbers of incorrect connections and connecting phrases (Comprehensiveness = 1, Organization/Links = 1, Correctness = 1, Total score = 3/9).

5.6.2 Concept inventory, SALG survey, and Data Collection - S18

The concept inventory consisted of a total of 16 multiple choice questions and was graded out of a total score of 16 (1 point for each item). This concept inventory was administered as a pre-test for students in the treatment and control groups in the first week of recitation. The pre-test was administered using the online quiz function in the course management system. Students were not able to access the quiz once it was completed and they were not allowed to view the questions or answers at any point during the quarter. The same questions from the concept pre-test were placed into the final exam for students

in both the treatment and control groups, and this post-test was used to measure the improvement in conceptual understanding throughout the ten-week quarter. Any subsequent references to the concept inventory "post-test" are restricted to the 16 concept inventory questions that were embedded in the final exam. Any references to the "final exam" include the entire exam, which consisted of the 16 concept inventory questions, 14 additional algorithmic multiple-choice questions, and five free response questions. Finally, the authors note it would have been ideal to use a previously validated concept inventory to probe student gains in conceptual understanding. However, the unique combination of topics covered in this course dictated the use of a customized concept inventory that better matched the respective course learning objectives. Because the concept inventory contained multiple dimensions (i.e., five different categories of concepts) the stratified alpha (α_s) reliability coefficient was calculated as described by Widhiarso and Ravand,[33] and an item-analysis was carried out for the 16 test items (Tables 5.12-5.14).

The collection of the SALG data at the beginning and end of the quarter was the same as described for the F17 implementation. Questions relating to the topics covered in the S18 implementation were added to the SALG model. These questions focused on conceptual learning gains, learning gains related to specific course topics, and learning gains related to applying course concepts in other contexts were chosen for the purpose of this study (see Table 5.23). Students were again informed that completion of both the pre- and post-SALG would provide a small amount of extra credit to their grade (5 out of 1000 total course points). Student responses to the previously validated SALG survey instrument were collected and compared between treatment and control groups. Unlike in the F17 implementation, the S18 implementation pre- and post-SALG responses were paired for all students.

Table 5.12: Stratified alpha reliability coefficient for concept inventory.

Stratified Alpha (α_s) = 1-[Σ(variance of each item)(1- α_i)]/(variance of all items)
0.661 (n = 16; variance of all test items = 8.127)
Test Item Concept Dimensions (α and variance of each dimension)
Items 1, 13 (Electrochemistry); α_i = 0.344; variance = 0.327
Items 2-4, 10 (Equilibrium); α_i = 0.285; variance = 1.257
Items 5-9 (Acid-base Chemistry); α_i = 0.454; variance = 1.591
Items 11-12 (Coordination Chemistry); α_i = 0.429; variance = 0.567
Items 14-16 (Nuclear Chemistry); α_i = 0.348; variance = 0.683

Table 5.13: Item analysis for concept inventory.

Question	Mean	Std. Deviation	N
Q1	0.8554	0.35245	242
Q2	0.6198	0.48643	242
Q3	0.5289	0.50020	242
Q4	0.4793	0.50061	242
Q5	0.3926	0.48933	242
Q6	0.7190	0.45042	242
Q7	0.8512	0.35659	242
Q8	0.7190	0.45042	242
Q9	0.6033	0.49023	242
Q10	0.5000	0.50104	242
Q11	0.7149	0.45241	242
Q12	0.4008	0.49108	242
Q13	0.8223	0.38304	242
Q14	0.8512	0.35659	242
Q15	0.6736	0.46988	242
Q16	0.7727	0.41994	242

Table 5.14: Item-total correlation for concept inventory.

Question	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q1	0.365	0.614
Q2	0.319	0.616
Q3	0.373	0.607
Q4	0.241	0.628
Q5	0.065	0.654
Q6	0.236	0628
Q7	0.128	0.641
Q8	0.243	0.627
Q9	0.144	0.643
Q10	0.110	0.648
Q11	0.232	0.629
Q12	0.206	0.633
Q13	0.368	0.612
Q14	0.310	0.620
Q15	0.384	0.606
Q16	0.407	0.605

5.6.3 Statistical Analysis - S18

New statistical analyses were incorporated into the experiment to provide a more complete interpretation of the data obtained. To test the quasi-experimental hypotheses on the S18 implementation, student performance on the pre- and post-test concept inventory, and self-reported gains in conceptual understanding on the pre- and post-SALG survey were compared between the treatment and control groups. The similarities and changes between the data analysis for the S18 implementation are discussed below.

The incoming academic preparation of the students in the concept map treatment and journal entry control groups was compared using an analysis of variance (ANOVA). The

ANOVA was also used to compare the concept map rubric scores between the F17 and S18 treatment groups. Concept inventory pre/post-test gains within the treatment and control groups were analyzed using paired t-tests, and a comparison of the performance on the post-test concept inventory between the treatment and control group was carried out using a multiple linear regression model. An independent samples t-test was used to compare the mean scores on the concept inventory pos-test between the groups to determine if conceptual learning gains were obtained for the treatment group over the control group. Multiple regression analyses were performed on the concept map and control groups as previously described for the F17 implementation. Due to the final exam containing a significant number of algorithmic problems that are likely less correlated to conceptual thinking, a multiple regression analysis with "final exam" as the dependent variable was not included in any of the analyses.

Changes were made between the F17 and S18 treatment of the SALG data. Following consultation with an expert in chemical education and statistics, it was determined that a chi-squared analysis of the survey question posed challenges in drawing appropriate conclusions. This was due to nine questions asked, each focusing on a different topic. This leads to complications with the chi-squared analysis and could result in a false positive being reported. As such, a qualitative approach is used to analyze the SALG survey responses. The relationship between the ability to construct high quality concept maps or journal entries and performance on the concept inventory post-test was also evaluated by determining the Pearson correlation coefficient. These additional statistical analyses were performed in order to improve the student data analysis and make more reliable conclusions about the data collected.

5.7 Results and Discussion - S18

Table 5.15 summarizes the descriptive statistics for the control and treatment groups participating in the quasi-experiment. The academic preparation of students in the treatment and control groups appeared to be equivalent when considering the distribution of pre-test scores and high school GPAs within each group, however the concept map treatment group did appear to have slightly higher math SAT scores compared to the journal control group. The concept inventory pre-test scores, math SAT scores, and overall high school GPAs were compared using an ANOVA, and these results appear to corroborate the notion there is no significant difference in the incoming academic abilities of the students between the two populations as measured by high school GPA and concept inventory pre-test (the null hypothesis stating there is no difference between the mean scores for all three of these independent variables cannot be rejected at the $p = 0.05$ level; see Table 5.16). However, the null hypothesis stating the math SAT scores are equivalent between groups can be rejected ($F = 6.681$, $p = 0.010$; see Table 5.16).

Table 5.15: Descriptive Statistics for the recitation section treatment and control groups (Control = journal group; Treatment = concept map).

