UCLA
UCLA Electronic Theses and Dissertations

Title
Towards the Development of New Bioconjugation Strategies for Peptides and Proteins

Permalink
https://escholarship.org/uc/item/3g1579pz

Author
Hakim Moully, Elamar

Publication Date
2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Towards the Development of New Bioconjugation Strategies for Peptides and Proteins

A thesis submitted in partial satisfaction

of the requirements for the degree of Master of Science

in Chemistry

by

Elamar Hakim Moully

2016

© Copyright by

Elamar Hakim Moully

2016

ABSTRACT OF THE THESIS

Towards the Development of

New Bioconjugation Strategies for Peptides and Proteins

by

Elamar Hakim Moully

Master of Science in Chemistry

University of California, Los Angeles, 2016

Professor Joseph Ambrose Loo, Chair

Incorporation of inorganic-based moieties into biological molecules can engender new properties and applications. In this thesis, two projects are described that aim to introduce new inorganic motifs into peptides. The first project describes attempts to incorporate palladium ion into the active site of the naturally-occurring metalloprotein Sso10b2 (from the archaeon *Sulfolobus solfataricus*). Since the native protein contains zinc, it cannot be obtained using recombinant technologies, since subsequent metal exchange requires harsh conditions. Using standard solid-phase peptide synthesis, we synthesized peptide fragments that would directly interact with Pd(II), and studied the metallation process via liquid chromatography-mass spectrometry and circular dichroism. The second project describes a new cysteine bioconjugation technique using organometallic boron-cluster reagents. A platinum-boryl complex was used to selectively label

the cysteine residue of an unprotected peptide with a carborane, via reductive elimination mechanism. These carborane labels will ultimately serve as tags for Raman microscopy, allowing us to track our target peptides within cells.

The thesis of Elamar Hakim Moully is approved.

Alexander Michael Spokoyny

Amy Catherine Rowat

Joseph Ambrose Loo, Committee Chair

University of California, Los Angeles

2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

SPPS	Solid-phase peptide synthesis
LC-MS	Liquid chromatography-mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
TIC	Total ion chromatogram
CD	Circular dichroism
NCL	Native chemical ligation
DMF	Dimethylformamide
DCM	Dichloromethane
EtOH	Ethanol
ACN	Acetonitrile
TFA	Trifluoroacetic acid
NMM	N-Methylmorpholine
EDT	1,2-Ethanedithiol
TIPS	Triisopropylsilane
Pd(II)	Palladum(II)
$Pd(OAc)_2$	Palladium acetate
HCL	Hydrochloric acid
NaOH	Sodium hydroxide
Pt	Platinum

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor Professor Alexander M. Spokoyny, for his constant support, encouragement, and guidance over the past two years. I am incredibly grateful for all his mentorship and advice. I would also like to thank the entire Spokoyny group, for the positive and fostering environment they built in our lab. I would especially like to thank Dr. Liban Saleh, Parker Beatty, Austin C. Rist, and Dr. Jonathan C. Axtell, whose help made this thesis possible. Furthermore, I would like to thank Azin Saebi, for her overwhelming moral support.

Finally, thank you to my family, for everything they have done and continue to do for me. Thank you for your endless motivation and inspiration.

Chapter 1: Design of Artificial Metalloenzyme Featuring Pd(II) Metal

1. 1. Introduction

The design of artificial metalloenzymes is a growing field, as these scaffolds posses unique properties that can be utilized in a variety of applications. The ability to mimic the enzymic abilities of these macromolecules could open the door to new, previously inaccessible mechanisms. Specifically, the incorporation of unnatural metals, such as palladium, could change the catalytic and electronic properties of the scaffolds. The goal of this research project was to develop a metalloenzyme scaffold containing Pd(II) metal ion in its active site, via direct coordination. Palladium possesses advanced catalytic abilities that could expand the functional scope of the scaffold.

Designing a functional, artificial metalloenzyme requires an ability to both recognize and incorporate all relevant aspects of the native scaffolds. Several approaches have been developed over the years to address these challenges, and these can be divided into two main subgroups: *de novo* design vs. native scaffold design.[1] *De novo* design refers to the assembly of a polypeptide that is not associated with any known, natural protein; additionally, these structures are also designed with a deliberate 3-dimensional structure in mind.[2] Because a large number of factors must be addressed when designing a *de novo* polypeptide (such as thermodynamic considerations and folding variability),[2] computational programs are increasingly used in the design process. The design of α-helical bundles containing metal-binding sites has been heavily studied and successfully used to elucidate the mechanisms of certain metalloproteins.[3, 4] Simple mutations in the design of a peptide can lead to changes in stability and specificity of the scaffolds, thus demonstrating the power of *de novo* design.[5]

The second approach utilizes native scaffolds as models for artificial metalloenzymes. This approach is more frequently used than *de novo* design because it requires fewer considerations, yet it still involves fundamental understanding of the correlation between structure and function.[6] Reetz et al. from the Max Planck Institute for Coal Research were the first to demonstrate the addition of a metal-binding active site into a protein scaffold that did not already contain an active site. Choosing copper(II) as their desired metal, the group incorporated a His/His/Asp sequence to form the coordinating site in imidazole glycerol phosphate synthase from *Thermotoga maritime*.[7] Another technique is the use of scaffolds already containing an active site, although the number of possibilities is more limited. For example, the biotin-avidin technology, as devised by Wilson and Whitesides, is one of the most commonly used methods to insert a non-natural group into the avidin protein.[8] Avidin is a natural protein that has high affinity for biotin, a co-factor for carboxylase enzymes; the interaction between avidin (as well as streptavidin and neutravidin) and biotin is one of the strongest known protein-ligand interactions. Wilson and Whitesides demonstrated that by attaching a certain moiety to biotin, the moiety will become anchored to the active site of avidin via the avidin-biotin interaction. Since the development of this technology, which is considered a non-covalent incorporation, numerous groups have utilized this interaction to anchor non-natural metal moieties inside avidin proteins.[9,10] Covalent incorporation of metal co-factors using functional groups on the scaffold and/or ligand, as well as direct metal coordination between amino acid residues and the meal ion are two other methods also employed in this field.[11] The ability to incorporate unnatural metals into enzymes will engender new catalytic properties in these scaffolds. These artificial metalloenzymes could be used for enantioselective catalysis that is not possible with small-molecule catalysts, since they offer additional selectivity.

Although all these possible approaches exist for the incorporation of unnatural metals, the field is still in its early stages, and so far only a few studies demonstrate the successful addition of palladium metal ions. Ward and coworkers used the avidin-biotin technology to incorporate different palladium ligands into the scaffold; these artificial complexes were then employed as catalysts in an asymmetric allylic alkylation to yield either *R* or *S* conformations in high *ee* yield.[12] However, the dependence on biotin-avidin restricts the widespread use of the catalysts, since they require avidin-type scaffolds. Palomo et al. designed a *p*-nitrophenyl-phosphonate palladium pincer that was covalently incorporated into immobilized lipase enzymes, which are commercially available.[13] The metalloenzyme was tested as a catalyst in a simple Heck cross-coupling reaction between iodobenzene and ethyl acrylate. Although the metalloenzyme complex showed high yield of the product (>99%), it could only be used for two cycles, retaining 70% activity. This simple Heck reaction has been shown to work with high turnover numbers ($>10^6$), suggesting that this specific metalloenzyme complex is extremely limited in its use.[14]

In this thesis, I will describe my attempts in designing a new Pd-based metalloenzyme that relies on direct coordination of the metal within the active site. This design would eliminate the dependence on linker-ligands that require unique design and synthesis for different scaffolds, thus simplifying the process and producing a more widespread approach. Our lab worked with collaborators in the Alexandrova lab at UCLA, who used a computational model to identify a protein that should—according to thermodynamic computations—bind to palladium(II). The protein, Sso10b2, is a DNA-binding protein from the archaeon *Sulfolobus solfataricus*.[15] This specific archaeon is thermoacidophilic, which suggests the enzyme scaffold should be able to withstand high temperatures and acidic environments while retaining its stable structure. The computer model indicates that palladium will coordinate to three residues within the active site:

to Lys11 and Asp13 in a monodentate fashion, and to Asp21 in a bidentate fashion (total coordination number of 4) (**Figure 1-1**).

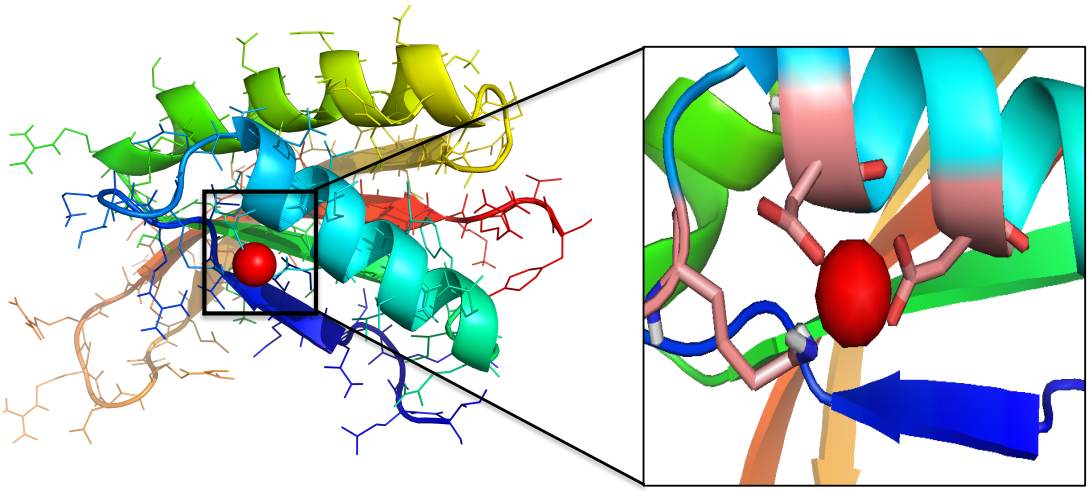

Figure 1-1. Computer model of Sso10b2, with Pd(II) ion incorporated within the active site.

However, the protein naturally binds to zinc(II) metal ion, and therefore cannot be obtained using recombinant methods, since subsequent metal exchange conditions are harsh. Solid-phase peptide synthesis (SPPS) is a method that allows for the step-wise addition of amino acids to build a peptide chain on a polystyrene resin.[16] Hood et al. established a rapid Fmoc-protected SPPS approach that uses 2-(6-Chloro-1H benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) as the coupling reagent.[17] We will utilize this approach to synthesize peptide segments of Sso10b2. Sso10b2 is comprised of 88 amino acids, but SPPS can only support chains of about 30 amino acids. Consequently, native chemical ligation (NCL) will be employed to construct the entire scaffold. NCL relies on a transthioesterification reaction between a cysteine residue on the N-terminus and a thioester on the C-terminus.

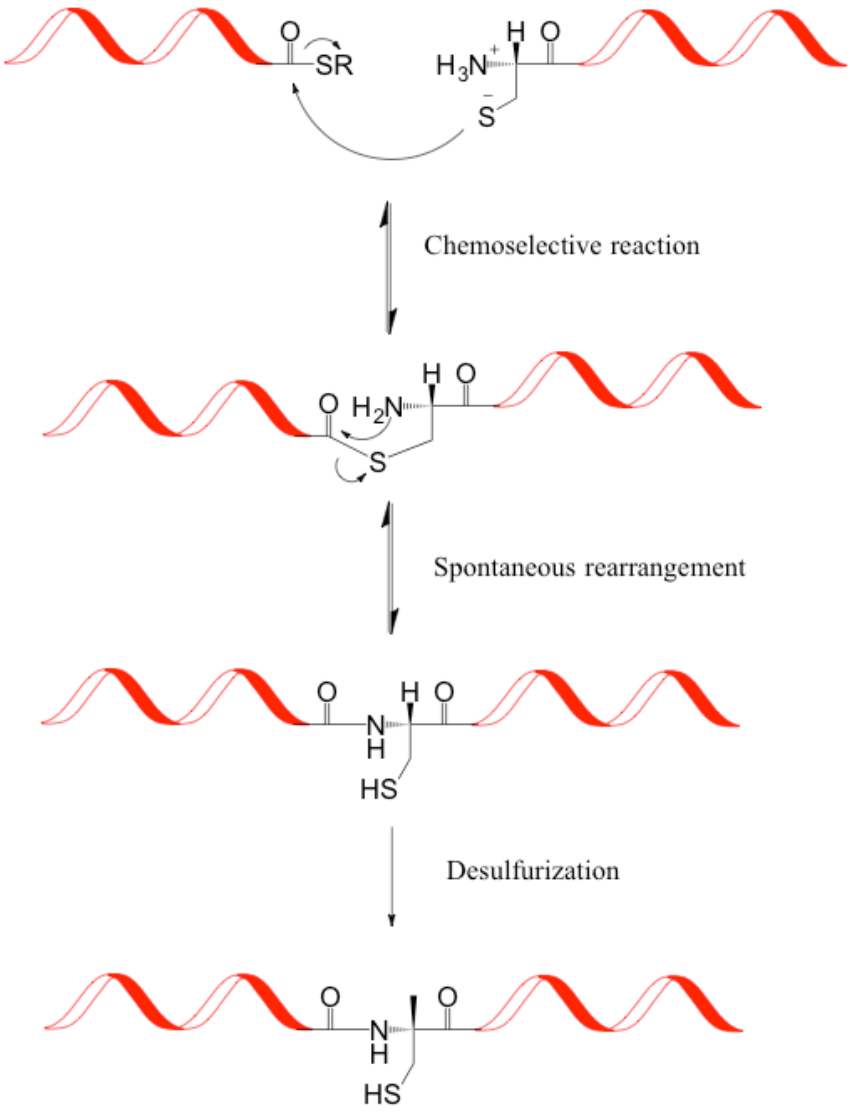

Figure 1-2. General scheme for NCL mechanism, resulting in a native peptide bond. Adapted from Yan and Dawson.[18]

Following transthioesterification, the intermediate rearranges to form a native amide bond between the two peptide segments (**Figure 1-2**).[18] In the case of our scaffold, no cysteines are present in the native sequence. However, Yan and Dawson described an approach to synthesize proteins without cysteine residues using desulfurization.[18] In this approach, cysteine undergoes desulfurization to yield an alanine residue within the sequence (**Figure 1-2**). Unfortunately, only

one alanine is present in Sso10b2, and its location is not convenient for a ligation site since it is in the middle of an α-helix. Therefore, two glycines (Gly29 and Gly59) can be mutated to alanines (Ala29 and Ala59) in the final sequence. We do not expect these mutations to generate any structural or functional changes in the scaffold, since both glycine and alanines are small residues.

1. 2. Methods

1. 2. 1. Peptide Synthesis

Peptides were synthesized on a 0.4 mmol scale using Fmoc-Solid Phase Peptide Synthesis (SPPS) (**Figure 1-3**). Synthesis was carried out in a 25 ml vessel connected to a vacuum. Amino acid residues were activated using 0.4M N-Methylmorpholine (NMM) in dimethylformamide (DMF) and coupled to either rink amide resin (**P1**, **P2**, and **P3**) or 2-chlorotrityl hydrazine resin (**P4** and **P5**) as follow: 1) a 5 minute coupling of the amino acid with 2.0 mmol equivalents of 0.4M HCTU in DMF; 2) Five washes of DMF; 3) 3 minute deprotection with 20% (v/v) piperidine in DMF; and 4) Five washes of DMF. After all amino acids were coupled and deprotected, the resin was washed with DMF, ethanol (EtOH), and dichloromethane (DCM). The peptide was cleaved from the resin using a solution of 18 mL trifluoroacetic acid (TFA), 0.8 mL triisopropylsilane (TIPS), 0.6 mL water, and 0.6 mL 1,2-Ethanedithiol (EDT) in 2.5 hours. The solvent was then evaporated under argon gas for approximately one to two hours. The peptide was triturated three times (50 mL total) with cold diethyl ether and was allowed to dry for 20 minutes. The obtained residue was dissolved first in H_2O with 0.1% TFA, and then acetonitrile (ACN) with 0.1% TFA was added until all solids were completely dissolved. The solution was lyophilized for two days.

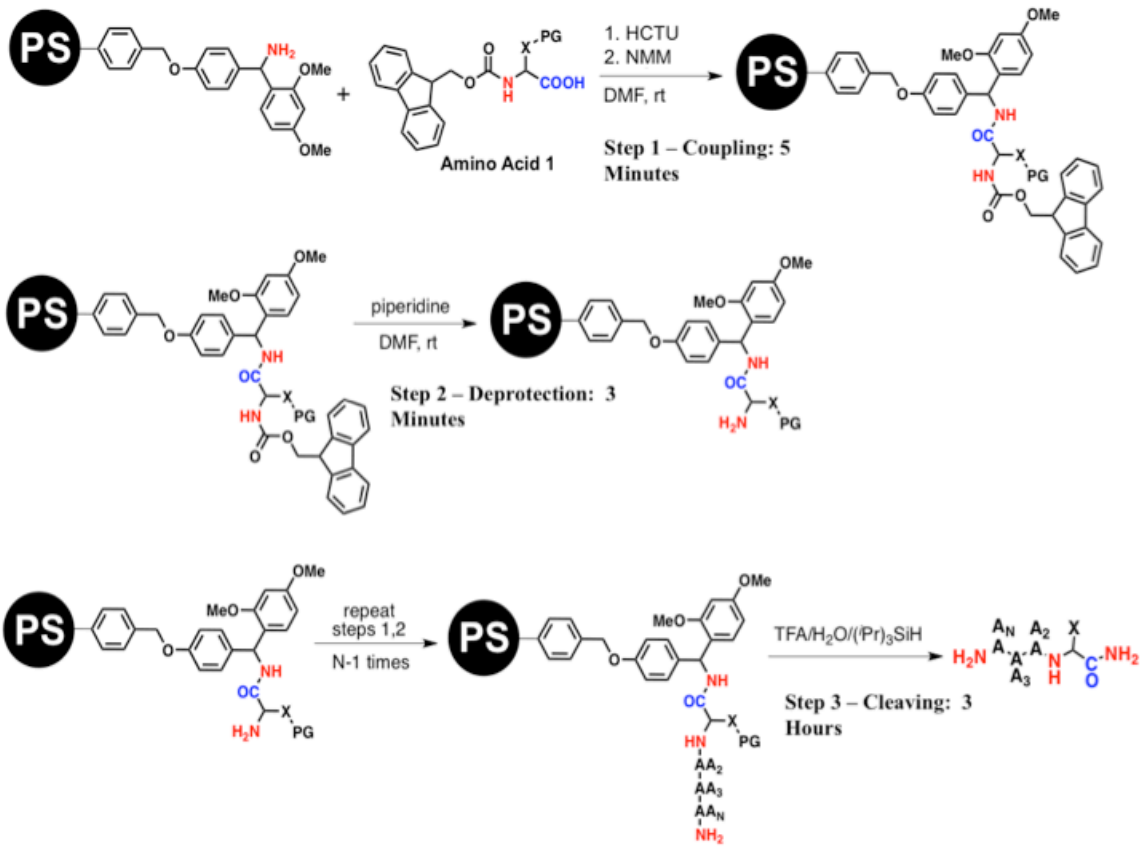

Figure 1-3. Solid-phase peptide synthesis (SPPS) general scheme. PS=Polystyrene resin.

