4-2017

Utilization of Unnatural Amino Acids in Novel Bioconjugates and Probing Applications

Christina A. Howard
College of William and Mary

Utilization of Unnatural Amino Acids in Novel Bioconjugates and Probing Applications

Christina Anne Howard

Manassas, Virginia

A Thesis Presented at the College of William and Mary in Candidacy for Departmental Honors

Chemistry Department

College of William and Mary

May, 2017

A thesis submitted in partial fulfillment of
the requirements for the degree of
Bachelor of Science in Chemistry from
The College of William and Mary

_____
Christina Anne Howard

Approved by the Committee, April 2017

_____
Committee Chair
Associate Professor Douglas Young, Chemistry
The College of William and Mary

_____
Garrett-Robb-Guy Professor Lisa Landino, Chemistry
The College of William and Mary

_____
Professor John Poutsma, Chemistry
The College of William and Mary

_____
Associate Professor Oliver Kerscher, Biology
The College of William and Mary

ABSTRACT

Bioconjugations are utilized in many fields including materials science, biochemistry and medicine, despite the limited chemistries available in biomolecules. Unnatural amino acids can be used to expand the chemical diversity in proteins, affording a greater variety of functional groups for bioconjugations which. The site-specific incorporation of unnatural amino acids confers greater control and specificity over the reactions. Applications of unnatural amino acid based bioconjugations will be explored in this thesis. Optimization of solid supported immobilization of GFP and the extension of the technology to a carboxylesterase will be described. Fluorescent labeling of a medically relevant enzyme, Utag, will be optimized as the first step in developing a novel screening system for prostate cancer. Development of reactions to functionalize divalent bioconjugates into multivalent complexes will also be discussed. The utility of GFP containing an unnatural amino acid will be extended using various conjugation partners. Additionally, the unique chemistry of 3-fluorotyrosine will be utilized to study the radical pathway of a multifunctional hemoglobin, Dehaloperoxidase. In short, unnatural amino acid technology will be leveraged to explore novel applications and reactions of bioconjugates and the radical pathway of Dehaloperoxidase.

TABLE OF CONTENTS

Chapter 7. Utilizing 3-Fluorotyrosine to Probe the Radical Pathway of Dehaloperoxidase

ACKNOWLEDGEMENTS

I'd like to dedicate this thesis to my grandmother, Ann Woodward, for making my college career possible.

Chapter 1: Introduction

Bioconjugations involve a reaction between a biological macromolecule, such as DNA, RNA or protein, and another secondary molecule. The secondary molecule can be of biological, organic or inorganic origin, such as a solid support, a fluorophore, a radioactive label or another biomolecule.[1] Bioconjugations are ubiquitous tools utilized in biochemistry, materials science, and medical applications, for everything from probes to drug delivery mechanisms to the development of novel materials.[1,3,4] Bioconjugations that utilize proteins make use of the few reactive chemical groups present in the 20 canonical amino acids. However, more useful bioconjugations exist that employ chemical functionalities often found outside of a biological context. This is primarily due to the orthogonality of these groups affording a higher degree of control over the reaction. Two extremely useful bioconjugation reactions are the 1,3-dipolar cycloaddition (CuAAC) between an alkyne moiety and an azide moiety and the Glaser-Hay reaction between two acetylene groups.[1] These reactions can be applied to proteins via inserting the alkyne or azide moieties into proteins via unnatural amino acids (UAAs). These UAAs can be site specifically incorporated into proteins using the Shultz methodology, which allows for greater control of bioconjugation reactions. This more sophisticated control allows for the development of novel protein-based bioconjugates, some of which will be described in this thesis.

Examples and Limitations of non-Bioorthogonal Bioconjugations

Many bioconjugations make use of the relatively limited chemical functionality of the 20 natural amino acids. The most commonly used functional groups used are the amine present in lysine residues and the thiol present in cysteines.