	Control (n = 115)	Treatment (n = 123)
High School GPA Avg.	3.74 +/- 0.263	3.77 +/- 0.27
SAT Math Avg.	582 +/- 80	612 +/- 85
Pretest Avg. (out of 16)	5.57 +/- 2.07	6.09 +/- 2.60
Post Test Avg. (out of 16)	9.91 +/- 2.81	10.76 +/- 2.72
Exam 1 Avg. (out of 100)	64.01 +/- 15.75	65.98 +/- 15.27
Exam 2 Avg. (out of 100)	75.09 +/- 15.90	78.45 +/- 14.48
Final Exam Total Avg. (out of 400)	244.07 +/- 59.96	253.11 +/- 59.02
Avg. rubric score on final concept/journal entry (out of 9)[a]	7.00+/- 1.23 (n = 100)	6.17 +/- 1.55 (n = 97)
Asian	n = 46	n = 58
Black or African American	n = 3	n = 0
Hispanic or Latino	n = 45	n = 36
Multi-Race/Unknown/Non-Resident Alien	n = 11	n = 13
White	n = 9	n = 16
Male	n = 55	n = 58
Female	n = 58	n = 61
Gender Not Reported	n = 2	n = 2

[a]Not all students submitted a final concept map/journal. The n value designates the number of students that completed the final concept map/journal assignment.

Table 5.16: ANOVA analyses; comparison of concept map inventory pre-test scores, math SAT scores, and high school GPA between selected groups.

Concept Map Treatment vs. Journal Control	F	p
Concept Inventory Pre-test Dependent Variable	2.731	0.100
SAT Math Dependent Variable	6.681	0.010
High School GPA Dependent Variable	0.753	0.386
Top 33% of Concept Map Students (n = 41) vs. Top 33% of Journal Students (n = 38)[a]	**F**	**p**
Concept Inventory Pre-test Dependent Variable	0.787	0.378
SAT Math Dependent Variable	0.070	0.793
High School GPA Dependent Variable	0.117	0.733
Concept Map Treatment (n=123) vs. Concept Map Top 33% (n=41)[a,b]	**F**	**p**
Concept Inventory Pre-test Dependent Variable	0.081	0.776
SAT Math Dependent Variable	0.057	0.811
High School GPA Dependent Variable	0.708	0.402

[a]The top 33% of students in the concept map treatment group and journal control group were determined using scores from the concept map rubric.
[b]mean values for dependent variables: concept map entire class SAT math = 612 +/ 85; concept map top 33% SAT math = 608 +/- 83; concept map entire class GPA = 3.77 +/ 0.27; concept map top 33% GPA = 3.81 +/- 0.24; concept map entire class concept inventory pre-test = 6.09 +/ 2.60; concept map top 33% concept inventory pre-test = 6.24 +/- 3.00.

Though a statistical comparison of the concept map rubric scores between the F17 and

S18 study participants is extremely limited due to the variety of independent variables that

cannot be controlled (e.g., courses taken prior to the S18 term, improved student maturity

from the F17 term to the S18 term, etc.), the average rubric scores for the final concept

map submission suggest the concept map treatment group from the S18 implementation

scored higher on the post-test than the concept map treatment group from F17 (F17 average = 4.04 +/- 2.85; S18 average = 5.02 +/- 2.78; see Tables 5.2 and 5.15). An ANOVA analysis examining the difference in concept map rubric scores between the F17 and S18 treatment groups was run to gain some sense if there was a significant improvement with the S18 implementation. The difference in average concept map rubric scores was significant at the p = 0.10 level (F = 2.755, p = 0.098; see Table 5.17). Though it is not possible to verify if the improved concept map intervention was the primary contributor to the improved rubric scores with the S18 cohort, these data do suggest the changes to the concept map implementation may have had some impact on helping students create more well-developed concept maps.

Table 5.17: ANOVA analysis of F17 vs S18 final concept map scores (dependent variable input as a numerical score of 0-9 on the concept map rubric).

	F	p	η^2
F17 concept map treatment versus S18 concept map treatment.	2.755	**0.098**	0.011

The comparison of the concept inventory pre- and post-test suggests students in the treatment and control groups made gains in conceptual understanding during the course of the term. A paired t-test in which the average pre- and post-test concept inventory scores were compared indicates a statistically significant increase in average score for the post-test within both the treatment and control groups (journal control: mean difference = 4.389; t = -15.058; p < 0.001; concept map treatment: mean difference = 4.610; t = -16.302; p < 0.001; see Table 5.18). These results are important to note because, although the concept inventory questions were not selected from a previously validated question set, this instrument appears to measure gains in conceptual understanding with statistically significant results for both groups. The general utility of the concept inventory

was confirmed by an item analysis and a single-administration internal consistency analysis. The item analysis indicates all of the questions on the inventory exhibited an item difficulty between 0.30-0.85 (Table 5.13), the item discrimination analysis revealed 12 of the 16 questions possess a point-biserial correlation greater than 0.20 (Table 5.14), and the stratified alpha reliability coefficient (α_s) was found to be 0.661, suggesting moderate reliability for the concept inventory (Table 5.12).

Table 5.18: Paired t-test comparing concept inventory pre-test and post-test scores within the treatment and control groups for the concept map implementation.

	n	Mean Difference	Std. Deviation	Std. Error Mean of Difference	p-value
Journal Control Pre- vs Post-test mean	108	4.389	3.029	0.291	**< 0.001**
Concept Map Treatment Pre- vs Post-test mean	118	4.610	3.072	0.283	**< 0.001**

To begin evaluating the first research hypothesis, in which it was stated the concept map treatment would yield greater gains in conceptual learning relative to the journal control group, an independent samples t-test was used to compare the mean scores on the concept inventory post-test between the two groups (see Table 5.19). Based on this independent samples t-test, it appears the concept map treatment group performed better on the post-test relative to the journal control group (mean difference = 0.844, p = 0.020). Though the mean difference between groups was less than one point, this level of improvement would translate to an approximately 5% increase in exam performance for the concept map treatment group. However, because the study groups in this quasi-experiment were not randomly assigned caution must be taken in over-interpreting these results. More specifically, because the assignment of the study groups was not

randomized, relying solely on the comparison of means in which no background academic traits were statistically controlled was not prudent. Therefore, it was desired to carry out further analysis in which the academic preparation characteristics among the treatment and control groups could be statistically controlled.

Table 5.19: Independent samples t-test comparing concept inventory post-test between the concept map treatment and journal control groups.

	t	df	Mean Difference	Std. Error Mean of Difference	p-value
Journal Control (n = 115; mean = 9.91) vs Concept Map Treatment (n = 123; mean = 10.76)	2.344	235	0.844	0.360	**0.020**

5.7.1 Multiple Regression Analyses - S18

Even though the quasi-experimental design is not able to control for the myriad of independent variables that likely impact student performance, including incoming academic preparation, a multiple regression analysis can be utilized to statistically control for these independent variables. Given the fact the math SAT scores appear to be slightly higher for the concept map treatment group, it was especially pertinent to hold this variable constant when comparing the concept inventory post-test scores between the treatment and control groups. The model employed here included three different independent variables: the concept inventory pre-test scores, overall high school GPAs, and math SAT scores. Because the quasi-experimental hypotheses are focused on determining the impact of the concept map treatment on conceptual understanding, the linear regression model included the concept inventory post-test as the dependent variable.