Peptide	Sequence	Calculated Mass (Da)	Observed Mass (Da)
P1	NH_2-T**K**NVEDHVLDV-$CONH_2$	1266.6568	1266.3060
P2	NH_2-T**K(alloc)**NVEDHVLDV-$CONH_2$	1350.6779	1350.2560
P3	NH_2-KLNEIVVRKTKNVEDHVLDVIVLFNQG-$CONH_2$	3117.7663	3117.8780
P4	NH_2-TEKLNEIVVRKTKNVEDHVLDVIVLFNQ-$CONHNH_2$	3305.8460	3305.6771
P5	NH_2-CIDEVILKGTGREISKAVDVYNSLKDRLGD-$CONHNH_2$	3319.7558	3119.5279

Table 1-1. List of peptides synthesized for this project, with calculated and observed masses.

***Note on Calculated/Observed Mass:** Calculated mass was determined using the Analysis function in ChemDraw Professional 15.0. The calculated mass reported here is the exact mass of the peptide sequence, as reported by ChemDraw. The observed mass was calculated using the highest-intensity $[M+2H]^{2+}$ peak (z=2). The observed peak was multiplied by 2, and the mass of

2 hydrogen atoms was subtracted from the product to give the observed mass. These methods were followed throughout for consistency purposes. If the $[M+2H]^{2+}$ peak was not observed, $[M+3H]^{3+}$ (or the next lowest charge peak) was used.

1. 2. 2. Peptide Purification

Post lyophilization, the crude peptide was dissolved in 20 mL of 95% H_2O with 0.1% TFA and 5% acetonitrile with 0.1% TFA. The solution was filtered using a 0.22 μm filter and then purified using a Prep-HPLC (XBridge™ Prep C18 5μm OBD™, 19x50 mm). Solvents: 0.1% TFA in H_2O (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). Depending on the sequence (**Table 1-1**), each peptide was purified using a different gradient. The HPLC fractions containing the peptide were collected and tested for purity on matrix-assistant laser desorption/ionization mass-spectrometry (MALDI-MS) and liquid chromatography-mass spectrometry (LC-MS). The pure fractions were lyophilized and stored at -4 °C.

1. 2. 3. LC-MS Analysis

LC-MS experiments were acquired using a Waters Acquity UPLC connected to a Waters LCT-Premier XE Time of Flight Instrument operated by MassLynx 4.1 software. The solvents used for analysis were 0.3% formic acid in H_2O (Solvent A) and 0.2% formic acid in acetonitrile (Solvent B). All solvents are LC-MS/MS grade and purchased from Fisher Scientific. The following methods were used:

Method 1-A LC conditions: BEH C18 1.7 μm column (2.1 x 50 mm, Waters), gradient: 0 – 5 min, 5 – 50% B; 5 – 5.10 min, 50 – 95% B, 5.10 – 7.50 min, 95% B, 7.50 – 7.51 min, 95 – 5% B, 7.51 – 10 min, 5%, flow rate: 0.4 mL/min. MS conditions: positive mode, 70 – 2000 *m/z*

Method 1-B LC conditions: BEH C18 1.7 μm column (2.1 x 50 mm, Waters), gradient: 0 – 5 min, 5 – 50% B; 5 – 5.10 min, 50 – 95% B, 5.10 – 7.50 min, 95% B, 7.50 – 7.51 min, 95 – 5% B, 7.51 – 10 min, 5%, flow rate: 0.4 mL/min. MS conditions: positive mode, 70 – 2000 *m/z*, 1.5 minutes solvent delay

1. 2. 4. Pd(II) Incorporation

Following peptide synthesis, a procedure was developed to incorporate Pd(II) into the active site. A 100 mM phosphate buffer solution (Na_2HPO_4) was prepared, both at pH 7 and pH 11, to serve as the solvent. We chose those two pH values as starting points: pH 7, since it is similar to physiological pH of 7.4; and pH 11, since the pKa of lysine is around 10, and the computational model predicted that the lysine within the active site has to be deprotonated for the coordination to occur. Hydrochloric acid and/or sodium hydroxide were added to the buffer until a pH of 7 or 11 was reached. A stock peptide solution of the peptide was then made at each pH by dissolving 1 mg of the pure peptide in 1 mL of the phosphate buffer solution. To prepare a Pd(II) stock solution, palladium acetate $Pd(OAc)_2$ was used. Over the course of the year, we prepared numerous 1:1 peptide:Pd(II) samples, at both pH values for comparisons. All samples were prepared at room temperature and then allowed to sit for approximately 2.5 days. The duration was determined based on various experiments and time points taken.

General procedure. The peptide was dissolved in the phosphate buffer solution (100 mM) of desired pH to give a 1 mg/mL stock solution. This solution was stored in the freezer at -4 °C. A stock solution of $Pd(OAc)_2$ also in phosphate buffer of desired pH was prepared to be 1:1 equivalents with the peptide solution (taking into account TFA counter ions when calculating mass of peptide). $Pd(OAc)_2$ did not readily dissolve in the phosphate buffer, and so it was

allowed to sit at room temperature for approximately 24 hours until no solids were observed, and the solution was a bright yellow/orange color. This solution was stored on the bench at room temperature. After the preparation of both solutions, 50 μL of each solution were mixed, yielding a 1:1 peptide:Pd(II) solution mixture at the desired pH. The solution was allowed to sit at room temperature for approximately 2.5 days, at which point LC-MS was used to confirm incorporation of Pd(II) into the active site.

1. 2. 5. Circular dichroism (CD)

CD was performed on JASCO J-715 Circular Dichroism spectrophotometer. CD was used to observe any changes in the secondary structure of the peptide scaffold upon addition of the palladium ion, since different secondary structures generate different elliptically polarized light. The computational model provided by the Alexandrova lab suggests that upon incorporation of Pd(II), the active side adopts a smaller helix conformation compared to the apo-form. Therefore, upon coordination with Pd(II), we expected a decrease in the mean residue ellipticty (MRE).

General Procedure. A solution of the desired peptide was prepared in 1:1 H_2O:ACN at a concentration of 0.02 M, taking into account TFA counter ions. A solution of 1:1 peptide:$Pd(OAc)_2$ was prepared in 1:1 H_2O:ACN so to have the same peptide concentration as the solution of the apo-peptide. A 1 mm cuvette was used. A blank of the solvent (1:1 H_2O:ACN) was run on the instrument prior to the experimental run, and the background was subtracted from the other acquired spectra. After acquiring a spectrum containing a peptide, the cuvette was cleaned in the following way: Wash 3X with Triton®, 2X with DI water, 2X with EtOH, 2X with DI water, and 2X with doubly DI water.

1. 2. 6. Native Chemical Ligation

The NCL synthesis approach was adapted from Pentelute et al.[19] We decided to assemble the entire complex from three segments, ^{1}Thr – ^{28}Gln, ^{29}Cys – ^{58}Asp, and ^{59}Cys – ^{88}Tyr. As mentioned in the introduction, ^{29}Cys and ^{59}Cys will undergo desulfurization to yield two alanines, resulting in the mutations G29A and G59A in the overall sequence. In the scope of my time in the lab, we only synthesized the first two segments, ^{1}Thr – ^{28}Gln (**P4**) and ^{29}Cys – ^{58}Asp (**P5**) **(Table 1-1)**. These were synthesized on 2-chlorotrityl hydrazine resin, so as to have a hydrazine on the C-terminus. The hydrazine terminals ultimately undergo oxidation during the reaction (using sodium nitrite) to form thioesters, following which ligation will occur.[19]

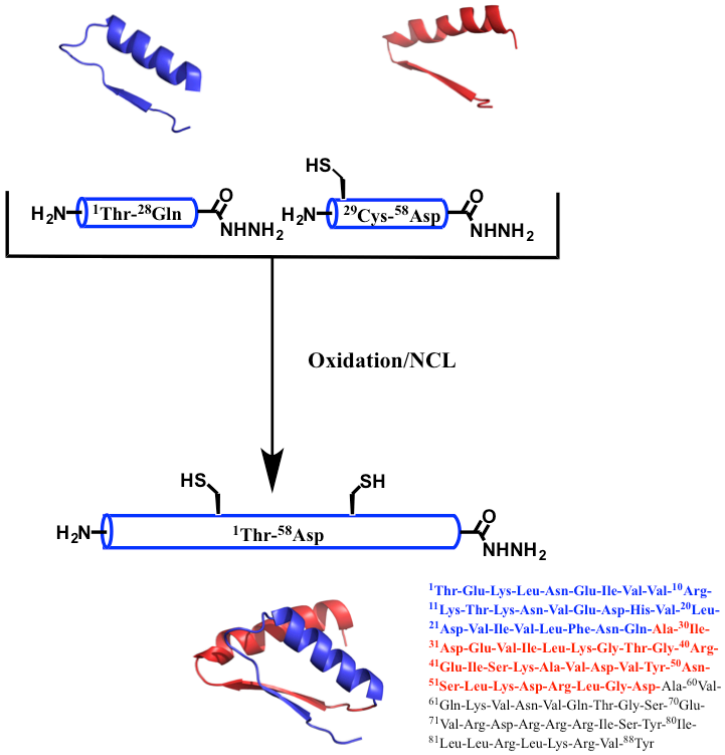

Figure 1-4. General scheme for NCL, linking P4 and P5.

Procedure for P4 and P5 Ligation. 0.58 μmol (1.9 mg) of **P4** was dissolved in 200 μL of ligation buffer (0.15 M Na_2HPO_4 buffer with 6 M Guanidinium•HCl) at pH of 3.1. The solution was stirred over an ice-salt bath at -12 °C. To the cold solution, 20 μL of 200 mM sodium nitrite ($NaNO_2$) was added dropwise in neat H_2O to oxidize hydrazine; after an additional 15 minutes, another 10 μL was added. The solution was allowed to sit for 30 minutes, maintaining -12 °C. To the mixture, 600 μL of 67 mM solution of 4-mercaptophenylacetic acid (MPAA) in ligation buffer at pH 7 was added dropwise to quench oxidation. The solution was kept at -12 °C for an additional 15 minutes, then allowed to warm up to room temperature. 0.63 μmol (1.3 mg) of **P5** was added to the mixture, and the pH was adjusted to 6.9 using 2 M NaOH to initiate NCL (**Figure 1-4**). LC-MS was used to monitor the ligation process at different time points.

1. 3. Results

1. 3. 1. LC-MS Analysis of Pd(II) Incorporation in P1

Our first round of experiments was performed with peptide **P1**, since it was relatively quick to synthesize and contained the active site. LC-MS data was acquired using Method 1-A (**Figure 1-5).**

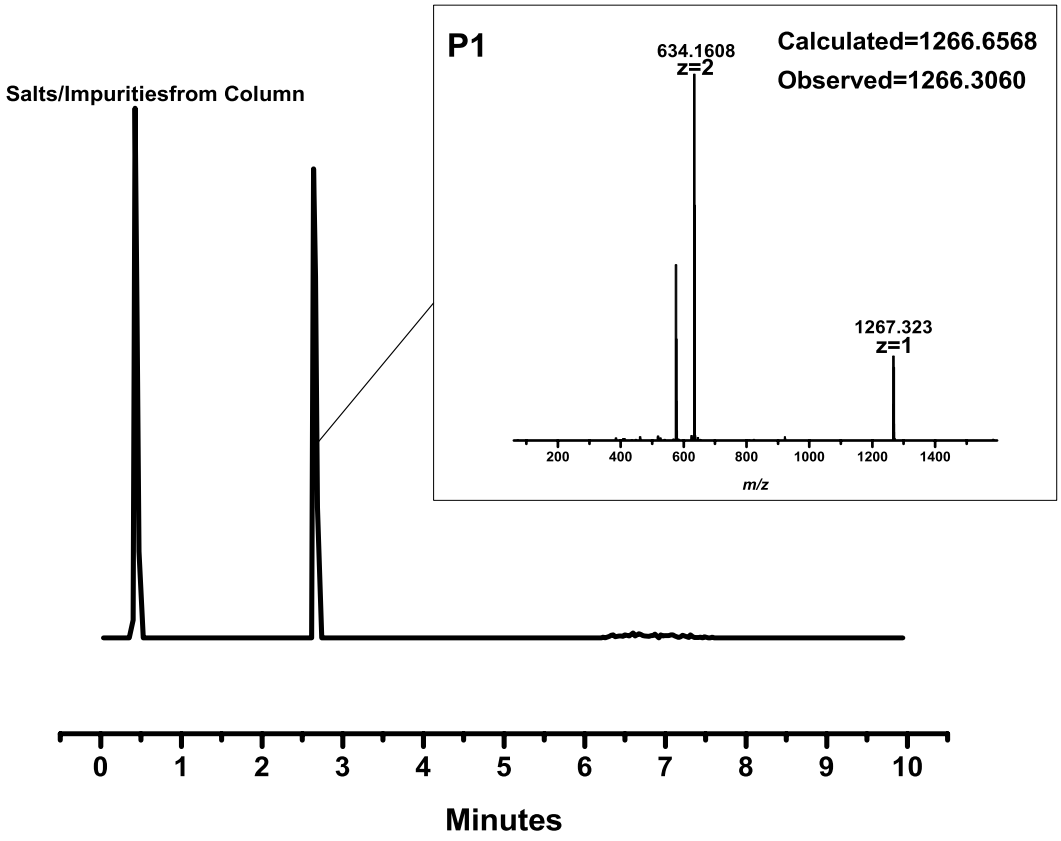

Figure 1-5. LC-MS TIC trace for P1, including mass spectrum. Obtained using Method 1-A.

As discussed earlier, based on the computational model we expected Pd(II) to coordinate to three residues within the active site: to Lys11 and Asp13 in a monodentate fashion, and to Asp21 in a bidentate fashion (total coordination number of 4) (**Figure 1-6**). We refer to the expected product as **P1-Pd(II)**.

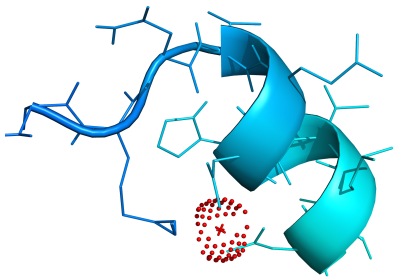

Figure 1-6. Computer model of P1 (blue) coordinated with Pd(II) (red). Total coordination number of 4.

***Note:** The Pd(II)-incorporation experiments were performed both at pH 7 and 11, but the LC-MS instrument was undergoing maintenance when running the pH 11 experiments. These experiments need to be repeated. Only the results from pH 7 are reported here.

➔ **Pd(II) Coordination with P1 at pH 7**

Following the procedure outlined in the methods section, we mixed $Pd(OAc)_2$ and **P1** at pH 7 and obtained the following results after 2.5 days. LC-MS data was acquired using Method 1-A (**Figure 1-7**).

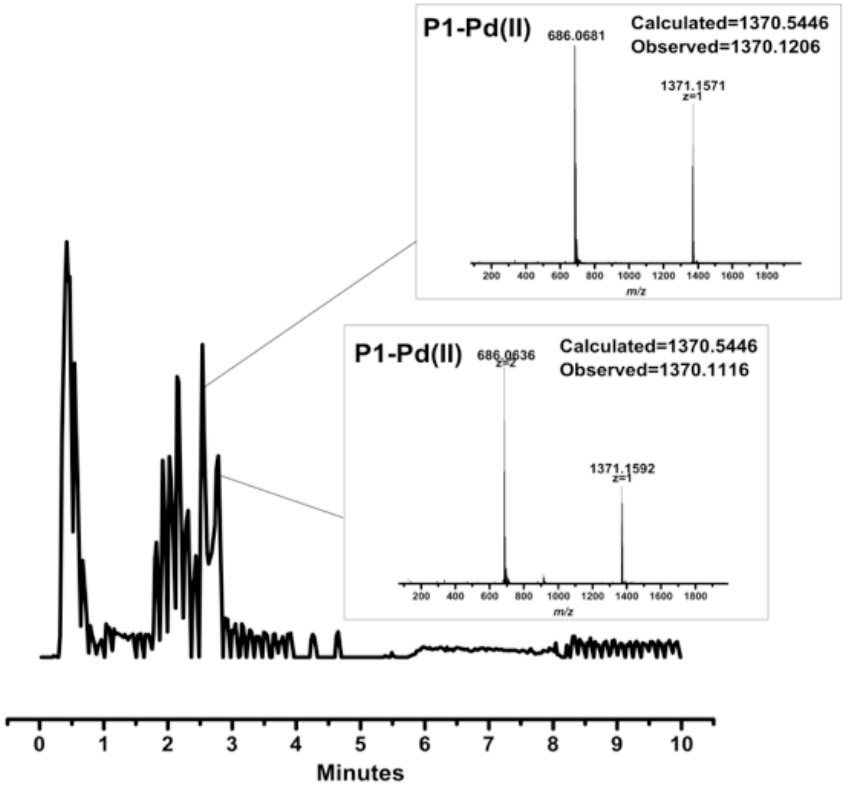

Figure 1-7. LC-MS TIC trace for P1-Pd(II), obtained following the general procedure described above. Multiple peaks correspond to the mass of P1-Pd(II), indicating isomers. Two representative mass spectra are displayed in the figure, but additional peaks contain the same mass spectra. Obtained using Method 1-A.