In building a linear regression model, it is imperative to avoid using a predictive model to answer an explanatory research question.[34] Thus, because it was hypothesized that

high school GPA, math SAT scores, and the concept inventory pre-test scores would likely correlate with performance on the concept inventory post-test, these co-variates were held constant when comparing the post-test score dependent variable between the treatment and control groups. Generally speaking, high school GPA is known to be a strong predictor of student success in college-level courses,[35] and in fact has been shown to be a stronger predictor of success than standardized tests such as the ACT and SAT.[36] However, a recent report suggests math SAT scores can be a strong predictor of student performance in general chemistry,[37] which suggests this variable is likely to be positively correlated to the dependent variable. These previous studies therefore suggest including both high school GPA and math SAT scores as co-variates in the regression model should help isolate the impact of the group participation (treatment vs. control) on the post-test scores. Including the concept inventory pre-test scores as a covariate fits within the creation of the explanatory model, as this assessment more directly measures existing conceptual understanding, and holding this variable constant is also expected to aid in isolating the impact of the concept map treatment on the final concept inventory performance. Though demographic characteristics can also positively correlate to student classroom performance, these were not included as co-variates in the explanatory regression model employed here because these independent variables would likely be redundant after accounting for the afore-mentioned predictors of academic performance. Indeed, when a model that also included gender and ethnic categorical variables was created, the fit of the regression did not appreciably improve and none of the correlation coefficients for the ethnic and gender categories were significant at the $p = 0.05$ level.

The results of the final multiple regression analysis are shown in Table 5.20. The concept inventory post-test dependent variable output was compared between all

participants in the treatment and control groups, while keeping constant the independent variables related to student academic preparation (high school GPA, SAT math scores, and the concept inventory pre-test scores). Though participation in the treatment group did have a positive correlation to performance on the concept inventory post-test, this result was not statistically significant (unstandardized B = 0.222, p = 0.540). The correlation of high school GPA, math SAT, and concept inventory pre-test also positively correlated to performance on the concept inventory post-test (unstandardized B = 0.174, 0.015, and 0.235, respectively), though interestingly the correlation of high school GPA was not statistically significant. Unsurprisingly, the concept inventory pre-test appeared to have the strongest correlation to performance on the post-test measure.

Table 5.20: Multiple regression analysis. Includes full class; dependent variable = **concept inventory post-test**. Group indicates coded treatment/control (journal control group = 0; concept map treatment group = 1).[a]

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	p
Constant	-1.034	2.918		-0.355	0.723
Group	**0.222**	**0.362**	**0.039**	**0.614**	**0.540**
High School GPA	0.174	0.674	0.016	0.258	0.796
SAT Math	0.015	0.002	0.439	6.715	0.000
Concept inventory pre-test	0.235	0.080	0.439	2.940	0.004

[a]$R = 0.534$; $R^2 = 0.285$; adjusted $R^2 = 0.270$; Standard error of the estimate = 2.462

To determine if more well-developed concept mapping skills might be correlated to improved performance on the concept inventory post-test, an additional multiple regression analysis was carried out in which only the top 33% of students from the treatment and control groups were included in the group independent variable (see Table

5.21 for the descriptive statistics for these treatment and control group populations). The same model used for the analysis of the entire treatment and control group populations was used for this analysis; the performance on the concept inventory post-test was identified as the dependent variable, while the concept inventory pre-test, math SAT scores, and high school GPAs were held constant. It is noted ANOVA was used to determine if there was a significant difference in the incoming academic preparation between the top 33% of students in the concept map group versus the top 33% of students in the journal control group, and these analyses indicate the null hypothesis stating there were no differences in the mean values of these independent variables could not be rejected (concept inventory pre-test, $F = 0.787$, $p = 0.378$; math SAT, $F = 0.070$, $p = 0.793$; high school GPA, $F = 0.787$, $p = 0.117$; see Table 5.16).

Table 5.21: Descriptive Statistics for the recitation section treatment and control groups; top 33% of students based on concept map/journal rubric score (Control = journal group; Treatment = concept map).

	Control (n = 38)	Treatment (n = 41)
High School GPA Avg.	3.78 +/- 0.24	3.81 +/- 0.24
SAT Math Avg.	614 +/- 84	608 +/- 83
Pretest Avg. (out of 16)	5.71 +/- 1.90	6.24 +/- 3.00
Post Test Avg. (out of 16)	10.65 +/- 2.20	11.83 +/- 2.42
Exam 1 Avg. (out of 100)	68.58 +/- 14.44	69.80 +/- 15.34
Exam 2 Avg. (out of 100)	80.08 +/- 14.89	82.32 +/- 12.82
Final Exam Total Avg. (out of 400)	263.05 +/- 55.32	276.41 +/- 52.39
Avg. rubric score on final concept/journal entry (out of 9)[a]	8.21 +/- 0.62	7.66+/- 0.79
Asian	n = 18	n = 16
Black or African American	n = 0	n = 0
Hispanic or Latino	n = 13	n = 12
Multi-Race/Unknown/Non-Resident Alien	n = 3	n = 6
White	n = 4	n = 7
Male	n = 29	n = 28
Female	n = 9	n = 13
Gender Not Reported	n = 0	n = 0

[a]Not all students submitted a final concept map/journal. The n value designates the number of students that completed the final concept map/journal assignment.

The results of the final multiple regression analysis for the top 33% of students from each group are shown in Table 5.22. The concept map treatment group appeared to attain higher scores on the concept inventory compared to the journal entry control group

(unstandardized B = 1.165, p = 0.021). This indicates students from the top 33% of students, as determined by the concept map rubric scores, scored on average 1.165 points higher on the concept map post-test than the top 33% of students in the control group. This is a roughly 7% improvement in the concept inventory post-test scores. These results suggest students who can create well-developed concept maps demonstrate greater learning gains related to conceptual understanding. Analogous to the multiple regression analysis that included the entire treatment and control populations, the correlation of high school GPA, math SAT, and concept inventory pre-test also positively correlated to performance on the concept inventory post-test (unstandardized B = 0.659, 0.012, and 0.297, respectively), though the correlation of high school GPA was again not statistically significant.

Table 5.22: Multiple regression analysis. Includes top 33% of students in concept map treatment (n = 41) and journal control group (n = 38); dependent variable = **concept inventory post-test**. Group indicates coded treatment/control (journal control group = 0; concept map treatment group = 1).[a]

	Unstandardized Coefficients		Standardized Coefficients		
	B	**Std. Error**	**Beta**	**t**	**p**
Constant	-0.680	4.152		-0.164	0.870
Group	**1.165**	**0.491**	**0.240**	**2.375**	**0.021**
High School GPA	0.659	0.988	0.067	0.667	0.507
SAT Math	0.012	0.003	0.384	3.483	0.001
Concept inventory pre-test	0.297	0.109	0.304	2.736	0.008

[a]R = 0.648; R^2 = 0.420; adjusted R^2 = 0.380; Standard error of the estimate = 1.928

The results from the analyses indicated students from the top 33% of the concept map treatment group performed significantly better on the concept map inventory post-test than the top 33% of students from the journal control group when the background characteristics of the students were statistically controlled (see multiple regression analysis in Table 5.22). These results suggest students who were able to create more well-developed concept maps made greater learning gains in conceptual understanding compared to students in the journal control group. However, a limitation in this study arises because it could not be determined if the ability to create well-developed concept maps resulted from the concept map intervention itself, or if students had a pre-existing ability to create high level concept maps upon entering the study. Because concept mapping skills were not measured for incoming students it is simply not possible to reconcile this question in the current study. Though incoming concept mapping could not be measured, the background academic preparation characteristics of the top 33% of students from the concept map population were compared to the overall concept map study group. A comparison of the mean values for the three academic preparation variables via ANOVA indicates the incoming scholastic abilities of the top 33% of students from concept map treatment group were representative of the entire population (concept inventory pre-test, $F = 0.081$, $p = 0.776$; math SAT, $F = 0.057$, $p = 0.811$; high school GPA, $F = 0.708$, $p = 0.402$; see Table 5.16). This analysis does not exclude the possibility that the top 33% of students from the concept map treatment group had some underlying advantage in creating better developed concept map, but does provide some assurance that evaluating the top 33% of students in the multiple regression analysis did not result in the creating a sub-population students with inherently better academic preparation.

The multiple regression analyses described above suggest the ability to create well developed concept maps correlates with performance on the final concept inventory post-test, thus it was of interest to determine how rubric scores on the final concept maps/journal entries correlated to final concept inventory post-test. The Pearson's correlation coefficient (ρ) suggests there is indeed correlation between well-developed concept mapping skills and performance on the final concept inventory post-test (ρ = 0.295, p = 0.003) whereas the quality of the journal entries did not correlate to scores on the final concept inventory post-test (ρ = 0.129, p = 0.209; see Figure 5.5). First, these results suggest the journal entries were an appropriate control group activity, as students who were more engaged in the journaling activity did not appear to benefit in terms of their final concept inventory performance. However, this can confound the analysis of the top 33% between the treatment and control groups as a random assortment of students from the journal group could have been selected. This could be due to the difference in the scoring rubric for the control group as it was not the same rubric used to score the concept maps. Thus, the correlation between the top 33% of the treatment and control may not be an appropriate way to compare the two groups and they should be examined separate from each other. Second, the Pearson's correlation appears to corroborate the notion that students who had better concept mapping abilities performed better on the final concept inventory, though the correlation is relatively weak. This positive correlation does not shed any light on the origin of the quality of concept mapping (i.e., it cannot be determined in this current study if the concept map treatment resulted in the ability to create more well-developed concept maps).
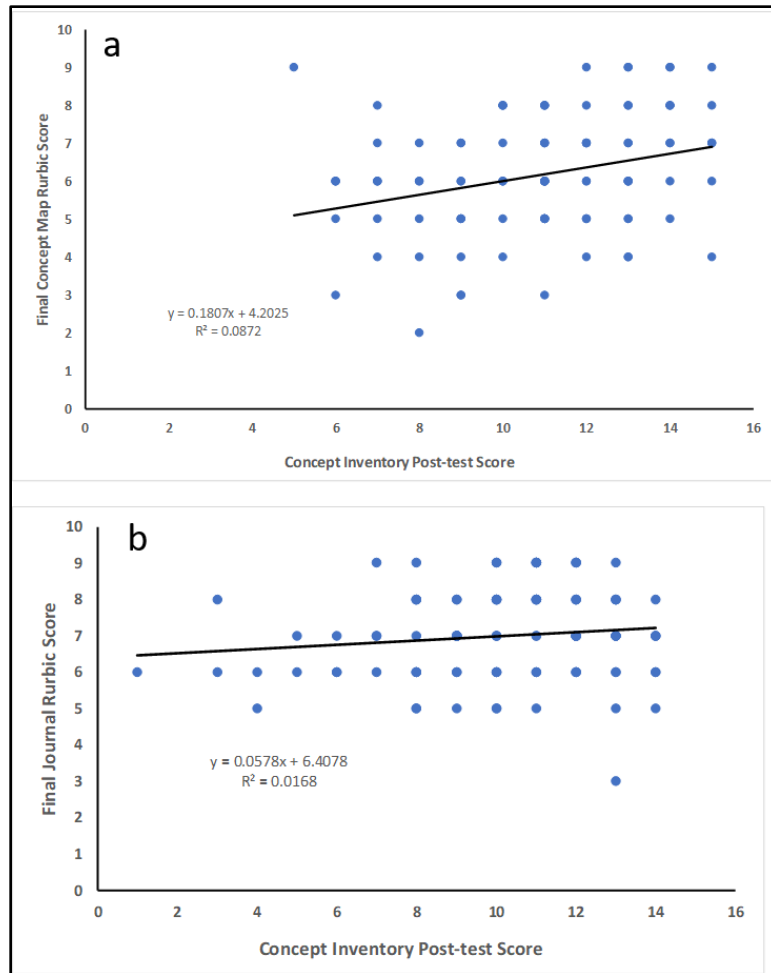
Figure 5.5: Correlation of final concept map/journal entry rubric score with concept inventory post-test score; A (concept map rubric scores vs. concept inventory post-test scores; two-tailed Pearson correlation = 0.295; p = 0.003); B (journal entry rubric scores vs. concept inventory post-test scores; two-tailed Pearson correlation = 0.129; p = 0.209).

Regardless of the limitations in identifying the source of improved concept mapping, it is noteworthy that better concept mapping correlates to conceptual understanding, and a statistically significant difference between the concept map treatment and journal control groups was observed on the concept inventory post-test performance while holding constant three independent variables related to incoming academic preparation. This data suggests finding a way to help students become more proficient in developing concept maps might lead to enhanced conceptual understanding.

5.7.2 Affective Perceptions of Students - SALG Survey - S18

The quasi-experimental design employed in this study relied on delivering the concept map treatment during co-curricular recitation sections taught by graduate TAs. Because this treatment was a relatively small part of the larger course structure, its impact on the students' ability to make gains in conceptual understanding of the course content may have been limited relative to an intervention carried out more frequently in the main lecture. Though the performance on the concept inventory did not appear to differ between the entire treatment and control group populations, it was possible students' self-perceived learning gains and responses to more general questions related to course concepts might be more sensitive to this limited treatment. Therefore, the Student Assessment of Learning Gains (SALG) was used to survey student self-perceived learning gains. Additionally, using a survey instrument such as this provided an opportunity to gauge how students view their ability to connect chemistry concepts to real world issues, their ability to connect chemistry concepts to concepts covered in other STEM disciplines, and their general interest in the course content. These types of affective learning outcomes cannot be measured in a content-based concept inventory but are certainly connected to long-term classroom success.[38] The SALG was administered in this study in a pre-post format in which students estimated their existing knowledge state and affective engagement in the course content prior to the beginning of the course, and then reported on their final learning outcomes at the end of the term.

The SALG assessment was originally developed in 1997 by Elaine Seymour.[39] Using an instrument such as the SALG to measure learning outcomes is supported by previous research that finds student self-reported learning gains strongly correlate to more traditional direct measures of learning.[40,41] The SALG survey questions cover topics

including student understanding of course content, self-evaluation of problem-solving skills and the ability to navigate complex problems, and the integration of their learning both within and outside of the chemistry course. The SALG was administered as part of this study because many of the aspects of learning and student engagement assessed in the SALG were expected to be impacted by a concept map intervention. Table 5.23 summarizes the nine SALG questions used to examine the effectiveness of the concept map treatment. Previous reports indicate students in the second half of the general chemistry course sequence often show difficulty in understanding and retaining information related to chemical equilibrium and acid/base chemistry,[42,43] therefore these specific learning objectives were added into the SALG instrument.

Table 5.23: SALG survey questions that were included in the quasi-experimental analysis.