The results suggest that the Pd(II) ion is binding to our short peptide **P1** at multiple sites because we observe different species eluting at different times that have the same mass spectra, corresponding to the mass of **P1-Pd(II)**; two mass spectra are showcased in **Figure 1-7** as a representation, but other smaller peaks within the LC trace have the same mass spectra as well. The isotope pattern of the peaks for **P1-Pd** is much different than when compared to **P1**, which also suggests that Pd(II) is somehow incorporating into the peptide (**Figure 1-8**).

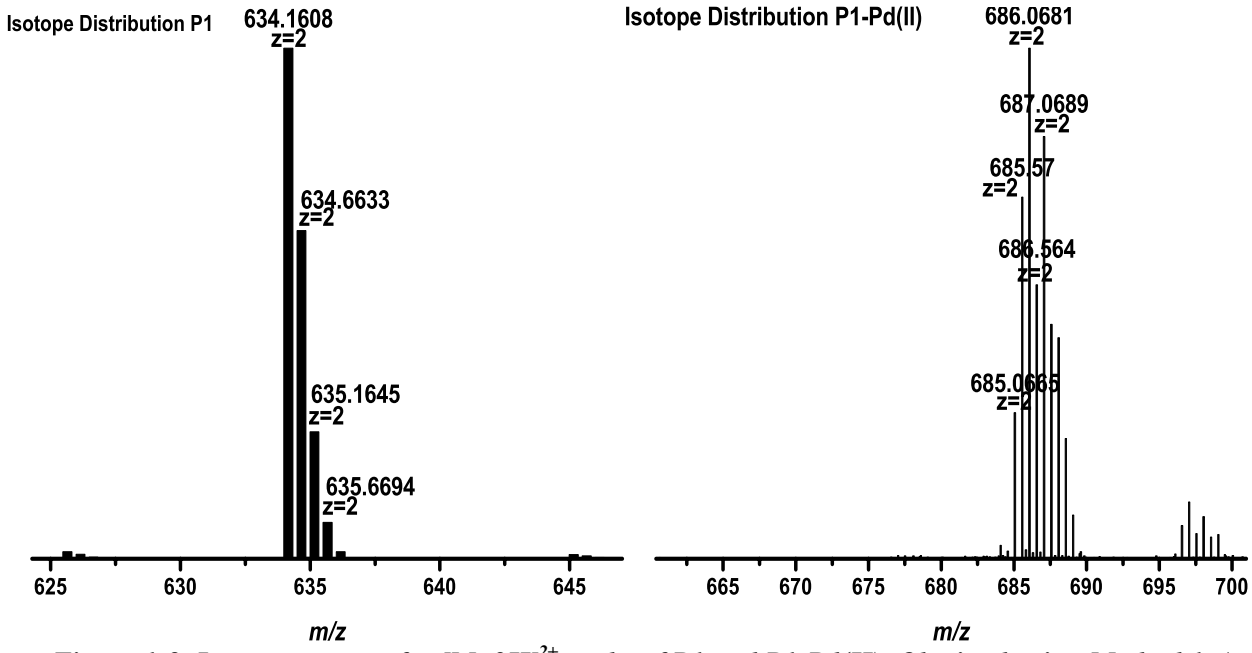

Figure 1-8. Isotope patterns for $[M+2H]^{2+}$ peaks of P1 and P1-Pd(II). Obtained using Method 1-A.

It is important to note that since the LC-MS takes place under acidic conditions, we cannot probe the native state of the scaffold. These conditions may interfere with the coordination of Pd(II) to the amino acid residues.

➔ **Effects of Excess Pd(II)**

One experiment was conducted where an excess of Pd(II) was mixed with the **P1** solution. The same general procedure was followed, but the Pd(II) stock solution prepared was with a higher concentration, that was approximately 5 equivalents to the peptide concentration. LC-MS data was acquired using Method 1-A (**Figure 1-9**).

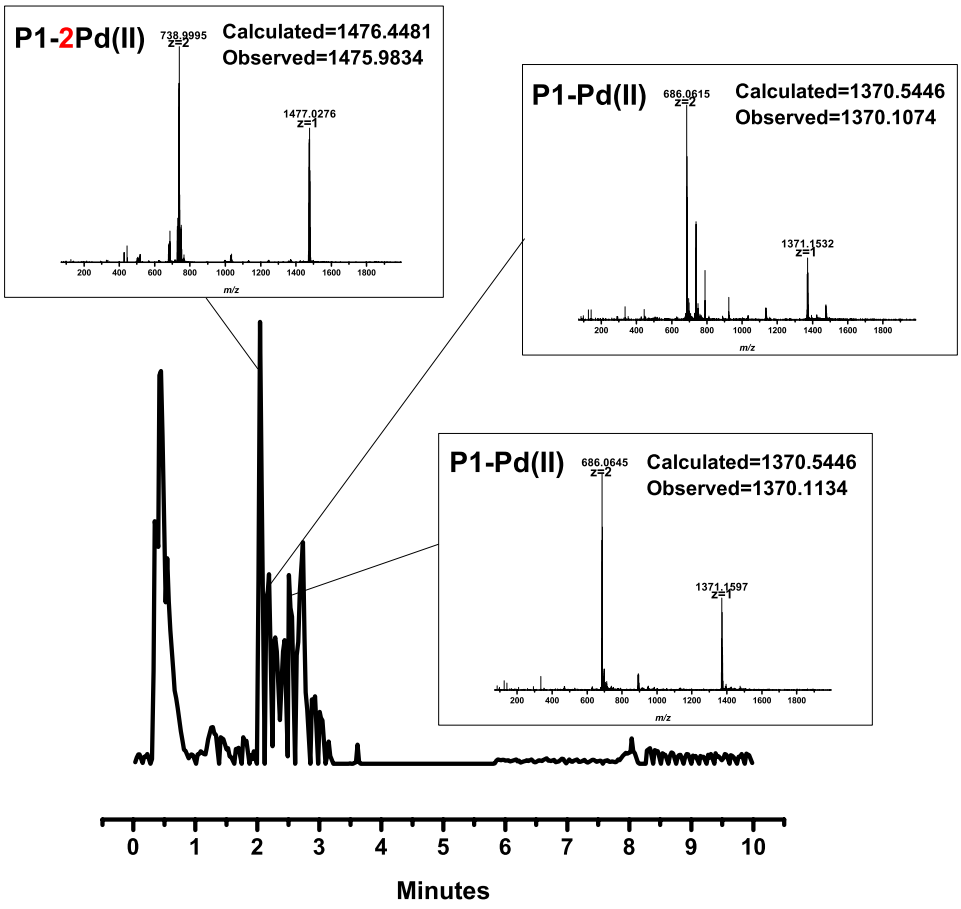

Figure 1-9. LC-MS TIC trace for reaction carried out with excess Pd(II) source. Main peak corresponds to P1 plus 2 Pd(II) ions. Other peaks also present with mass corresponding to P1-Pd(II). Obtained using Method 1-A.

The largest peak in the LC trace corresponds to the peptide plus two Pd(II) ions based on its *m/z*. And, we again observe several smaller peaks corresponding to the peptide plus one Pd(II) ion (multiple isomers).

➔ **Control Reaction with Protected Lysine in Active Site**

To investigate whether the three residues (Lys 11, Asp 13, and Asp 21) are indeed involved in the coordination of Pd(II), we conducted one experiment with a peptide containing alloc-protected Lysine11 (**P2**). The protected lysine in **P2** was the only difference compared to **P1** (**Table 1.1**). The same procedure was followed as discussed above. LC-MS data was acquired using Method 1-A (**Figure 1-10**).

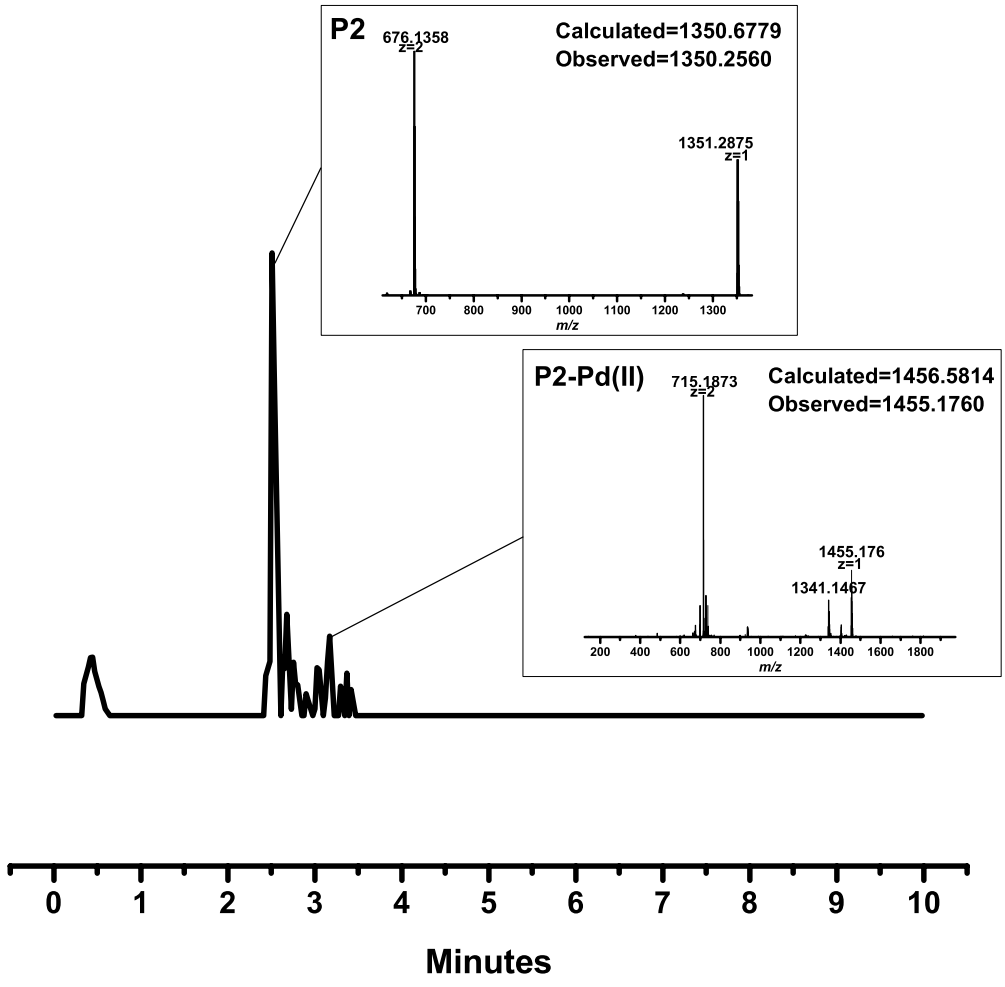

Figure 1-10. LC-MS TIC trace for the control reaction carried out with P2, which contains a protected lysine. Main peak corresponds to P2 alone, with no Pd(II) binding. Other smaller peaks are also present with mass corresponding to P2-Pd(II). Obtained using Method 1-A.

The largest peak in the LC trace corresponds to **P2** alone (without binding to Pd(II)). Some smaller peaks are present that correspond to the peptide plus Pd(II), but these are not as significant, suggesting that Lys11 is highly involved in the coordination of the metal ion.

1. 3. 2. LC-MS Analysis of Pd(II) Incorporation in P3

We then decided to investigate the coordination of Pd(II) with a longer peptide segment **P3**, still containing the active site (**Figure 1-11**).

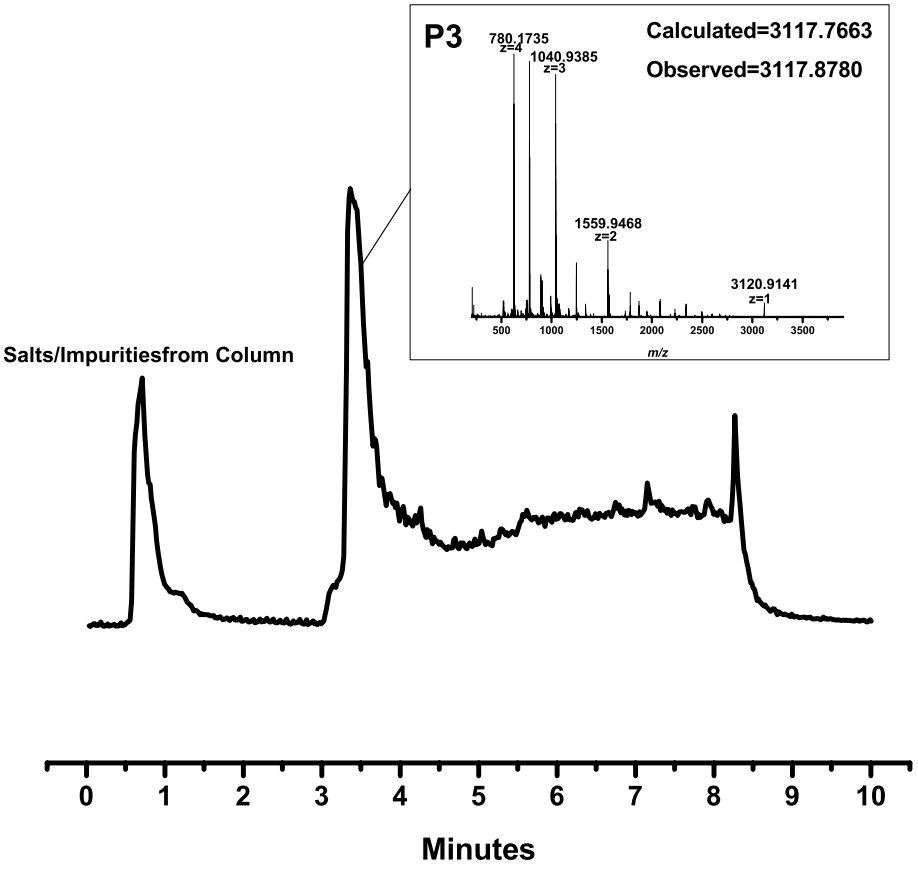

Figure 1-11. LC-MS TIC trace of P3, including mass spectrum. Obtained using Method 1-A.

In natural enzymes and proteins, the environment surrounding the active site is also significant in directing binding. Therefore, we hoped that the increased rigidity of the longer peptide would

better direct the Pd(II) to bind only at the active site, thus eliminating the isomers we observed with **P1**. We refer to the expected product as **P3-Pd(II) (Figure 1-12**).

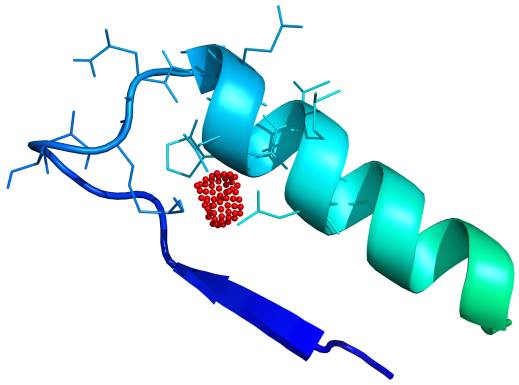

Figure 1-12. Computer model of P3 coordinated with Pd(II) (red). Total coordination number of 4.

➔ **<u>Pd(II) Coordination with P3 at pH 7</u>**

We repeated the procedure with a longer peptide segment **P3**, still containing the active site. We hoped that the longer segment would allow for more selective Pd(II) coordination. LC-MS data was acquired using Method 1-A (**Figure 1-13**).

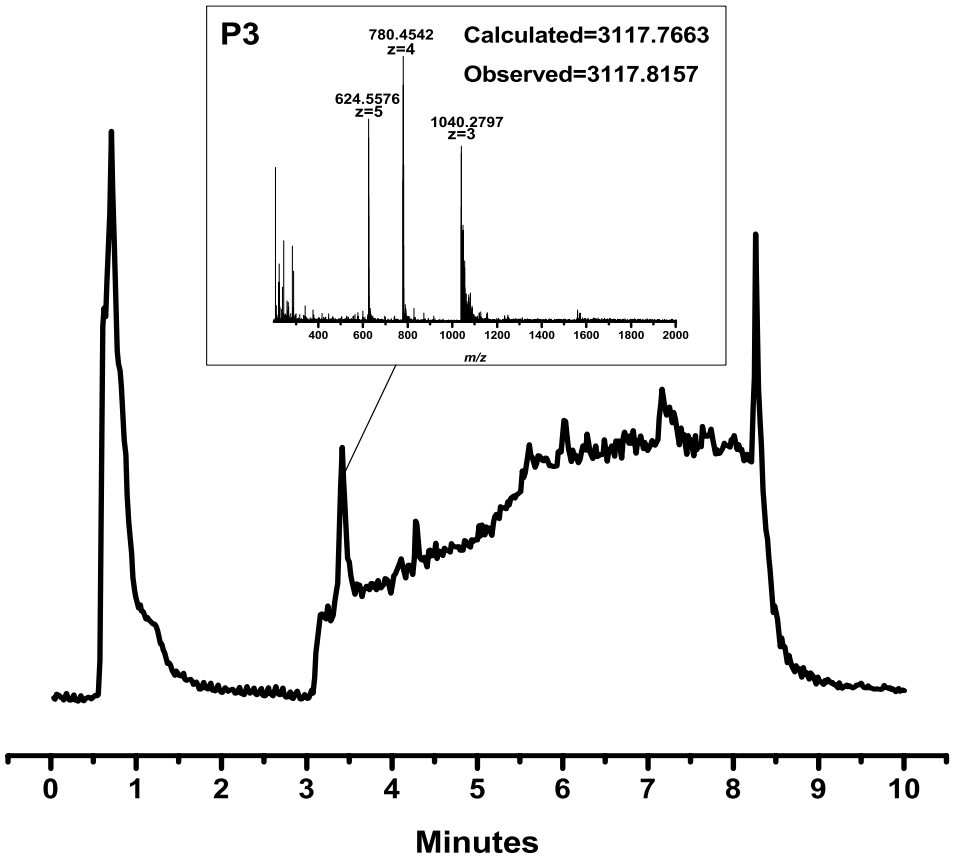

Figure 1-13. LC-MS TIC trace for reaction performed with P3 and Pd(II). No peaks observed corresponding to the expected product P3-Pd(II), only a peak for P3 alone. Obtained using Method 1-A.