Survey Questions
Pre-SALG:[a] Presently, I understand… **Post-SALG:[b] As a result of your work in this class, what gains did you make in your understanding of…**
1. Chemical Equilibrium.
2. Acid Base Chemistry.
3. How ideas we will explore in this class relate to ideas I have encountered in other classes within this subject area.
4. How ideas we will explore in this class relate to ideas I have encountered in classes outside of this subject area.
5. How studying this subject helps people address real world issues.
Pre-SALG:[a] Presently, I am in the habit of… **Post-SALG:[b] As a result of your work in this class, what gains did you make in the following skills…**
6. Connecting key ideas I learn in my classes with other knowledge.
7. Applying what I learn in classes to other situations.
8. Using systematic reasoning in my approach to problems.
Pre-SALG:[a] Presently, I… **Post-SALG:[b] As a result of your work in this class, what gains did you make in the following skills…**
9. Feel(ing) comfortable working with complex ideas.

[a]Likert Scale: 1 = not at all, 2 = just a little, 3 = somewhat, 4 = a lot, 5 = a great deal.
[b]Likert Scale: 1 = no gains, 2 = a little gain, 3 = moderate gain, 4 = good gain, 5 = great gain.

The SALG was administered to the entire concept map treatment and journal entry control groups, and the survey respondents represent a sub-population of each group due

to compliance issues related to survey completion. Though the pre- and post-survey responses are paired, the students are identified within the survey in an anonymous fashion, therefore it was not possible to disaggregate survey responses from the top 33% of students from the concept map treatment or journal entry groups (pre/post paired respondents: n = 36 for treatment; n = 18 for control). The average Likert scale responses for all of the pre- and post-SALG questions among both study groups are summarized in Table 5.24, and the distributions of Likert scale responses on the nine post-SALG survey questions are illustrated in Figure 5.6 (questions 1-9). These descriptive results indicate the concept map treatment group had higher proportions of positive responses (5 = strongly agree; 4 = agree) than the control group on all but one of the post-SALG survey questions. However, any interpretation of these survey responses must be tempered given the fact the response rate on the post-SALG survey was quite low and the pre-SALG baseline responses could not be matched for a large number of these respondents. In an effort to compare how many students from each study group made gains on the various survey questions, the percentage of students who made gains was plotted against the number of questions for which an improvement was reported (see Figure 5.7). This analysis indicates that 78% of the concept map treatment students reported improvement in the pre- to post-SALG responses for five or more of the nine questions, whereas 72% of the journal control group students reported such gains. Because of the limited number of students who responded to both the pre- and post-SALG survey these results do not provide unambiguous evidence as to whether the concept map treatment impacted student perceptions of conceptual thinking or affective outcomes.

Table 5.24: Summaries of Liker-scale averages and standard deviations for pre- and post-SALG questions analyzed for the concept map treatment and journal entry control groups.

| Question	Concept pre-SALG	Concept post-SALG	Journal pre-SALG	Journal post-SALG
1. Chemical Equilibrium.	4.0+/-0.97	4.4+/-0.64	3.9+/-1.01	4.3+/-0.71
2. Acid Base Chemistry.	3.6+/-1.03	4.1+/-0.71	3.3+/-0.92	3.8+/-0.90
3. How ideas we will explore in this class relate to ideas I have encountered in other classes within this subject area.	4.0+/-0.78	4.2+/-0.81	4.0+/-1.05	3.9+/-0.90
4. How ideas we will explore in this class relate to ideas I have encountered in classes outside of this subject area.	3.6+/-0.91	4.1+/-0.98	3.6+/-1.04	4.1+/-0.92
5. How studying this subject helps people address real world issues.	4.0+/-1.02	4.3+/-0.73	4.0+/-1.27	4.2+/-0.90
6. Connecting key ideas I learn in my classes with other knowledge.	4.1+/-0.91	4.0+/-0.84	4.0+/-0.96	3.8+/-0.95
7. Applying what I learn in classes to other situations.	4.1+/-0.95	3.8+/-0.94	4.0+/-1.04	3.8+/-1.03
8. Using systematic reasoning in my approach to problems.	4.5+/-0.87	3.9+/-0.92	4.3+/-0.88	3.8+/-1.04
9. Feel(ing) comfortable working with complex ideas.	4.2+/-0.93	3.9+/-1.05	4.1+/-1.03	4.0+/-0.93
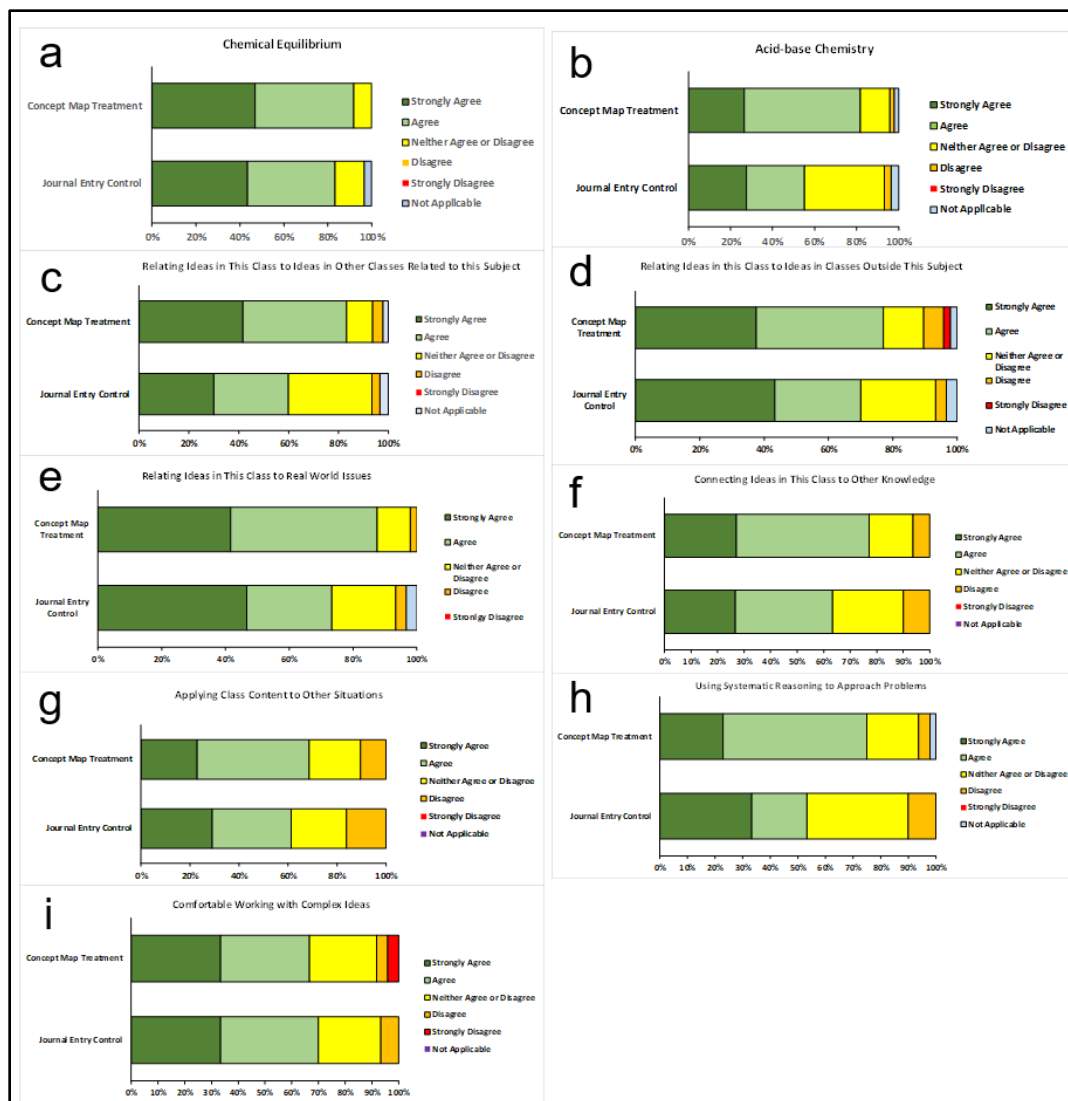
Figure 5.6: Post-SALG responses to SALG questions: a) #1; b) #2; c) #3; d) #4; e) #5; f) #6; g) #7; h) #8; i) #9. Post-SALG sample sizes: treatment = 49, control = 30.
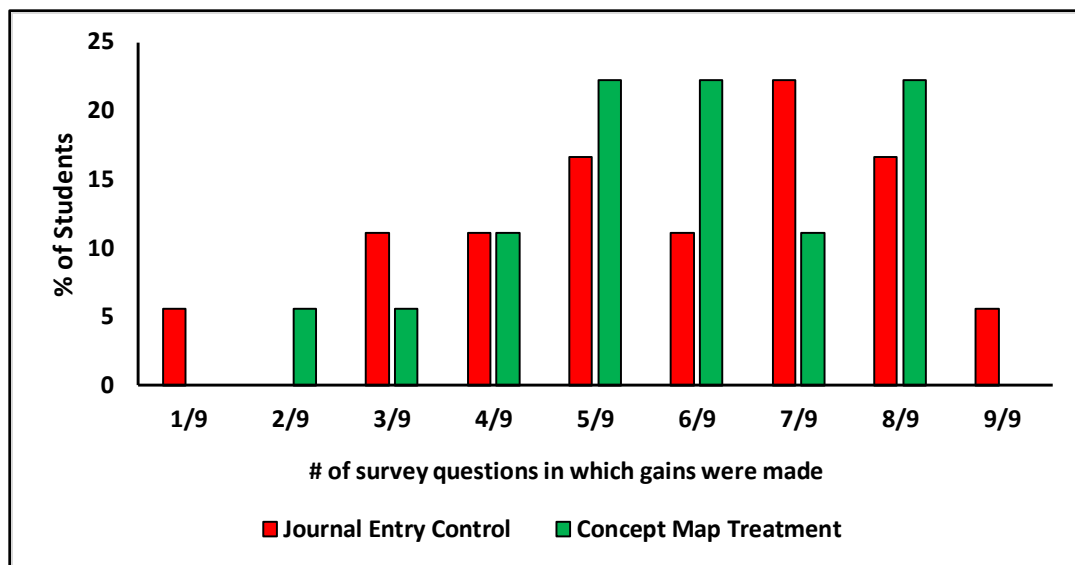
Figure 5.7: The percentage of students who made gains from pre- to post-SALG plotter versus the number of questions in which gains were made (out of a total of 9 questions. The number of respondents for both the pre- and post-SALG was: treatment = 36; control = 18.

Though the SALG survey data did not explicitly reveal the concept map students were impacted in their affective learning outcomes, using this type of data to supplement more quantitative exam or concept inventory data could provide useful insights in future studies. Students who feel they are better prepared may have additional confidence, which in turn helps them overcome test anxiety and perform better on high stakes assessments.[44] One could take the view that even if a concept map treatment were to yield equivalent exam scores compared to traditional assignments, if this type intervention results in gains related to measures of self-confidence, self-efficacy, and the ability to connect chemistry to real world ideas it is likely to positively impact interest/retention in this important gateway course.[38,45]

5.7.3 Limitations of the S18 Study

As described above in the discussion, one of the primary limitations of this study was the fact it was not possible to determine if students found to generate the most well-

141

developed concept maps gained this skill as a result of the study treatment, or if they had pre-existing traits that resulted in higher scores on the concept map rubric. Future studies could attempt to measure incoming concept mapping skills, though this would still pose problems from an experimental standpoint; initial skills in the concept mapping of a pre-existing knowledge state may not translate to concept mapping skills related to the learning objectives covered in the study, and the preliminary concept mapping evaluation itself could impact the performance on the final concept map created by students. Alternatively, concept mapping skills could be gauged more frequently throughout the term to track growth in concept mapping and then be correlated to conceptual understanding, however that could create a barrier in terms of grading workload for a large enrollment course. Nevertheless, creating an experimental design that addresses this limitation should be a priority in future studies.

More traditional threats to internal validity were inherent due to the fact this was a quasi-experimental study conducted within a 10-week long course. Perhaps the most obvious limitation was the fact the treatment and control groups were taught by two different graduate TAs, therefore the instructor effects could not be controlled. Despite possible biases in instructional effectiveness that might have been inherent between the two study groups, the two TAs did have similar experience teaching the discussion group sections and had collaborated in preparing classroom activities for previous offerings of the discussion group sections. These factors likely minimized the potential differences in student learning outcomes usually linked to instructor effects.

History and maturation effects are related to numerous uncontrolled variables that students experience throughout the academic term, and include the different types of course content students may have learned outside of the study, psychological factors that

may impact student confidence and self-efficacy (e.g., poor performance in other courses), and personal life experiences that may limit students' performance in the course related to the study. Though the two-group design helps mitigate these types of validity threat factors, these various independent variables could not be tracked and may have not been equivalent between the two study groups. Tracking and controlling for these types of factors is certainly a daunting task, and often remain unknown in classroom-based studies that take place over quarter- or semester-long time frames. The quasi-experimental design used within the real classroom population in this study also had limitations in terms of non-random selection of study groups and cross-group contamination. The fact the math SAT scores appeared to be non-equivalent between the treatment and control groups suggests the selection process was indeed a limitation in this study (though it is noted incoming academic preparation was statistically controlled in the multiple regression analysis). Design contamination was possible due to the fact students from the treatment and control groups were integrated in the main lecture course and likely discussed their discussion group activities with each other. This cross-group contamination could have also resulted in compensatory rivalry between the study group (i.e., the control group may have felt the treatment group was receiving more desirable classroom activities). Though these design contamination factors were not controlled, the student comments in the final course evaluations did not illuminate any issues with cross-group contamination or compensatory rivalry.

The experimental results from this study are also limited by external threats to validity, therefore caution should be taken when attempting to generalize the impact of the concept map treatment on general chemistry students at other institutions of higher education. The University of California-Riverside (UCR) has a very distinct student body in terms of

demographic make-up.[d] Because UCR has one of the most diverse student populations among larger American research universities, the impact of the concept map treatment observe here may indeed translate differently to classes with less diverse student populations. This study was also conducted in a large enrollment course that had mandatory discussion group recitations associated with the main lecture. Courses with smaller enrollments and lectures lacking discussion groups/recitations may also find the concept map treatment would yield different results in terms of student conceptual understanding.