No change in mass was observed in the mass spectra, which means Pd(II) was not detected in the sample containing 1:1 **P3**:Pd(II). Only the peptide alone (apo-form) was detected in the sample. One possible explanation is that the longer peptide is less soluble in our buffer. Some precipitation of the peptide was apparent when it was added to the buffer. Furthermore, upon addition of $Pd(OAc)_2$, further precipitation occured. Upon centrifugation, a clear, sizeable pellet with a slightly yellow hue (from palladium) was present at the bottom of the tube. The presence of the large pellet and its yellow color suggests that our longer peptide was precipitating out, together with much of the Pd(II). The experiment needs to be repeated under optimized conditions to ensure full solubility.

1. 3. 3. Circular Dichroism

We obtained the CD spectrum for **P3** alone as well as a for a **P3**:$Pd(OAc)_2$ mixture that we hoped would yield **P3-Pd(II) (Figure 1-14)**. Calculations were performed as described in Price et al.[20] to obtain the MRE. In a CD spectrum, an ideal α-helix has two minima at 222 and 208 nm, and a MRE at 222 nm of about -30,000 deg cm^2 $dmol^{-1}$.[20] When more than one secondary structure is present in a scaffold, the spectrum is a combination of both.

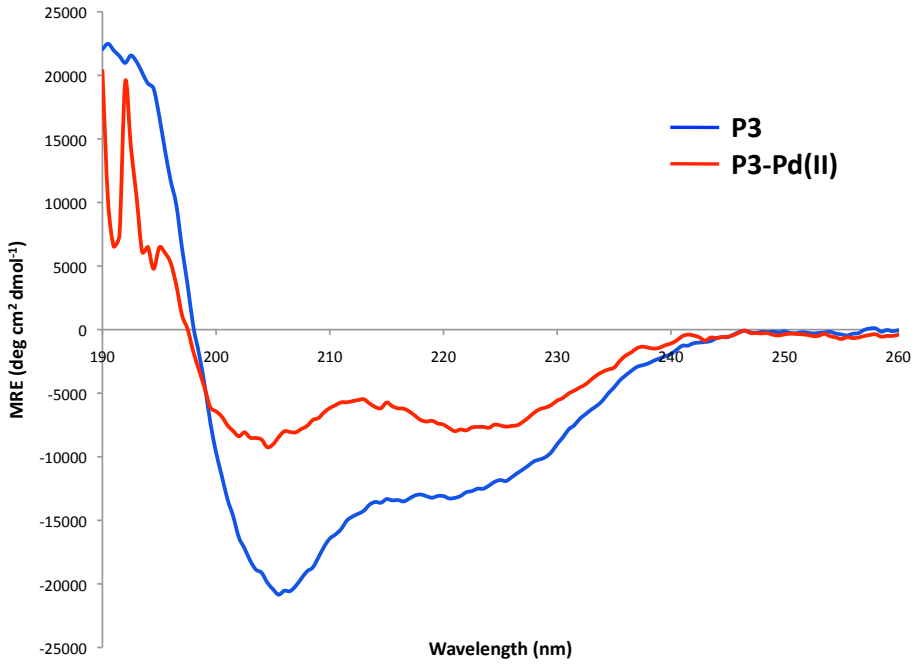

Figure 1-14. Circular dichroism spectrum for P3 and P3-Pd(II). The two minima points are representative of an α-helix. As the computation model predicted, incorporation of Pd(II) decreases the helicity.

The two minima that are representative of an α-helix are observed in both samples, since **P3** is long enough to have a secondary structure containing an α-helix. As predicted by the computational model, we see a decrease in MRE for the **P3**:Pd(II) sample, suggesting that the

Pd(II) is coordinating within the active site, even though we observe no change in the LC-MS mass spectra.

1. 3. 4. Native Chemical Ligation

P4 and **P5** were synthesized using SPPS on hydrazine resin and isolated individually. LC-MS data was acquired using Method 1-A (**Figure 1-15**).

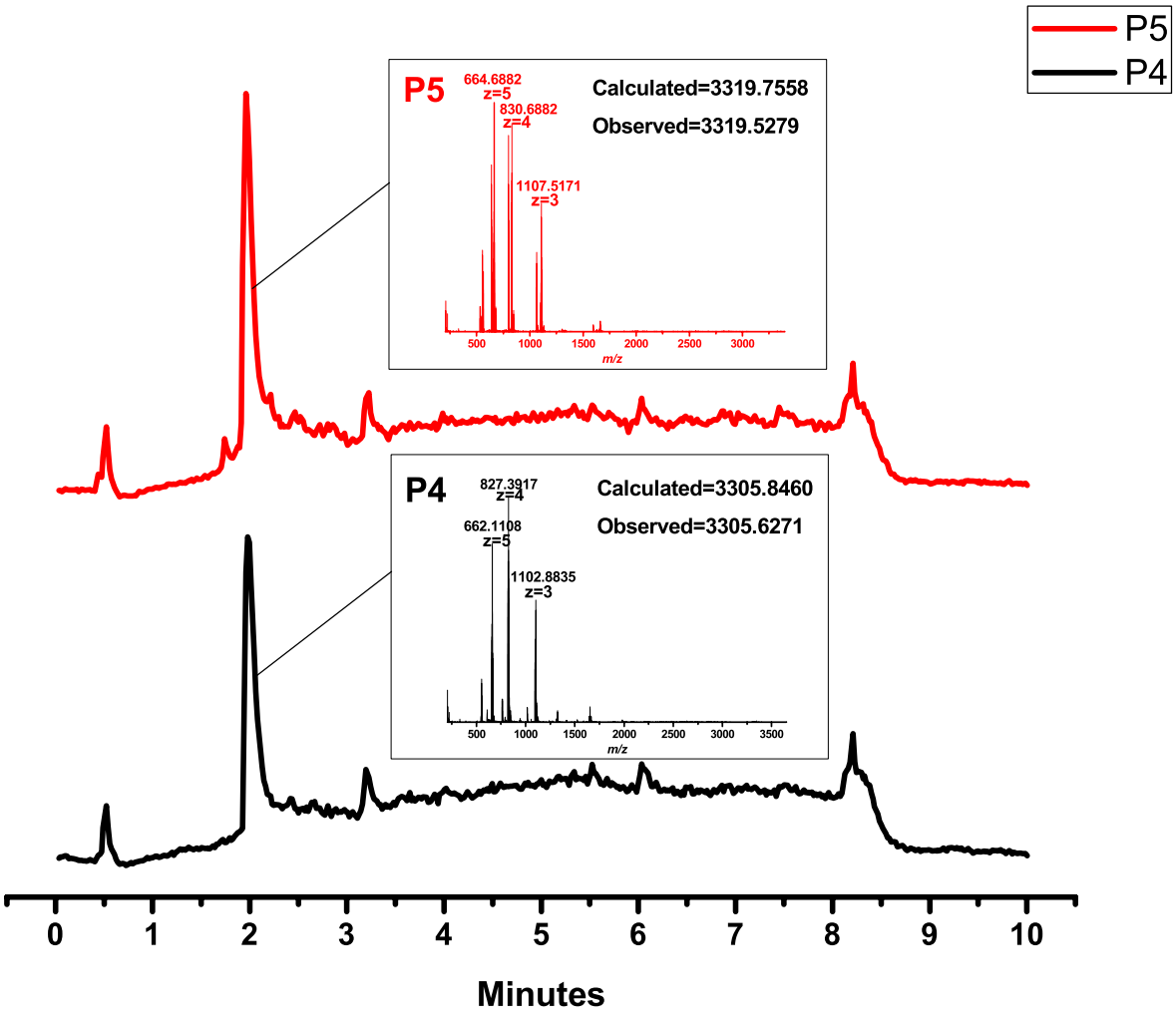

Figure 1-15. LC-MS TIC traces of P4 and P5, including mass spectra. Obtained using Method 1-A.

The NCL reaction was performed as described in the methods section. Aliquots were taken at two different time points, 2 hours and 9 hours, and analyzed using LC-MS to determine whether the reaction was complete. LC-MS Method 1-B was utilized, including a solvent delay to ensure that no salts reached the mass detector.

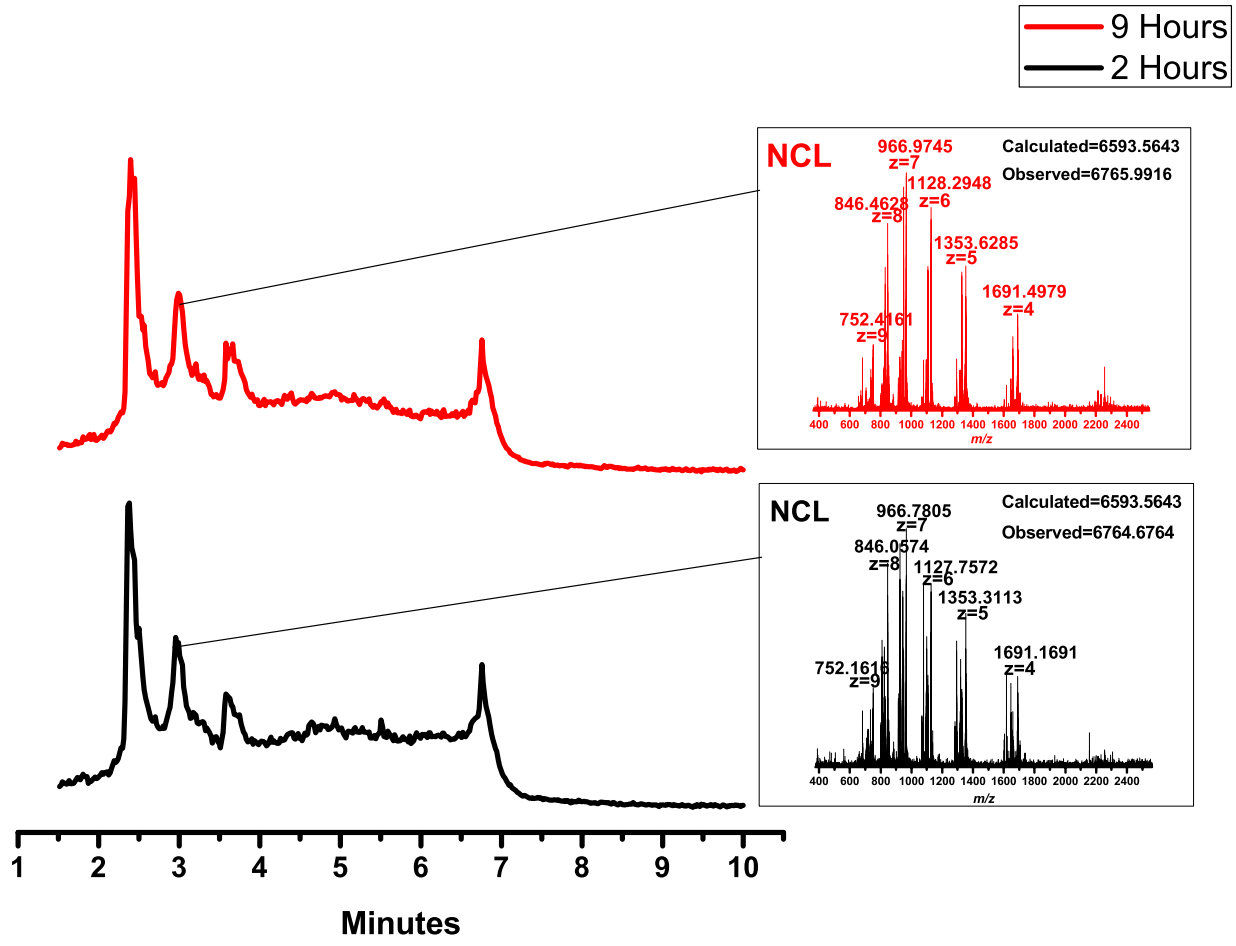

Figure 1-16. NCL results after 2 and 9 hours. LC-MS TIC trace. Peak at retention time of 2.99 minutes displays a charge envelope, indicating that ligation occurred. Obtained using Method 1-B.

There was no change between 2 and 9 hours, as is observed in **Figure 1-15**. In both spectra, a new peak at 2.99 minutes emerged with a mass spectrum showcasing a charge envelope, suggesting that ligation did occur. The observed mass was approximately 200 Da higher than the expected mass for the ligation product; however, there is a cluster of peaks for each charge,

making the observed mass harder to calculate accurately. We need to use a deconvolution software to obtain an accurate mass. In the scope of my research, I did not have time to thoroughly analyze the ligation product or repeat the experiment. However, these results show promising potential for full synthesis of the protein scaffold in the future since NCL is possible.

1. 4. Discussion and Future Steps

By studying the interaction between our scaffold and Pd(II) metal ion, we will gain insight into how a natural macromolecule scaffold behaves when it is introduced to an unnatural cofactor. The preliminary work presented here suggests that Pd(II) metal ion is incorporated into the active site of the synthesized scaffold, though not selectively. These results are promising, since they showcase that it is possible to incorporate an unnatural metal into a biological scaffold. As we progress with this project, we need to better understand where Pd(II) is binding, so we can engender specific regiochemical control associated with the metallation. We need to develop conditions where we can investigate the native binding conditions. We can also perform fragmentation experiments on mass spectrometry to help us answer these questions. Moreover, based on the CD spectra, we observe that the incorporation of Pd(II) induces a conformational change in the active site (smaller helix conformation), as predicted by the computational model. After we optimize the conditions for NCL, it will become vital to investigate if other conformational changes take place upon incorporation of the metal ion with the entire scaffold, and how those changes affect the overall function of the complex. Ultimately, after evaluating the properties of our artificial metalloenzyme, we want to evaluate its potential as an enantioselective catalyst.

1. 5. References

1. Lu, Y., Yeung, N., Sieracki, N., and Marshall, N. M. *Nature*, **2009**, 460, 855–862.

2. DeGrado, W.F. and Summa, C. M. *Annu. Rev. Biochem*., **1999**, 68, 779–819.

3. Choma, C, T., Lear, J. D., Nelson, M. J., Dutton, P. L., Robertson, D. E., and DeGrado W. F. *J. Am. Chem. Soc*. **1994**, 116, 856–865.

4. Touw, D. S., Nordman, C. E., Stuckey, J. A., and Pecoraro, V. L. *Proc. Natl. Acad. Sci*., **2007**, 104 (29), 1969–11974.

5. Litowski, J. R. and Hodges, R. S. *J. Biol. Chem.,* **2002**, 277, 37272–37279.

6. Drienovsku, I. and Roelfes, G. *Isr. J. Chem.* **2015**, 55, 21–31.

7. Podtetenieff, J., Taglieber, A., Bill, E., Reijerse, E. J., and Reetz, M. T. *Angew. Chem.,* **2010**, 122, 5277–5281.

8. Wilson, M. E. and Whitesides, G. M. *J. Am. Chem. Soc.,* **1978**, 100**,** 306–307.

9. Collot, J., Gradinaru, J., Humbert, N., Skander, M., Zocchi, A., and Ward, T. R. *J. Am. Chem. Soc.,* **2003**, 125, 9030–9031.

10. Ward, T. R. *Chemistry,* **2005**, 11 (13), 3798–804.

11. Lewis, J. C. *ACS Catal.* **2013**, 3, 2954−2975.

12. Pierron, J., Malan, C., Creus, M., Gradinaru, J., Hafner, I., Ivanova, A., Sardo, A., and Ward, T. R. *Angew. Chem. Int. Ed.,* **2008**, 47, 701–705.

13. Filice, M., Romero, O., Aires, A., Guisan, J. M., Rumbero, A., and Palomo, J. M. *Adv. Synth. Catal.,* **2015**, 357, 2687–2696.

14. Farina, V. *Adv. Synth. Catal.*, **2004**, 346, 1553–1582.

15. Chou, C.-C.; Lin, T.-W.; Chen, C.-Y.; and Wang, A. H.-J. *J. Bacteriol*., **2003**, 185 (14), 4066–4073.

16. Merrifield, R. B. *J. Am. Chem. Soc.,* **1963**, 85 (14), 2149–2154.

17. Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; and Park, J. H. *J. Pept. Sci.,* **2008**, 14, 97–101.