There was also a limitation in using the series of nine SALG survey questions to compare the self-reported learning gains between the treatment and control groups. Because each question was evaluated independently, any significant differences between concept map and journal entry groups on SALG questions may have arisen due to a family-wise error.[46] In other words, by using a larger number of independent survey questions to measure a general learning outcome, the probability that the analysis of one of these survey questions results in mistakenly rejecting the null hypothesis increases (i.e., a type I error or "false positive"). The use of nine survey questions would result in a family-wise error rate of approximately 20% [family-wise error rate = $1-(1-\alpha)^c$, where α is the desired level for the p-value and c is the number of tests; family-wise error rate for nine survey questions = $1-(1-0.05)^9 = 0.204$]. With that said, one could argue because each SALG survey question was evaluating a different content-based learning objective or affective learning outcome, significant difference in one question/outcome might not necessarily be an artifact of a family-wise error. Finally, though the SALG instrument was used in an attempt to gain insight about student perceptions of affective learning outcomes, compliance issues resulted in a limited subset of study participants who

144

provided feedback. Caution is therefore warranted in interpreting any general trends in the student responses to the SALG survey provided herein. Future studies should aim to increase the participant response rate and might consider including focus group interviews to gather additional data on students' perceptions of affective learning outcomes that could help corroborate data collected in the SALG.

A final limitation of the study is lied in the fact it was not possible to determine if students found to generate the most well-developed concept maps gained this skill as a result of the study treatment, or if they had pre-existing traits that resulted in higher scores on the concept map rubric. Future studies could attempt to measure incoming concept mapping skills, though this would still pose problems from an experimental standpoint; initial skills in the concept mapping of a pre-existing knowledge state may not translate to concept mapping skills related to the learning objectives covered in the study, and the preliminary concept mapping evaluation itself could impact the performance on the final concept map created by students. Nevertheless, creating an experimental design that addresses this limitation should be a priority in future studies.

5.8 Conclusions

In summary, the purpose of this study was to create a streamlined concept map implementation and generate new evidence about the impact of concept map assignments on large enrollment general chemistry courses. Students who were able to generate more well-developed concept maps appear to perform better on concept inventory questions, and self-reported gains in understanding acid/base chemistry and comfort in applying chemistry to real-world issues were also observed by students who participated the concept map treatment. Though the results presented here suggest the ability to create well-developed concept maps correlates to improved learning gains in conceptual

145

understanding, future research should aim to ascertain if there is indeed a causal relationship between conceptual learning and creating well-developed concept maps and how students can be coached into creating more advanced concept maps. These results do suggest that the improved S18 implementation was successful in improving, though slightly, the student conceptual gains and self-confidence, both of which were unobserved in the F17 implementation.

It is also noted the concept map assignment was structured in a way that would allow easy integration into a traditional large enrollment course. More specifically, the goal was to balance making the concept map assignment significant enough to impact student performance without making the administration and grading of concept maps a burden in terms of TA/instructor workload. Additionally, it was desired to avoid creating student resistance to the concept map assignments, which can arise if the subjective grading rubric is used to assign grades for the concept maps. The S18 concept map implementation described in this study can act as a template for instructors who wish adopt this type of learning intervention and assessment, however future implementations might be re-designed to improve student engagement and comfort using concept maps while maintaining the ease of implementation. Increasing student engagement with concept mapping without relying on a subjective grading rubric could be accomplished by coupling the concept map assignment with a close-ended task that assesses students' ability to properly connect concepts. For instance, linking the concept map assignments described here with the Measure of Linked Concepts (MLC's) described by Ye and Lewis[20] or the more recently reported Creative Exercises (CE's) designed by Ye and coworkers[47] would provide a less subjective method of evaluating the students. This would likely reduce student anxiety about being judged on their concept maps while simultaneously increasing

the incentive for students to take the concept map assignment seriously. Including the MLC's or CE's in a quasi-experimental research design might also provide a means to determine if concept mapping skills can be taught to selected student populations and if true gains in concept mapping lead to improved conceptual understanding. Regardless of how instructors might adapt the use of concept maps into their course, making this tool a more prominent component of the instructional toolbox can help students attain the type of meaningful learning described in Ausubel's assimilation theory of learning.

5.9 Endnotes

[a]SALG survey: https://salgsite.net/about

[b]Cmap concept mapping software: https://cmap.ihmc.us/

[c]IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

[d]University of California – Student Diversity Statistics.

https://diversity.ucr.edu/statistics/studentstat.html

5.10 References

[1] D. Cros, M. Maurin, R. Amouroux, M. Chastrette∗, J. Leber, M. Fayol, Eur. J. Sci. Educ., 8 (1986) 305–313.

[2] D.R. Mulford, W.R. Robinson, J. Chem. Educ., 79 (2002) 739.

[3] A.G. Harrison, D.F. Treagust, Sci. Educ., 80 (2018) 509–534.

[4] M.M. Cooper, L.M. Corley, S.M. Underwood, J. Res. Sci. Teach., 50 (2013) 699–721.

[5] J.D. Novak, Instr. Sci., 19 (1990) 29–52.

[6] D. Rickey, A.M. Stacy, J. Chem. Educ., 77 (2000) 915.

[7] E. Cook, E. Kennedy, S.Y. McGuire, J. Chem. Educ., 90 (2013) 961–967.

[8] A.H. Johnstone, J. Chem. Educ., 70 (1993) 701.

[9] D. Gabel, J. Chem. Educ., 76 (1999) 548.

[10] K.R. Galloway, S.L. Bretz, J. Chem. Educ., 92 (2015) 2019–2030.

[11] J.D. Novak, J. Chem. Educ., 61 (1984) 607.

[12] G. Nicoll, J.S. Francisco, M. Nakhleh, J. Chem. Educ., 78 (2001) 1111.

[13] J.S. Francisco, M.B. Nakhleh, S.C. Nurrenbern, M.L. Miller, J. Chem. Educ., 79 (2002) 248.

[14] S.A. Kennedy, J. Chem. Educ., 93 (2016) 645–649.

[15] M.-P. Chevron, Perspect. Sci., 2 (2014) 46–54.

[16] D.P. Ausubel, New York: Holt, Rinehart and Winston, Inc. (1968).

[17] J.C. Nesbit, O.O. Adesope, Rev. Educ. Res., 76 (2006) 413–448.

[18] N. Turan-Oluk, G. Ekmekci, Chem. Educ. Res. Pract., 19 (2018) 819–833.

[19] J.D. Novak, B.B. Gowin, Cambridge University Press, Cambridge. (1984).

[20] L. Ye, R. Oueini, S.E. Lewis, J. Chem. Educ., 92 (2015) 1807–1812.

[21] A. Regis, P.G. Albertazzi, E. Roletto, J. Chem. Educ., 73 (1996) 1084.

[22] P.G. Markow, R.A. Lonning, J. Res. Sci. Teach., 35 (1998) 1015–1029.

[23] M. Besterfield-Sacre, J. Gerchak, M.R. Lyons, L.J. Shuman, H. Wolfe, J. Eng. Educ., 93 (2013) 105–115.

[24] C.J. Luxford, S.L. Bretz, J. Chem. Educ., 91 (2014) 312–320.

[25] N.L. Burrows, S.R. Mooring, Chem. Educ. Res. Pract., 16 (2015) 53–66.

[26] A.J. Cañas, R. Carff, G. Hill, M. Carvalho, M. Arguedas, T.C. Eskridge, J. Lott, R. Carvajal, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005: pp. 205–219.

[27] J.D. Novak, A.J. Cañas, Florida Inst. Hum. Mach. Cogn., (2008).

[28] M.R. Blanchard, S.A. Southerland, J.W. Osborne, V.D. Sampson, L.A. Annetta, E.M. Granger, Sci. Educ., 94 (2010) 577–616.

[29] M.R. Mack, C. Hensen, J. Barbera, J. Chem. Educ., 96 (2019) 401–413.

[30] J. Burdo, L. O'Dwyer, Adv. Physiol. Educ., 39 (2015) 335–340.

[31] N.L. Leech, J.A. Gliner, G.A. Morgan, R.J. Harmon, J. Am. Acad. Child Adolesc. Psychiatry, 42 (2003) 738–740.

[32] J.F. Eichler, J. Peeples, Chem. Educ. Res. Pract., 17 (2016) 197–208.

[33] W. Widhiarso, H. Ravand, Rev. Psychol., 21 (2014) 111–121.

[34] E.J. Pedhazur, Thomson Learning, USA. (1997).

[35] R. Zwick, J.C. Sklar, Am. Educ. Res. J., 42 (2005) 439–464.

[36] S. Geiger, M.V. Santelices, University of California, Berkeley Center for Studies in Higher Education, (2007), https://escholarship.org/uc/item/7306z0zf.

[37] P. Vincent-Ruz, K. Binning, C.D. Schunn, J. Grabowski, Chem. Educ. Res. Pract., 19 (2018) 342–351.

[38] C.H. Middlecamp, T. Jordan, A.M. Shachter, K. Kashmanian Oates, S. Lottridge, J. Chem. Educ., 83 (2006) 1301.

[39] S.M.D. Elaine Seymour, Douglas J. Wiese, Anne-Barrie Hunter, Pap. Present. Natl. Meet. Am. Chem. Soc., San Franci (2000).

[40] N. Falchikov, D. Boud, Rev. Educ. Res., 59 (1989) 395–430.

[41] G.D. Kuh, (Bloomington, IN, Indiana University Center for Postsecondary Research), (2001).

[42] J.H. van Driel, W. Gräber, Springer Netherlands, Dordrecht, 2003: pp. 271–292.

[43] M.M. Cooper, H. Kouyoumdjian, S.M. Underwood, J. Chem. Educ., 93 (2016) 1703–1712.

[44] R.T. Jana Hackathorn, Kathryn Cornell, April Garczynski, Erin Solomon, Katheryn Blankmeyer, J. Scholarsh. Teach. Learn., 12 (2012) 78–87.

[45] T. Lindstrom, C. Middlecamp, J. Chem. Educ., 94 (2017) 1036–1042.

[46] S. Olejnik, J. Li, S. Supattathum, C.J. Huberty, J. Educ. Behav. Stat., 22 (1997) 389–406.

[47] A. Gilewski, E. Mallory, M. Sandoval, M. Litvak, L. Ye, Chem. Educ. Res. Pract., 20 (2019) 399–411.

Chapter 6: Concluding Remarks

Mass spectrometry has become an incredibly powerful tool for the characterization of peptides and proteins. With the development of new analytical tools, the information which can be obtained will only continue to grow. An appreciation of this can be obtained in Chapters 2 through 4, each of which focuses on their own aspect of information that can be ascertained. In Chapter 2, a new wavelength, 213 nm, was coupled with mass spectrometry. It was revealed that 213 nm led to both specific and nonspecific fragmentation. Systems which contain common labile bonds at 266 nm, show enhanced dissociation at 213 nm. Importantly, many native and non-native carbon-sulfur bonds were found to dissociate readily following 213 nm photoexcitation. These included C-S bonds adjacent to disulfide bonds, leading to the identification of a new signature triplet for improving accurate identification of disulfide bonded peptides. Future applications will include the examination of large proteins which contain numerous disulfide bonds. Photodissociation of many C-S bonds leads to the formation of an alaninyl radical which initiates beta dissociation at the radical site, leading to the formation of signature d-ions. These d-ions prove vital for the site-specific identification of phosphorylated residues. Further experiments will hopefully expand on the bond-specific dissociation observed herein and apply it to future biomolecular structure characterization.

Action-excitation energy transfer (action-EET) is a wonderful analytical tool which can provide distance specific information about the conformation of a peptide or protein in the gas phase. Previous action-EET systems have utilized intrinsic donors and disulfide bond acceptors for characterization of three-dimensional peptide structures. Chapter 3 focused on the identification of two new energy acceptors and their utility in action-EET experiments was probed. The first system, C-S bonds, was inspired due to the enhanced

C-S bond dissociation that was observed in the 213 nm UVPD experiments. Methionine contains two such energy acceptors, and while dissociation of these bonds was observed, the yields were often very low, prompting an investigation of selenomethionine. C-Se bonds proved to be quite useful as energy acceptors. Energy transfer from tyrosine/tryptophan to the C-Se bonds led to significant bond-specific dissociation. It was determined that distance was the primary factor contributing to energy transfer efficiency. Due to the dietary uptake of selenomethionine and its incorporation into peptides and proteins for X-ray crystallography studies, this action-EET system could see significant use in the future for gas phase protein structure studies.

Photoinitiated radical chemistry is often utilized for bond breaking. The work in Chapter 4 reveals that radical chemistry can also lead to gas phase bond formation. We know that 266 nm light will selectively break S-S and C-I bonds in the gas phase. In one such system, unreactive sulfur radicals are used, while in the other system, aromatic phenyl radicals are used, each of which leads to the formation of new gas phase bonds. Sulfur-centered radicals will react selectively with another nearby S-S bond. The formation of this new bond was shown to occur both intermolecularly and intramolecularly. The formation of these new bonds shows that caution should be taken in methods like ECD and UVPD where sulfur radicals are commonly generated, as disulfide scrambling can occur even with a single disulfide homolytically cleaved. In the second system, diradical recombination is shown to be useful for identifying sites of noncovalent interaction between a crown ether and either a peptide or protein. Future studies with the crown ether system may focus on applications to different proteins. Diradical covalent coupling of the crown ether to peptides may prove useful for identifying sites of solvation as a protein is transferred from liquid to gas phase.

While often not viewed in the same light, teaching and research often go together. Chapter 5 focuses on a re-examination of concept maps into first- and third-quarter chemistry courses. Results of the study revealed that students who were able to produce well-developed concept maps performed slightly better on a series of concept inventory questions. It also appeared that students in the treatment group generally reported high self-reported gains in a series of survey questions than those in the control. These results show that it is show that the introduction of concept maps can improve student learning gains and the proper implementation into the classroom is key for their success. Future work should examine the relationship between conceptual learning and concept mapping and if further instruction on how to make advanced concept maps can further improve student learning gains.

The work in this dissertation was two-fold, the development of novel spectroscopy coupled mass spectrometry experiments and the implementation of concept maps into a classroom setting. Research is not, and should not, be viewed as work in a laboratory setting only. While they are often viewed as unrelated, recent evidence has shown that growth of students in the classroom will lead to the production of better scientists. Continued development of analytical techniques for protein characterization and student learning gains go together in helping solve future scientific inquiries.