18. Yan, L. Z. and Dawson, P. E. *J. Am. Chem. Soc.*, **2001**, 123 (4), 526–533.

19. Mong, S. K., Vinogradov, A. A., Simon, M. D., and Pentelute, B. L. *ChemBioChem,* **2014**, 15, 721–733.

20. Kelly, S. M., Jess, T. J., and Price, N. C. *Biochimica et Biophysica Acta*, **2005**, 1751 (2), 119–139.

Chapter 2: Cysteine Labeling of Unprotected Peptides with Organometallic Boron-Cluster Reagents

2. 1. Introduction

Fluorescence techniques are the most widely used methods to tag and track target molecules and proteins within cells. Despite the popularity and great utility of these methods, fluorescence has two fundamental inherent limitations: (1) fluorescent tags are often not photostable, and tend to photobleach over time, therefore limiting the experiment time; and (2) shining UV light at the fluorescent tags can result in the formation of reductive oxygen species (ROS), which can then react and interfere with other components within the cell. Raman imaging is an alternative technique that circumvents these limitations, allowing us to visualize molecules based on their vibrational modes. Because spontaneous Raman scattering generates a weak signal, two Raman imaging techniques have been developed to enhance the signal: Stimulated Raman Scattering (SRS) and Coherent Anti-Stokes Raman Scattering (CARS). SRS utilizes two laser beams (pump and Stokes) that focus on the sample. When the energy difference between the lasers matches the vibrational frequency of a given bond within the sample, it results in a non-linear interaction between pump and Stokes photons. The non-linear interaction then promotes the chemical bonds into an excited, vibrational state. The transition also allows for an energy transfer from the pump to the Stokes beam, resulting in higher intensity of the Stokes beam.[1] CARS also utilizes a pump and Stokes laser, but the pump beam has a adjustable frequency. Similar to SRS, when the energy difference between the lasers matches the vibrational frequency of a chemical bond, the chemical bonds will become excited. High-energy pump photons will then interact with the excited chemical bonds, resulting in the emission of Anti-Stokes photons. The photons give a coherent, high-intensity signal; however, due to the high frequency of the Anti-Stokes photons, CARS has background vibrational noise even at non-

resonant frequencies.[1] X. Sunney Xie at Harvard University and Wei Min at Columbia University are leading the field of Raman microscopy for biological applications. Xie et al. pioneered the development of label-free imaging using SRS, which relies on the inherent vibrations of biological molecules at different frequencies.[2–4] Min et al. developed alkyne tags and isotope tags for SRS imaging, expanding the scope of the technique. [5, 6]

The typical Raman spectrum of a cell contains a silent region between approximately ~2200 and 2800 cm^{-1}, since no natural elements within the cell vibrate at this frequency (**Figure 2-1**). If a molecule containing frequencies within this range could be introduced to an element within the cell, the utility of Raman imaging in biological applications could be substantially improved. Typically, B-H bond stretches occur at a frequency of approximately ~2500 cm^{-1}, falling right within the silent region. While this would provide the ideal modification to a molecule within the cell, for Raman imaging purposes multiple B-H stretches would be required to allow for rapid imaging, as well as any modifications retaining the cell permeability of the original molecule targeted for modification. In this regard, a class of compounds known as carboranes, which are boron-rich, 3-D aromatic analogues of benzene, could be ideal (**Figure 2-1**). The boron-rich nature of carboranes provides many B-H stretches in one relatively compact molecule. In addition, carboranes possess a number of other advantages: (1) their small size, which avoids interference with protein folding and function; (2) their hydrophobic nature, which allows for cell permeability;[7] and (3) their bioorthogonality, which ensures the tag does not interfere or react with components within the cell.

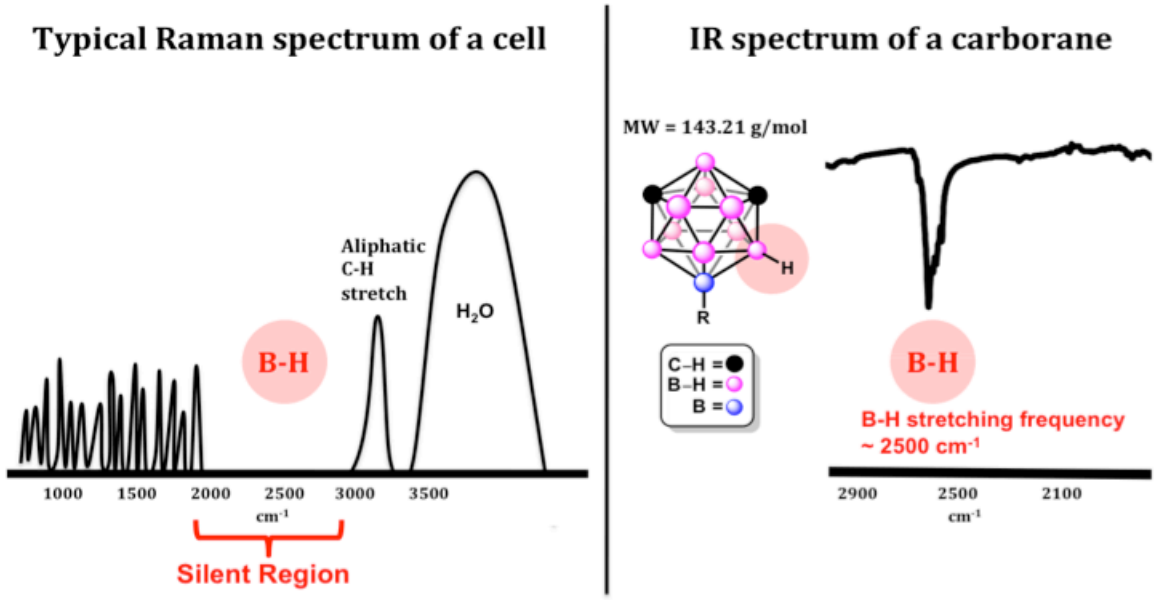

Figure 2-1. Representation of the typical Raman spectrum of a cell, highlighting the silent region between ~2200 and ~2800 cm^{-1}. The B-H stretch in a carborane molecule has a frequency of ~2500 cm^{-1}.

Recently, Pentelute and Buchwald reported the arylation of unprotected peptides via cysteine functionalizing using an organometallic palladium(II) reagent.[8] The reaction proceeded quickly (under five minutes) in mild, aqueous conditions via reductive elimination**.** A similar organometallic reagent containing carborane could allow for functionalization in the same vein. With that in mind, we used the platinum boryl organometallic complex [$PtCl(PPh_3)_2$(9-B-*m*-$C_2B_{10}H_{11}$)] (***m*-1**) that was previously reported by our lab.[9] We had shown that the platinum boryl complex can facilitate B–S bond formation in a facile manner.[9] A mild base deprotonates the thiol, allowing for a salt metathesis reaction to occur and generating a S–Pt bond containing intermediate. Reductive elimination then takes place, generating an S–B bond (**Figure 2-2**).

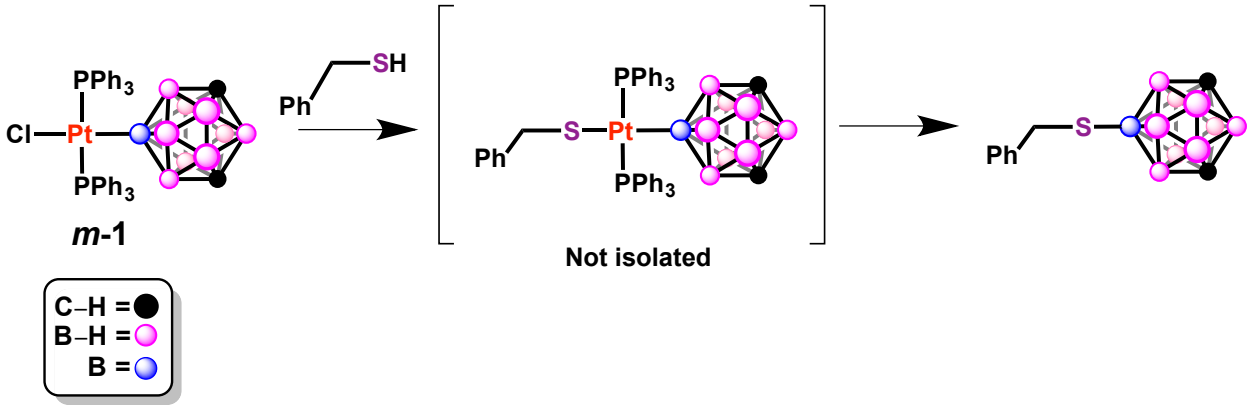

Figure 2-2. Reductive elimination performed with benzyl mercaptan as the thiol, previously reported by our lab.[9] Reaction conditions: K_3PO_4, C_6H_6, RT.

We hypothesized that we can use the same chemistry highlighted in **Figure 2-2** with biological molecules containing cysteine, which has a free thiol (**Figure 2-3**).

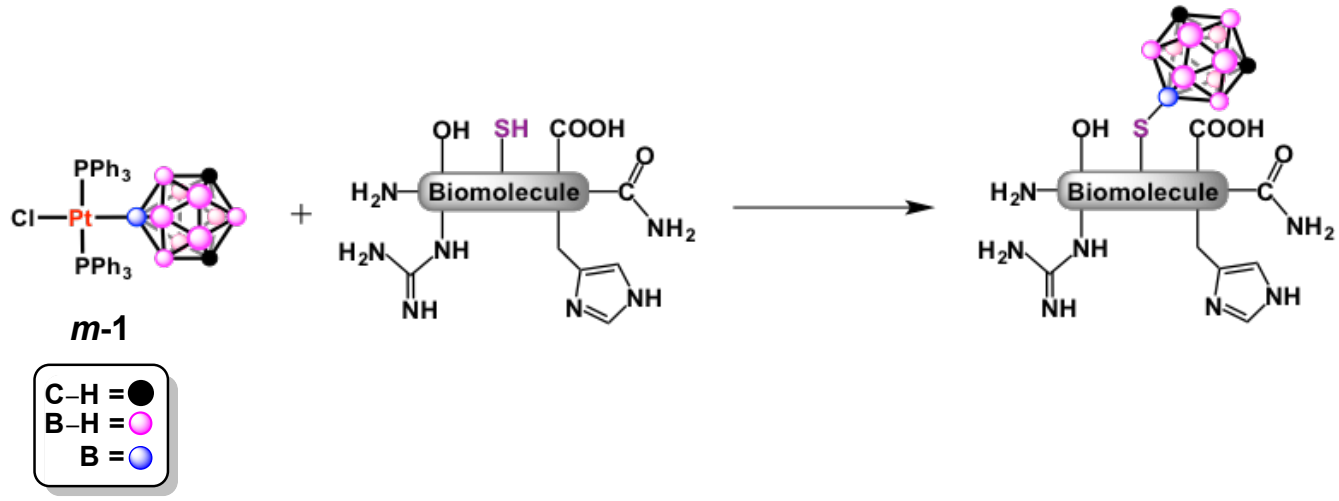

Figure 2-3. General scheme for our proposed carborane tagging of biological molecules via reductive elimination.

The strategy developed in this work allows for facile bioconjugation of peptides with boron-rich clusters. To the best of our knowledge, this is the first reported boron-cysteine bond (**Figure 2-4 A**). Currently, other strategies towards appending boron-rich clusters onto biological molecules are carried out via a CH vertex, and often require more complicated

synthetic strategies. Incorporation of carboranylalanine—an analogue of phenylalanine containing *o*-carborane in place of the phenyl group—is the common approach to introduce carboranes into amino acid sequences (**Figure 2-4 B**).[7] Carboranylalanine was designed as a Boron Neutron Capture Therapy (BNCT) agent, but has also shown enhanced properties over phenylalanine due to its hydrophilic nature.[7, 10] Other carborane analogues also exist, most commonly of unnatural amino acids.[7] However, this approach is limited in its scope, since it requires total synthesis of the sequence; thus, it cannot be easily applied to proteins and other large biomolecules. Moreover, its incorporation may alter the properties of the biomolecule, which makes it unsuitable to study the native functions.

Another bioconjugation strategy involves the functionalization of the carboranes with thiols, and having these react with maleimide-functionalized peptides or proteins (**Figure 2-4 C**).[11] However, this technique requires an additional step of maleimide bioconjugation to the peptide, introducing additional complexities such as selectivity. Hawthorne et al. developed a way to append carboranes on antibodies via an azo-coupling reaction, without interfering with the function of the antibodies (**Figure 2-4 D**).[12] Starting with a 1-phenyl-1,2-dicarba-*closo*-dodecaborane, this material was converted to a diazonium ion, and then reacted with an IgG antibody.[12]

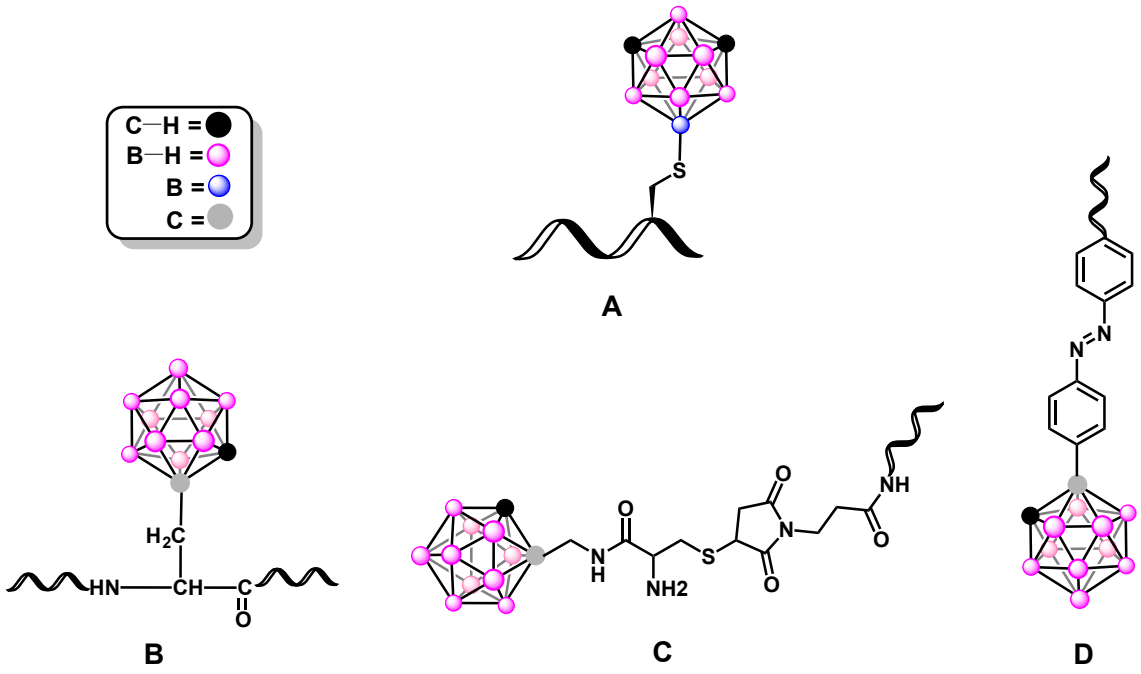

Figure 2-4. Different approaches for the bioconjugation of biomoleculess with carboranes. A) The reductive elimination approach described in this thesis, resulting in a direct S–B bond. B) Carboranylalanine as an analogue of phenylalanine within a biomolecule. C) Conjugation of carborane-thiol with maleimide-functionalized biomolecule. D) Carborane appended to biomolecule via azo-phenyl coupling.

Numerous other techniques have been reported for the conjugation of boron-rich clusters on biomolecules, which are described in the reviews by Stephenson et al. and Morin.[7, 13] While these approaches have demonstrated promising results, each requires specific preparation and is limited in its scope. The bioconjugation technique we report here is suitable for both small and large biomolecules, and offers control and selectivity over the incorporation of the carboranes.

2. 2. Methods

2. 2. 1. Peptide Synthesis

Peptides were synthesized on a 0.4 mmol scale using Fmoc-Solid Phase Peptide Synthesis

(SPPS) (**Figure 2-5**). Synthesis was carried out in a 25 mL vessel connected to a vacuum. Amino acid residues were activated using 0.4M NMM and coupled to rink amide resin as follow: 1) a 5 minute coupling of the amino acid with 2.0 mmol equivalents of 0.4M HCTU; 2) Five washes of DMF; 3) 3 minute deprotection with 20% (v/v) piperidine in DMF; and 4) Five washes of DMF. After all amino acids were coupled, the resin was washed with DMF, EtOH, and DCM. The peptide was cleaved from the resin using a solution of 18 mL TFA, 0.8 mL TIPS, 0.6 mL water, and 0.6 mL EDT in 2.5 hours. The solvent was then evaporated under argon gas for approximately one to two hours. The peptide was triturated three times (50 mL total) with cold diethyl ether and was allowed to dry for 20 minutes. The obtained residue was dissolved first in H_2O with 0.1% TFA, and then acetonitrile with 0.1% TFA was added until all solids were completely dissolved. The solution was lyophilized for two days.

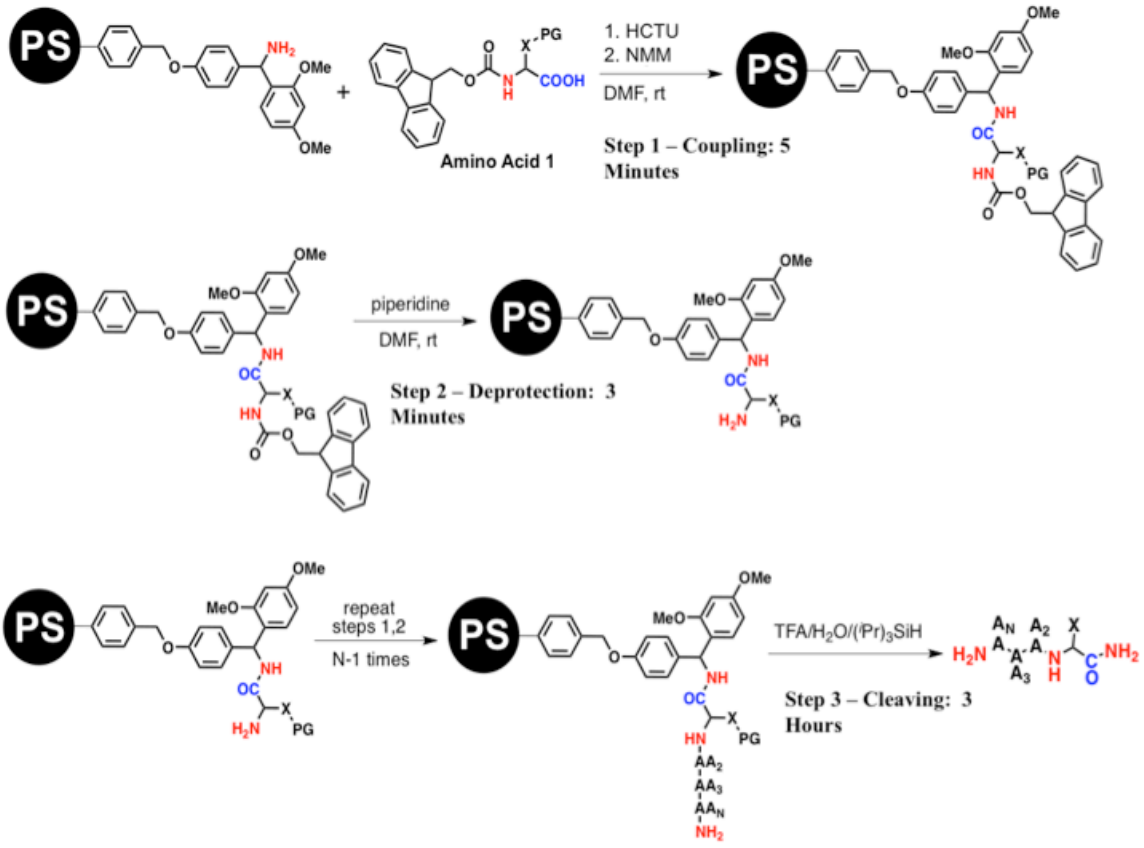

Figure 2-5. Solid-phase peptide synthesis (SPPS) general scheme. PS=Polystyrene resin.

Peptide	Sequence	Calculated Mass	Observed Mass
S1	NH_2-TKNVKDHVL**C**GDV-$CONH_2$	1425.7398	1425.7800
S2	NH_2-TKNVKDHVL**S**GDV-$CONH_2$	1409.7627	1409.7200

Table 2-1. List of peptides synthesized for this project, with calculated and observed masses.

***Note on Calculated/Observed Mass:** Calculated mass was determined using the Analysis function in ChemDraw Professional 15.0. The calculated mass reported here is the exact mass of the peptide sequence, as reported by ChemDraw. The observed mass was calculated using the highest-intensity $[M+2H]^{2+}$ peak (z=2). The observed peak was multiplied by 2, and the mass of 2 hydrogen atoms was subtracted from product to give the observed mass. These methods were followed throughout for consistency purposes.

2. 2. 2. Peptide Purification

Post lyophilization, the crude peptide was dissolved in 20 mL of 95% H_2O with 0.1% TFA and 5% acetonitrile with 0.1% TFA. The solution was filtered using a 0.22 µm filter and then purified using a Prep-HPLC (XBridge™ Prep C18 5µm OBD™, 19x50 mm). Solvents: 0.1% TFA in H_2O (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). Depending on the sequence (**Table 2-1**), each peptide was purified using a different gradient. The HPLC fractions containing the peptide were collected and tested for purity on MALDI-MS and LC-MS. The pure fractions were lyophilized and stored in the freezer at -4 °C.

2. 2. 3. LC-MS Analysis

LC-MS experiments were acquired using a Waters Acquity UPLC connected to a Waters LCT-Premier XE Time of Flight Instrument operated by MassLynx 4.1 software. The solvents used for analysis were 0.3% formic acid in H_2O (Solvent A) and 0.2% formic acid in acetonitrile

(Solvent B). All solvents are LC-MS/MS grade and purchased from Fisher Scientific. The following methods were used:

Method 2-A LC conditions: BEH C18 1.7 µm column (2.1 x 50 mm, Waters), gradient: 0 – 8 min, 0 – 50% B; 8 – 8.10 min, 50 – 95% B, 8.10 – 10 min, 95% B, 10 – 10.01 min, 95 – 5% B, 10.01 – 12 min, 5%, flow rate: 0.3 mL/min. LC trace was plotted from 1.5 – 10.5 minutes. MS conditions: positive mode, 70 – 4000 *m/z*, solvent delay of 1.5 minutes.

Method 2-B LC conditions: BEH C18 1.7 µm column (2.1 x 50 mm, Waters), gradient: 0 – 8 min, 2 – 50% B; 8 – 8.10 min, 50 – 95% B, 8.10 – 10 min, 95% B, 10 – 10.01 min, 95 – 5% B, 10.01 – 12 min, 5%, flow rate: 0.3 mL/min. LC trace was plotted from 1.5 – 10.5 minutes. MS conditions: positive mode, 70 – 4000 *m/z,* solvent delay of 1.5 minutes

2. 2. 4. Procedure Development

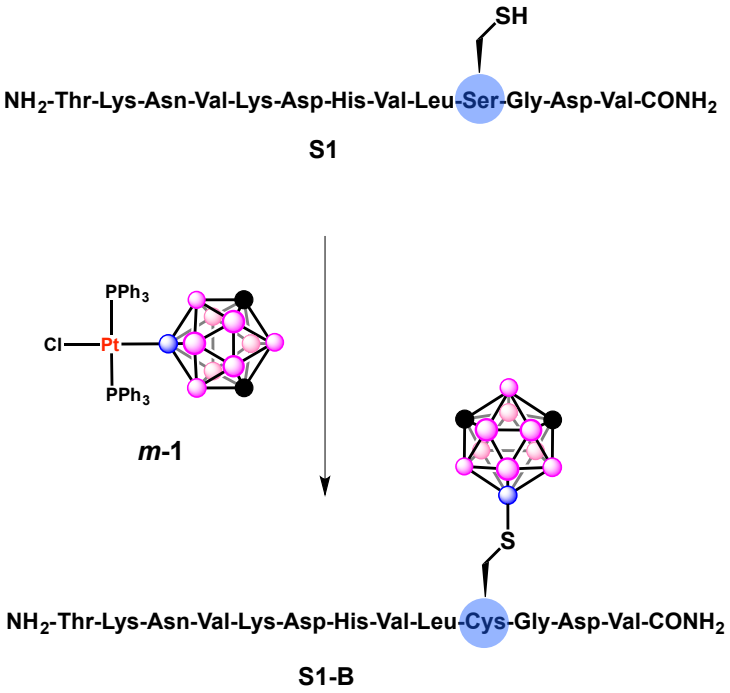

Figure 2-6. General scheme for cysteine conjugation of S1 with carborane tag using m-1.

We chose **S1** as the first peptide to use since we previously used the sequence in the work described in Chapter 1, and we knew it was a well-behaved sequence that can be easily isolated. We then incorporated a cysteine residue followed by a glycine residue in the sequence (**Figure 2-6, Figure 2-7**).

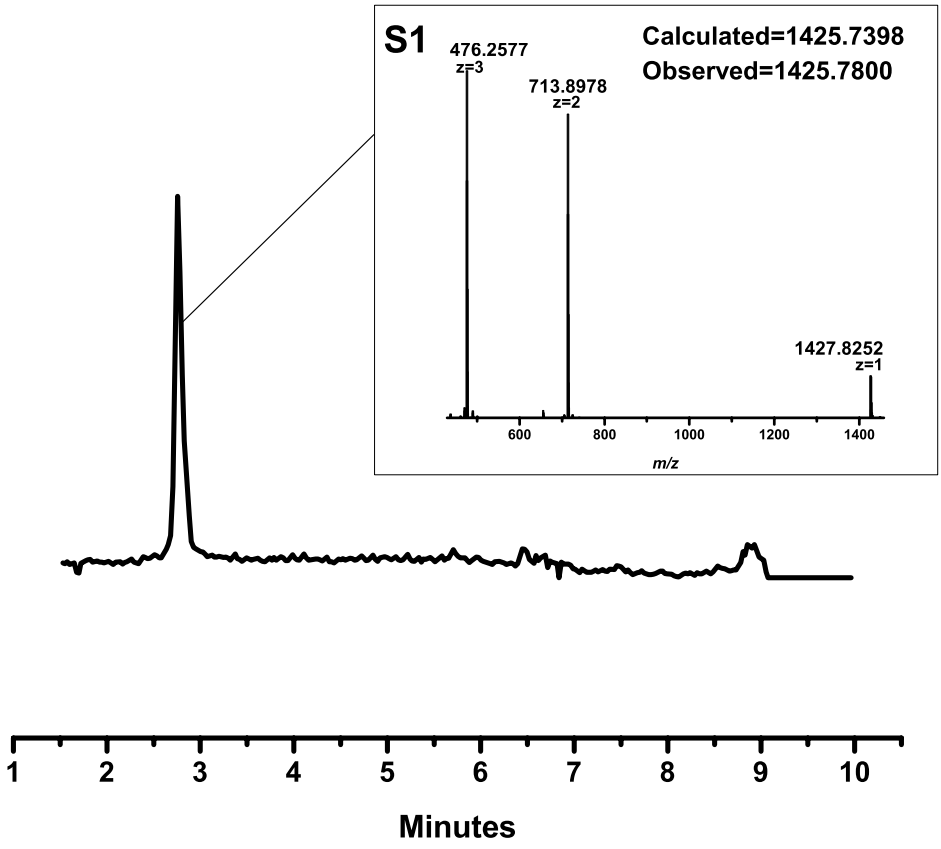

Figure 2-7. LC-MS TIC trace of S1, including mass spectrum. Obtained using Method 2-A.

Initially, we started by combining **S1** in an excess of [$PtCl(PPh_3)_2$(9-B-*m*-$C_2B_{10}H_{11}$)] (referred to as ***m*-1**) together with 100 mM Tris•HCl buffer to assess the reactivity of the reaction. We used approximately 3 equivalents of ***m*-1** per peptide and ran the reaction at 40 °C to ensure full conversion. It is important to note, that when calculating the moles of peptide used, we included the mass of TFA counter ions. We used DMF as the solvent because it dissolved both the peptide

(hydrophilic) and ***m*-1** (hydrophobic). A small amount of water was added to fully dissolve the peptide in the solution mixture. The final pH of the solution (although predominantly DMF) was approximately 7. We refer to the expected product as **S1-B** (**Figure 2-6**). After a few hours, aliquots were taken at 2 and 5 hours and quenched with 2:1 H_2O:ACN with 0.1% TFA. The aliquots were analyzed using Method 2-A by LC-MS (**Figure 2-8**).

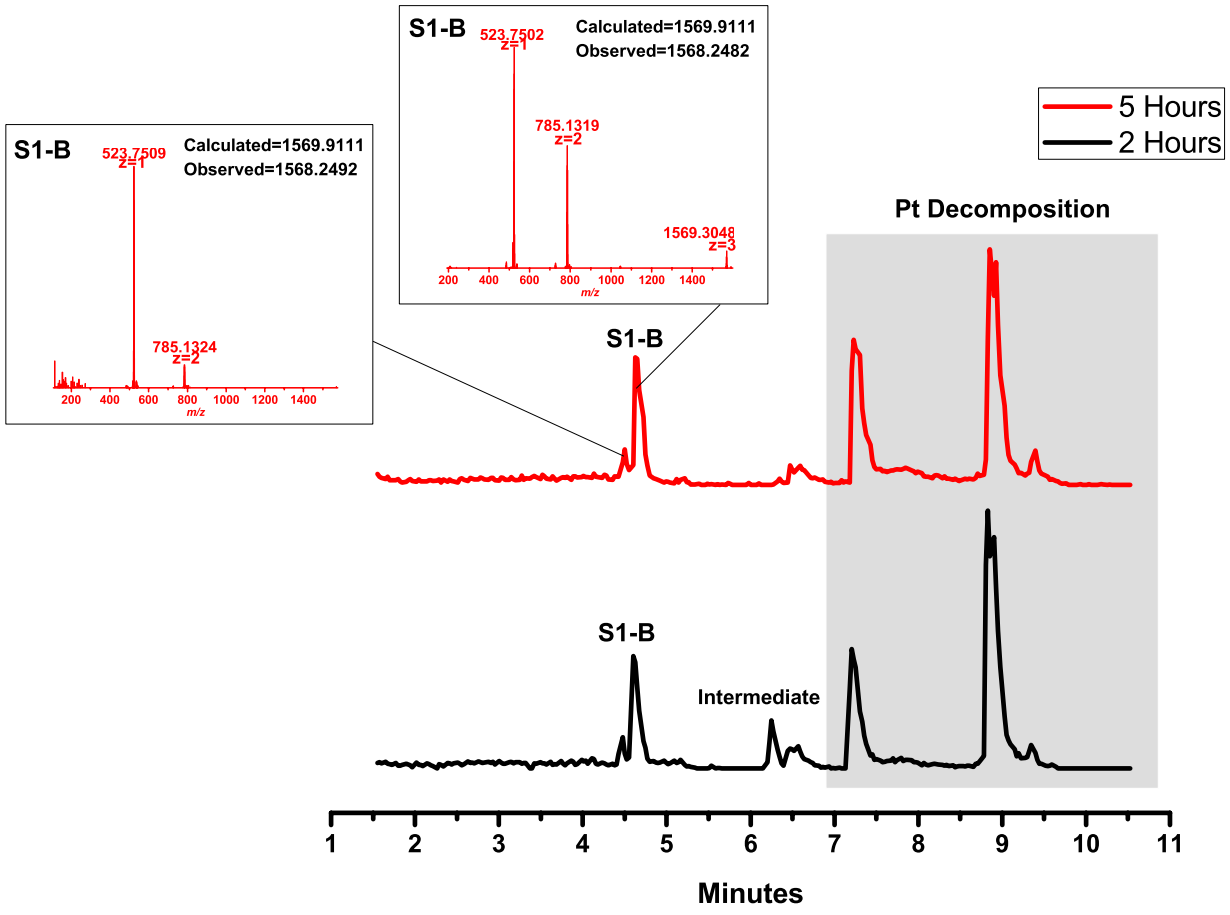

Figure 2-8. LC-MS TIC trace from preliminary experiment. Two peaks correspond to the mass of S1-B product, suggesting isomers are present. All of the peptide reacted, as there is no S1 peak present. Obtained using Method 2-A.

We noted that two peaks in the chromatogram corresponded to the mass of **S1-B**, suggesting the presence of two isomers. One peak is significantly larger than the other, indicating that the second isomer is a minor product. Since no isomers are observed in the **S1** peptide run (**Figure 2-7**), we are still investigating the root of the isomers. Moreover, some intermediate was observed at 2 hours, with mass corresponding to **S1**-$Pt(PPh_3)_2$. The peptide is still bound to platinum, suggesting that reductive elimination with the carborane did not take place. The carborane is most likely displaced when the reaction is quenched with TFA. However, at 5 hours, the intermediate peak was essentially gone. To optimize the reaction and maximize product conversion, we explored the following variables:

1) Solvent
2) Base type
3) Reagent excess
4) Temperature
5) Reaction duration

After evaluating the different variables, we came up with the following general procedure:

Optimized General Procedure. S1 was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, 1.2 equivalents of ***m*-1** was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred at room temperature. Reaction is complete within 1 hour (peptide has completely reacted). 20 µL aliquot of reaction mixture was quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using LC-MS. When the reaction was

confirmed to be complete by LC-MS, the remaining reaction mixture was quenched with 400 µL 2:1 H_2O:ACN with 0.1% TFA and stored in the freezer -4 °C.

2. 3. Results

2. 3. 1. Optimization Results for Reactions Using [PtCl(PPh$_3$)$_2$(9-B-m-C$_2$B$_{10}$H$_{11}$)]

Entry	Base	Reagent Excess (equivalents)	Solvent	Temperature
Entry 1	100 mM Tris•HCl	3	~9:1 DMF:H_2O	40 °C
Entry 2	100 mM Tris•HCl	3	**DMF**	40 °C
Entry 3	100 mM Tris•HCl	3	DMF	40 °C
Entry 4	**100 mM Trizma® base**	3	DMF	40 °C
Entry 5	100 mM Tris•HCl	3	DMF	40 °C
Entry 6	100 mM Tris•HCl	**1.2**	DMF	40 °C
Entry 7	100 mM Tris•HCl	1.2	DMF	40 °C
Entry 8	100 mM Tris•HCl	1.2	DMF	**RT**
Entry 9	100 mM Tris•HCl	1.2	DMF	RT
Entry 10 – Serine Control	100 mM Tris•HCl	1.2	DMF	RT

Table 2-2. Evaluation of reaction conditions. Highlighted in blue are the variables that were changed from the previous reaction.

*Note: Each time a variable was changed, a second reaction was run simultaneously with the conditions from the previous best result, for comparison purposes. These comparison reactions are also included here.

➔ **Effects of Solvent**

Procedure for Entry 1. S1 (0.8 mg, 4.24E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL H_2O was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 3.0 equivalents of ***m*-1** (1.0 mg, 0.0011 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (2 hours, 5 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-9**).

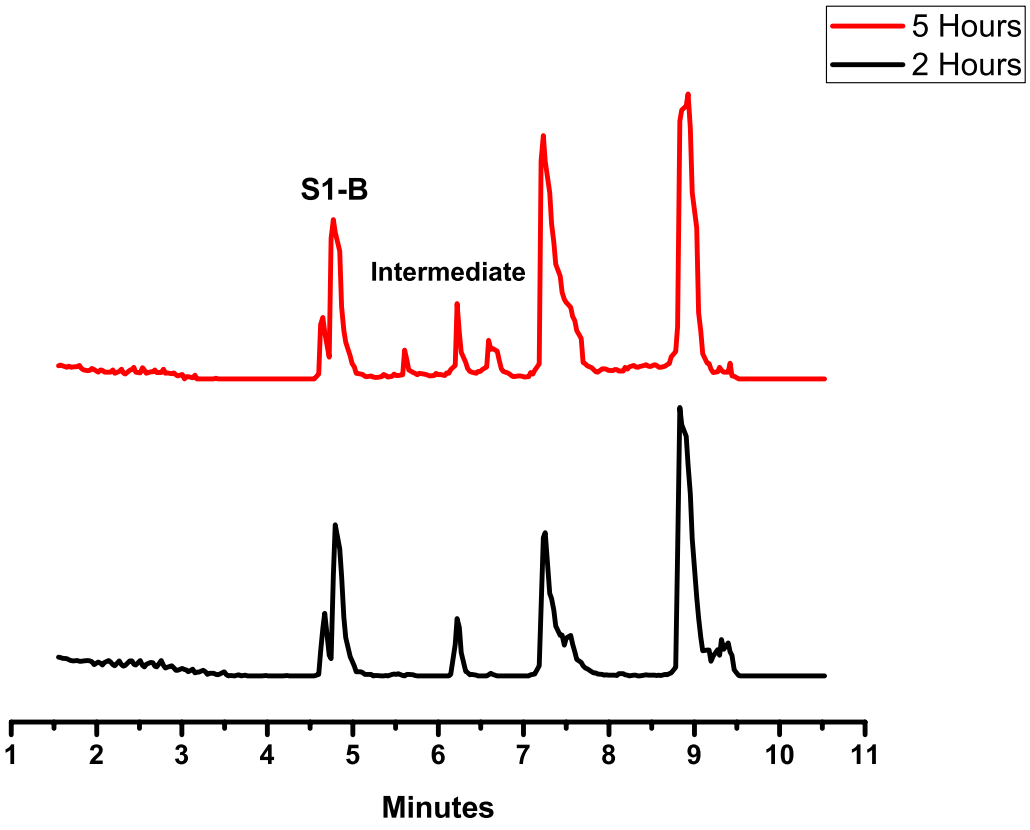

Figure 2-9. LC-MS TIC trace for Entry 1. Obtained using Method 2-A.

Procedure for Entry 2. S1 (1.1 mg, 5.84E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 3.0 equivalents of ***m*-1** (1.6 mg, 0.0018 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (2 hours, 5 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-10**).

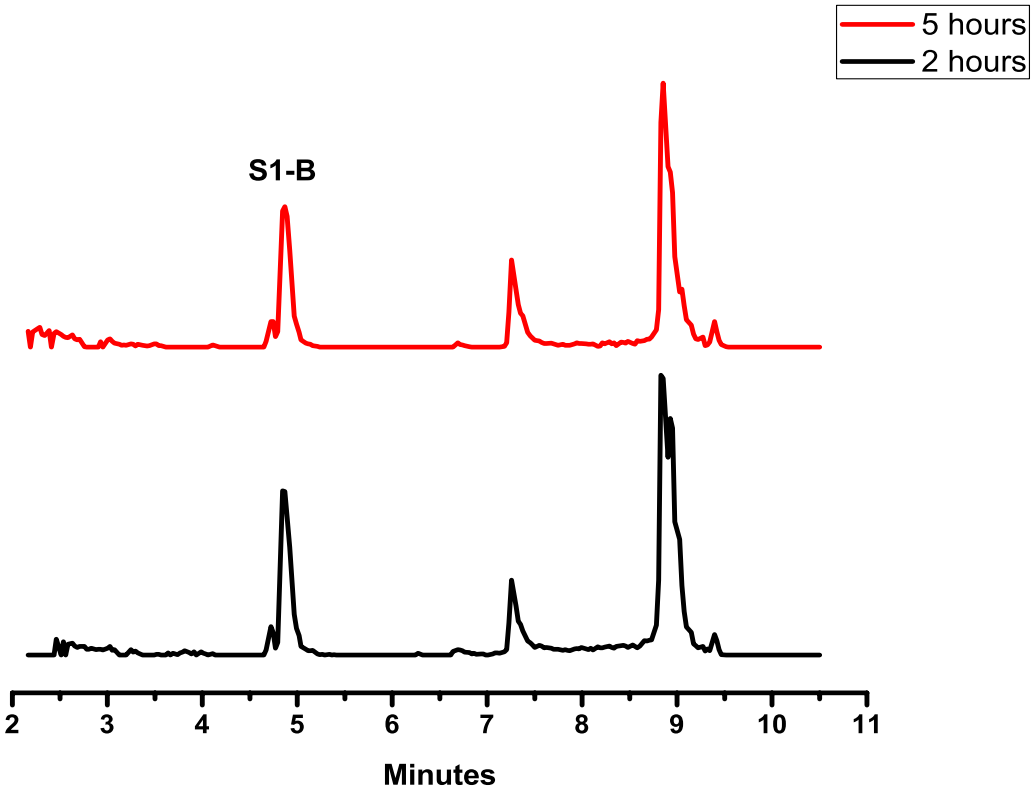

Figure 2-10. LC-MS TIC trace for Entry 2. Obtained using Method 2-A.

★ Based on these results, we chose DMF to be the primary solvent, since it resulted in minimal intermediate formation. Water appeared to interfere with the reductive elimination step,

➔ **Effects of Base Type**

Procedure for Entry 3. S1 (1.3 mg, 6.91E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 3.0 equivalents of ***m*-1** (1.6 mg, 0.0018 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (30 minutes, 1 hour, 2 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-11**).

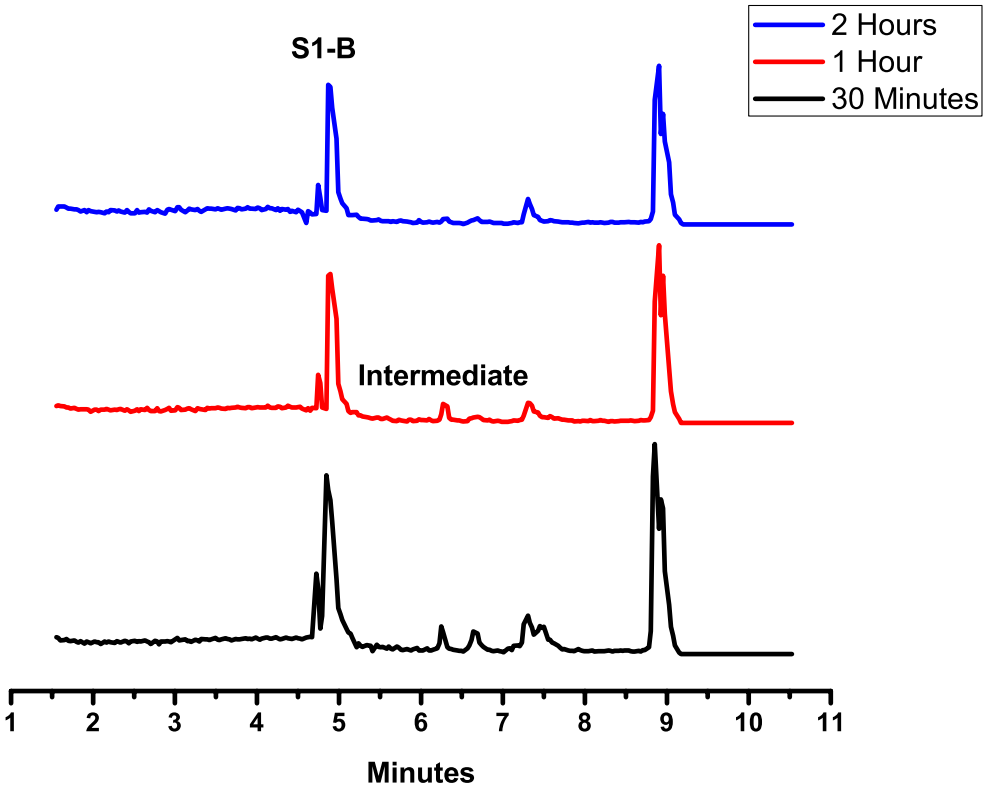

Figure 2-11. LC-MS TIC trace for Entry 3. Obtained using Method 2-A.

Procedure for Entry 4. S1 (1.1 mg, 5.84E-4 mmol) was dissolved in 100 µL Trizma® base in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 3.0 equivalents of ***m*-1** (1.6 mg, 0.0018 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (30 minutes, 1 hour, 2 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-12**).

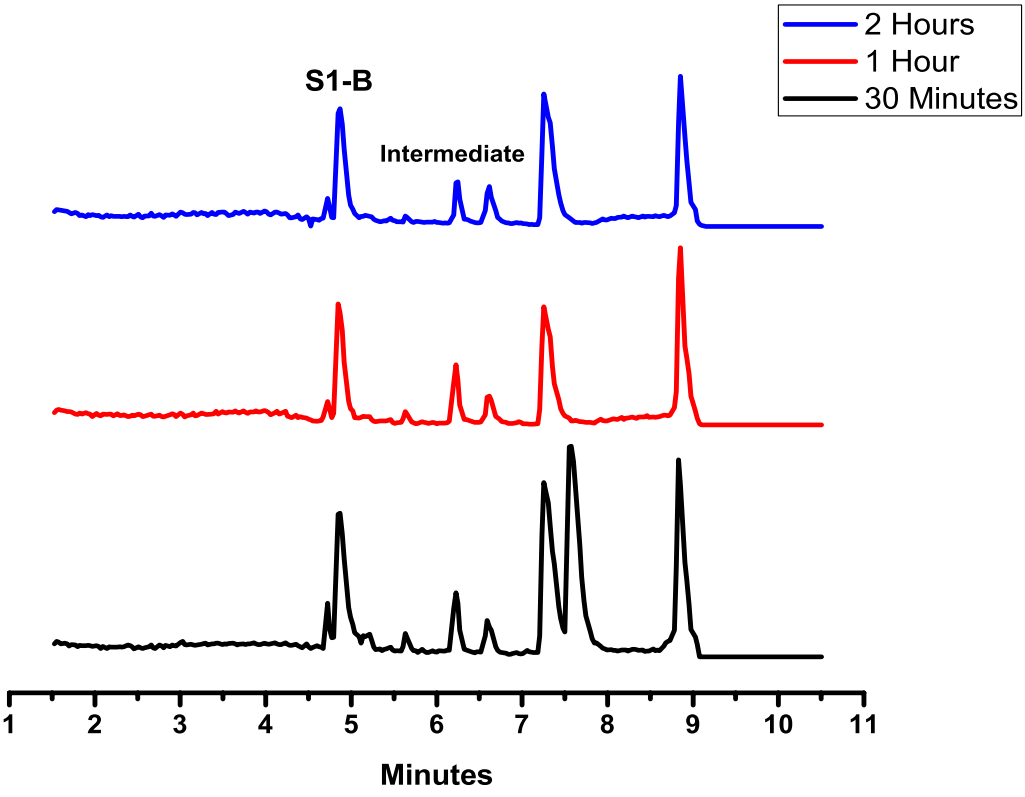

Figure 2-12. LC-MS TIC trace for Entry 4. Obtained using Method 2-A.

★ Based on these results, we chose to run the reactions with Tris•HCl, since it lead to less intermediate formation compared to Trizma® base.

➔ **Effects of Reagent Excess**

Procedure for Entry 5. S1 (1.3 mg, 6.91E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 3.0 equivalents of ***m*-1** (1.6 mg, 0.0018 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (5 minutes, 20 minutes, and 1.25 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-13**).

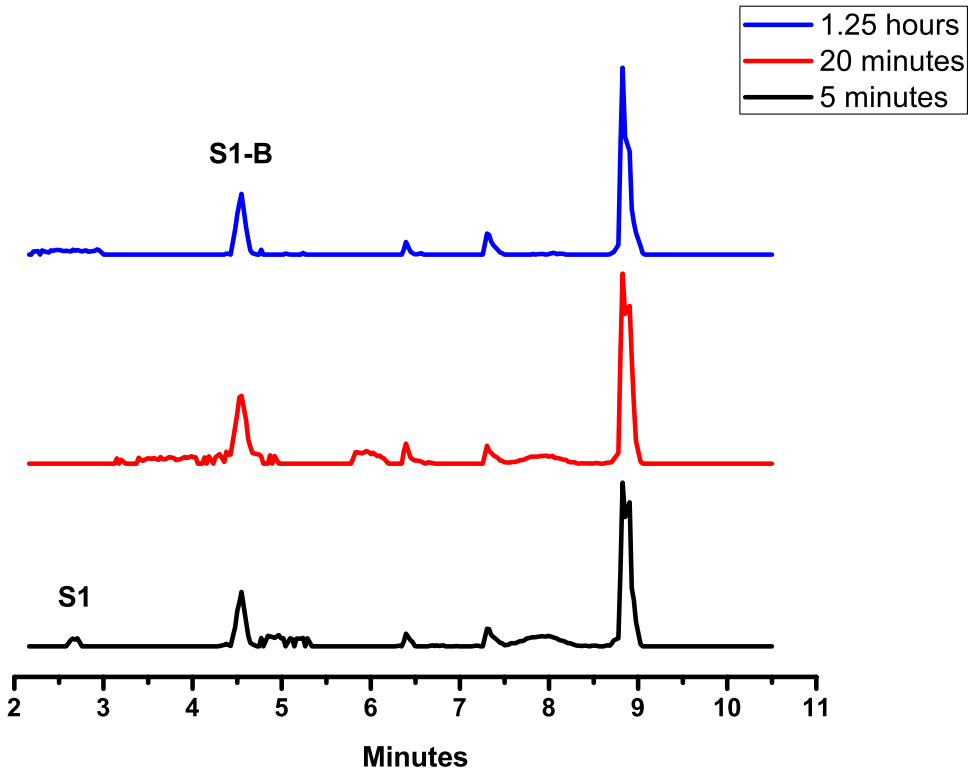

Figure 2-13. LC-MS TIC trace for Entry 5. Obtained using Method 2-A.

Procedure for Entry 6. S1 (1.0 mg, 5.31E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 1.2 equivalents of ***m*-1** (0.7 mg, 7.79E-4 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (5 minutes, 20 minutes, and 1.25 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-14**).

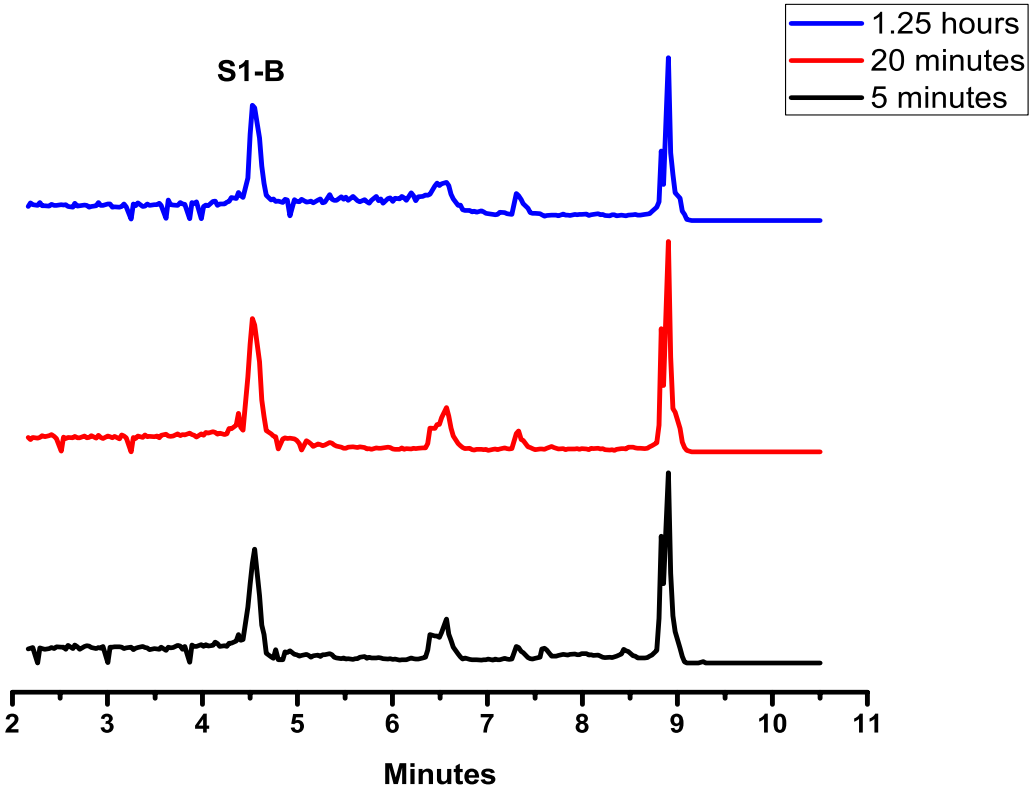

Figure 2-14. LC-MS TIC trace for Entry 6. Obtained using Method 2-A.

★ Based on these results, we reduced the reagent excess to 1.2 equivalents, since we still saw full conversion using the reduced amount.

➔ **Effects of Temperature**

Procedure for Entry 7. S1 (0.9 mg, 4.78E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 1.2 equivalents of ***m*-1** (0.8 mg, 8.91E-4 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred in a hot bath of 40 °C. 20 µL aliquots were sampled at different time points (10 minutes, 30 minutes, 1 hour) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-A on the LC-MS (**Figure 2-15**).

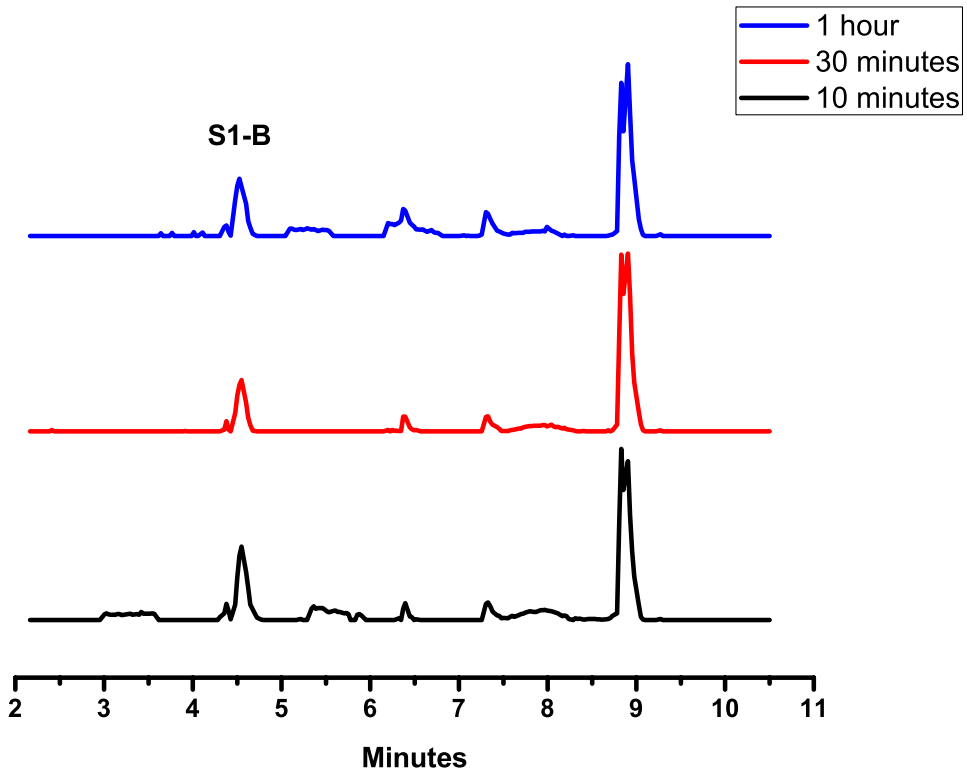

Figure *2-15. LC-MS TIC trace for Entry 7. Obtained using Method 2-A.*

Procedure for Entry 8. S1 (1.0 mg, 5.31E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 1.2 equivalents of the platinum complex (0.8 mg, 8.91E-4 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred at room temperature. 20 µL aliquots were sampled at different time points (10 minutes, 30 minutes, 1 hour) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method A on the LC-MS (**Figure 2-16**).

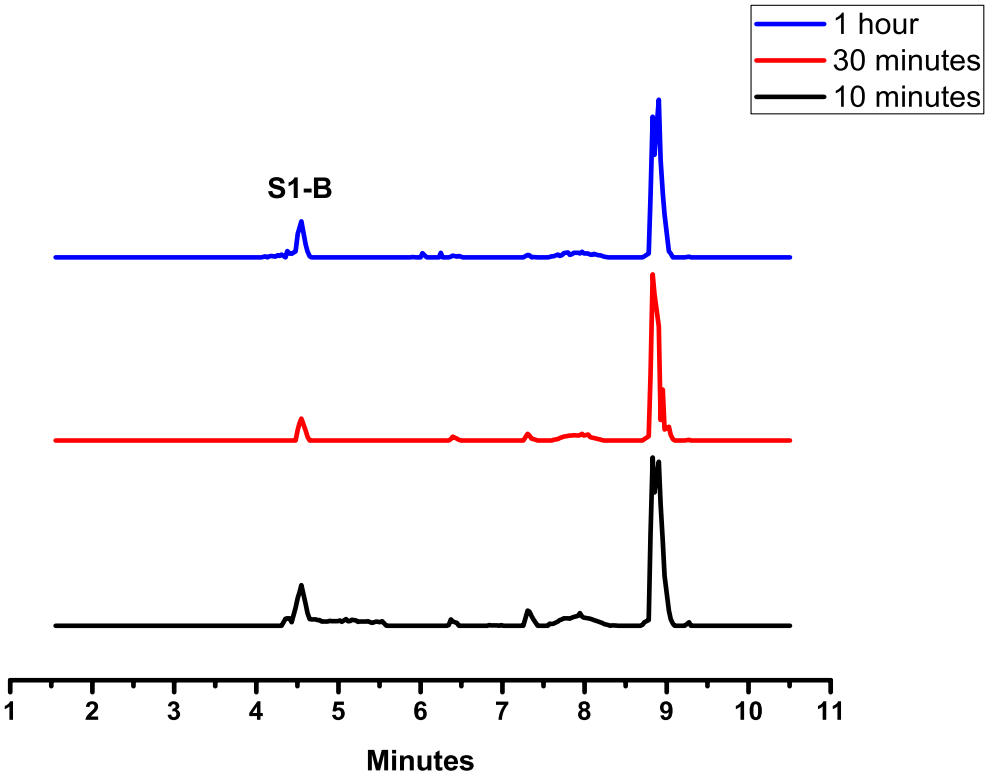

Figure 2-16. LC-MS TIC trace for Entry 8. Obtained using Method 2-A.

★ Based on these results, we started running the reactions at room temperature. While 40 °C is suitable for peptides, proteins may begin to degrade or denature at such temperatures.

➔ **Effects of Reaction Time**

Procedure for Entry 9. S1 (3.3 mg, 0.0018 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 1.2 equivalents of ***m*-1** (1.9 mg, 0.0021 mmol) was dissolved in 200 µL DMF and sonicated for 10 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred at room temperature. 20 µL aliquots were sampled at different time points (10 minutes, 30 minutes, 1 hour, 2 hours) and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-B on the LC-MS (**Figure 2-17**).

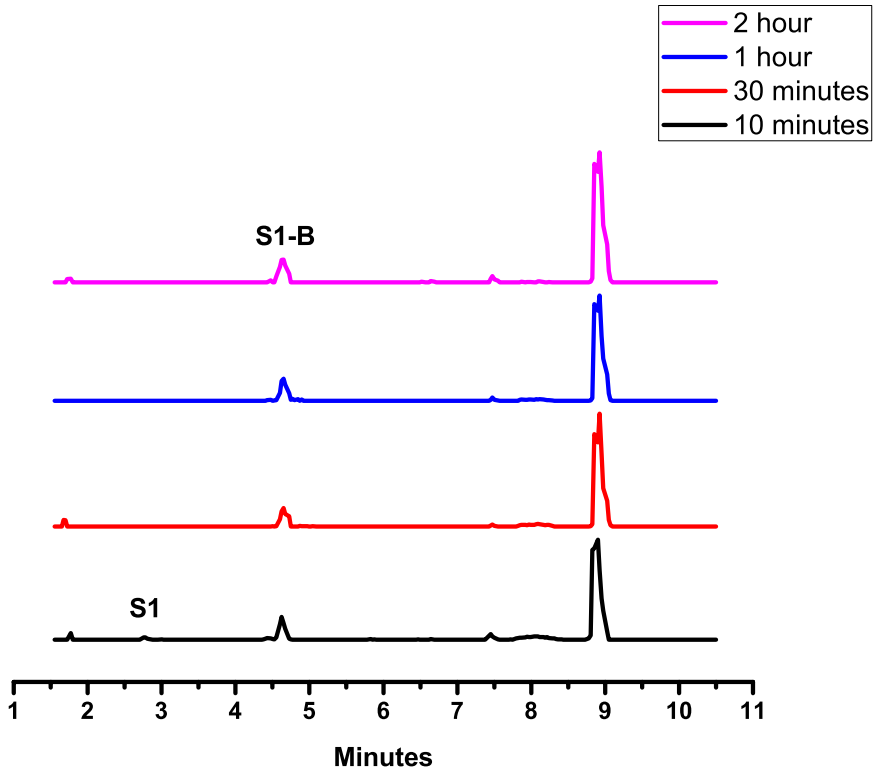

Figure 2-17. LC-MS TIC trace for Entry 9. Obtained using Method 2-B.

★ Based on these results, we see full conversion of the peptide by 1 hour—which is significantly faster than the 5 hour reaction we used in our preliminary experiments.

➔ Serine Control Reactions

In order to ensure that the reaction is selective for cysteine, we ran a control using peptide **S2**, which contained a serine residue in place of the cysteine (**Table 2-1**). We expected to see no product formation. **P2** was run on LC-MS using Method 2-B, and the TIC was extracted to show ions from 300 – 4000 *m/z* (**Figure 2-18**).

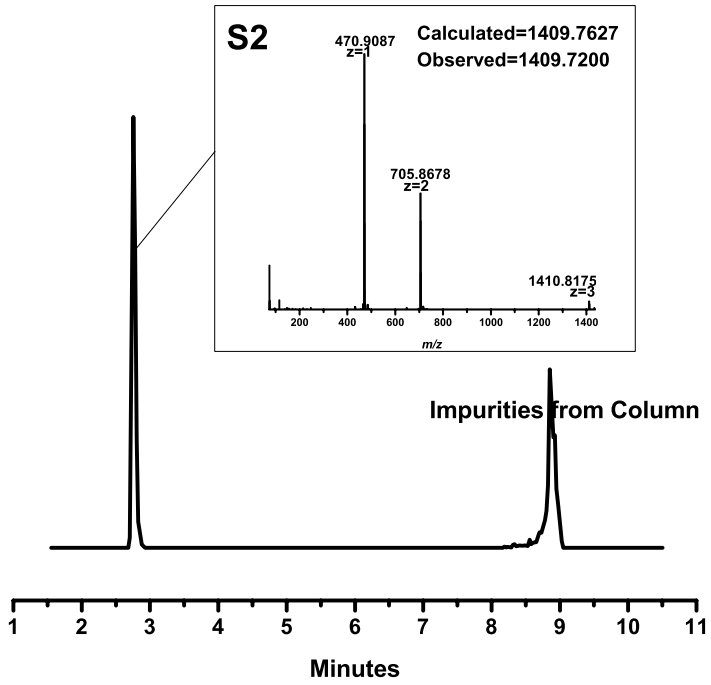

Figure 2-18. LC-MS TIC trace for S2, including mass spectrum. Obtained using Method 2-B. TIC was extracted to show ions from 300 – 4000 m/z.

Procedure for Entry 10 – Serine Control. S2 (1.0 mg, 5.36E-4 mmol) was dissolved in 100 µL Tris•HCl in DMF (100 mM) in an Eppendorf tube. An additional 40 µL DMF was added, and the resulting solution was vortexed for 5 seconds. In a separate vial, approximately 1.2 equivalents

of ***m*-1** (0.58 mg, 6.46E-4 mmol) was dissolved in 200 µL DMF and sonicated for 20 seconds. The two solutions were combined in a new reaction vial (using a micropipette) and stirred at room temperature. 20 µL aliquot was sampled at 1 hour and quenched with 200 µL 2:1 H_2O:ACN with 0.1% TFA. The quenched aliquots were analyzed using Method 2-B on the LC-MS, and the TIC was extracted to show ions from 300 – 4000 *m/z* (**Figure 2-19**).

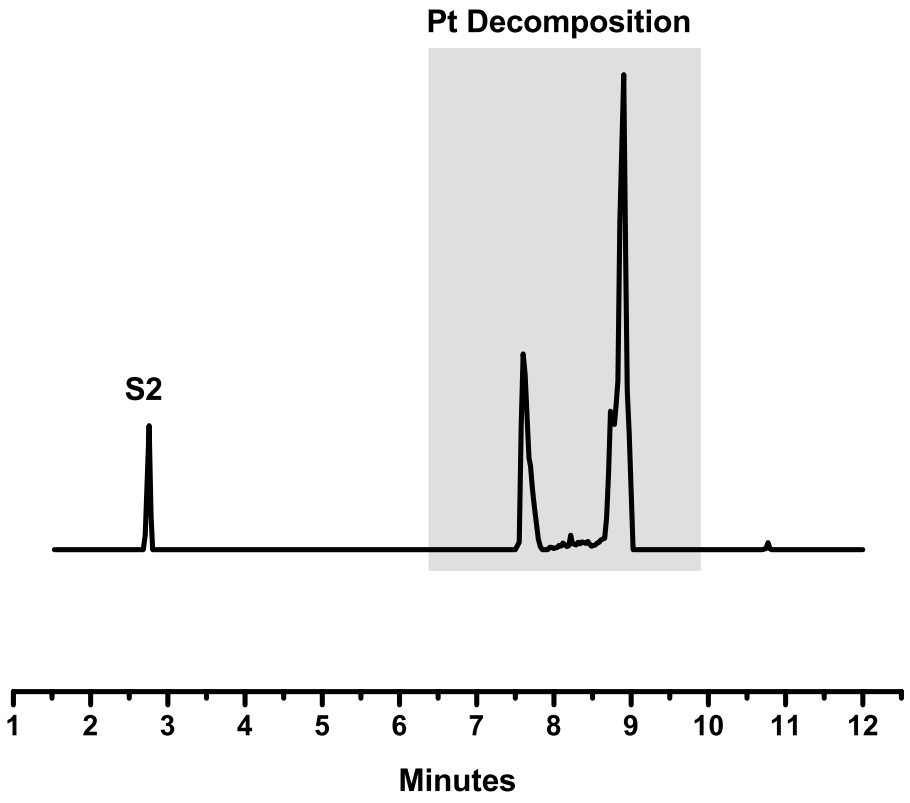

Figure 2-19. LC-MS TIC trace for Entry 10, control reaction with serine. Obtained using Method 2-B. TIC was extracted to show ions from 300 – 4000 m/z.

As expected, we see no product formation. We see only peaks corresponding to **S2** alone, as well as Pt decomposition products. These results confirm that the reaction is selective for cysteine.

Moreover, it verifies that the isomer peaks we previously observed do no stem from the reagent reacting with other functional groups on the peptide.

2. 3. 2. Product Isolation

After confirming the reaction was complete via LC-MS, the remaining reaction mixture was quenched with 400 μL of 2:1 H_2O:ACN with 0.1% TFA and stored in the freezer. To isolate product, the quenched reaction mixture was lyophilized for approximately five hours. Once dry, the residue was dissolved in water and purified using a 0.22 μm filter to remove the insoluble particles (most likely reagent decomposition products). The solution was then purified using Prep-HPLC (ZORBAX Rx-C8, 9.4 mm x 25 cm). Solvents: 0.1% TFA in H_2O (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). Gradient: 0 – 80 min, 0 – 50% B, 80 – 95 min, 50 – 95% B, 95 – 105 min, 95 – 5% B, 105 – 110 min, 5% B, flow rate: 10 mL/min. The fractions containing the product were run on LC-MS to determine purity. The pure fractions were combined and lyophilized. Once dry, the product was stored in the freezer at -4 °C.

2.4. Discussion and Future Steps

The results we obtained so far indicate that we can successfully and selectively label cysteine residues using ***m*-1**. Based on the control performed with **S2**, we can conclude that the reaction is selective to cysteine, and does not react with other functional groups on the peptide side-chains. Another control would be to run a peptide digest, such as trypsin, to ensure that the carborane only adds where we expect. As we continue work on this project, we plan to label additional sequences, such as a nuclear localization signal (NLS). These sequences tag proteins for delivery into the nucleus; by conjugating carboranes onto these sequences, we can visualize our targets

within the nucleus. We also intend to work with sequences containing multiple cysteines, to see if we can tag the sequence with multiple carboranes. Additionally, we can staple the peptide using a biscarborane motif (**Figure 2-20**). It has been reported that stapled peptides show increased stability against proteasomal degradation and enhanced cell permeability.[11]

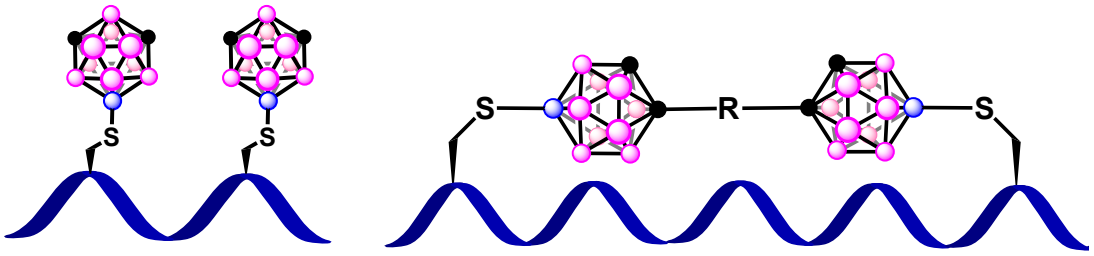

Figure 2-20. Models of peptides containing two cysteines, highlighting potential future directions.

We also intend to assess other organometallic reagents containing different ligands and/or *ortho*-carboranes. We would like to understand how the design of the reagent affects the progress of the reaction. Ultimately, we want to use these developed conditions to tag larger biological molecules such as proteins and potentially antibodies, and use Raman imaging to track them within the cell.

References

1. Folick, A., Min, W., and Wang, M. C. *Current Opinion in Genetics & Development,* **2011**, 21, 585–590.

2. Freudiger, C. W., Min, W., Saar, B. G. Lu, S., Holtom, G. R., He, C., Tsai, J. C., Kang, J. X., Xie, X. S. *Science*, **2008**, 322 (5909), 1857−1861.

3. Evans, C. L., and Xie, X. S. *Annu. Rev. Anal. Chem.,* **2008**, 1, 883–909.

4. Wang, M. C., Min, W., Freudiger, C. W., Ruvkun, G., Xie, X. S. *Nat. Method.*, **2011**, 8(2) 135–138.

5. Wei, L., Hu, F., Shen, Y., Chen, Z., Yu, Y. Lin, C-C., Wang, M. C., and Min, W. *Nat. Method.*, **2014**, 11, 410–412.

6. Wei, L., Yu, Y., Shen, Y., Wang, M. C., and Min, W. *Proc. Natl. Acad. Sci*., **2013**, 110 (28), 11226–11231.

7. Valliant, J. F., Guenther, K. J., King, A. S., Morel P., Schaffer, P., Sogbein, O. O., and Stephenson, K. A. *Coordination Chemistry Reviews*, **2002**, 232 173–30.

8. Vinogradova, E. V., Zhang, C., Spokoyny, A. M., Pentelute, B. L., and Buchwald, S. L. *Nature,* **2015**, 526, 687–691.

9. Saleh, L. M. A., Dziedzic, R. M., Khan, S. I., Spokoyny, A. M. *Chem. Eur. J.* **2016,** online.

10. Leukart, O., Caviezel, M., Eberle, A., Escher, E., Tun-Kyi, A., Schwyzer, R. *Helvetica Chimica Acta*, **1976**, 59 (6), 2184–2187.

11. Neirynck, P., Schimer, J., Jonkheijm, P., Milroy, L.-G., Cigler, P., and Brunsveld, L. *J. Mater. Chem.* B, **2015**, 3, 539–545.

12. Mizusawa, E., Dahlman, H. L., Bennett, S. J., Goldenberg, D. M., and Hawthorne, M. F. *Proc. Natl. Acad. Sci*., 1982, 79, 3011–3014.

13. Morin, C. *Tetrahedron*, **1994**, 50 (44), 12521–12569.

14. Spokoyny, A. M., Zou, Y., Ling, J. J., Yu, H., Lin, Y.-S. and Pentelute, B. L. *J. Am. Chem. Soc.*, **2013**, 135 (16), 5946–5949.