IgG

Toxin, drug, radioisotope or enzyme
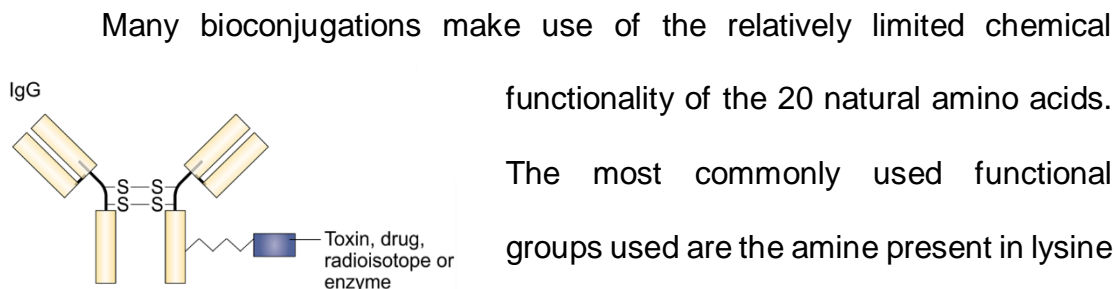
Figure 1.1: A LTT bioconjugate.[2]

These chemistries are utilized in in ligand-targeted therapeutics (LTTs) to covalently attach a therapy molecule to a targeting biomolecule, such as an antibody (IgG, Figure 1.1).[2] The antibody will only bind to the target cell, where it is typically endocytosed, delievering the toxic payload only to diseased cells. This prevents the drug from affecting healthy cells, allowing for higher dosage of the drug and an increased effectiveness of the therapy.[2]

Bioconjugations are also applied in the developing field of protein microarrays. These consist of a layer of proteins immobilized on a solid support, typically a flat surface. These microarrays have been developed for use in clinical applications to replace expensive diagnostic tools in poor countries.[3] Laying down a uniform surface of proteins on these "biochips" is important to create reproducible and stable binding of the substrate of interest, but this can prove challenging without the use of UAAs. Researchers are forced to use either reactive natural amino acids or chemical groups that are added to the protein via a chemical reaction with these reactive amino acids.[3] Neither of

these methods affords control over the orientation of the protein, resulting in a heterogenous layer of the protein.

Utilizing the naturally occurring lysine and cysteine allows for facile production of the protein of interest, but limits the scope of available chemistries for bioconjugation. In addition to being chemically limited, this method of bioconjugation allows
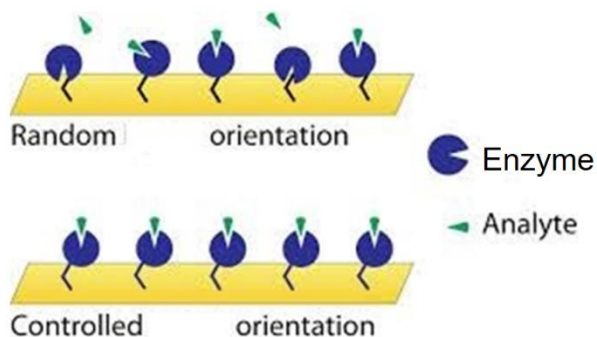
Figure 1.2: Two methodologies for enzyme immobilization. Utilization of natural reactive groups results in a heterogenous layer of randomly oriented enzymes, while immobilization using a site specifically incorporated UAA results in a homogenous layer of protein in a controlled orientation.[4]

for a heterogenous mixture of products. For example, when utilizing a reactive lysine residue to immobilize an enzyme, there is no way to control which lysine residue is reacted with the surface. Some proteins will be conjugated to the surface in such a way that the active site will be blocked, rendering that protein useless in the reaction (Figure 1.2). This defeats part of the purpose of immobilization, by decreasing the effective concentration of active enzyme. Additionally, bioconjugation reactions that utilize naturally occurring chemistries are not useful for conjugation in biological systems. Since these cysteine and lysine residues are common to all proteins, side reactions will occur with other proteins besides the protein of interest. UAAs used in conjunction with bioorthogonal reactions, such as the CuAAC the Glaser-Hay, offer a solution to this cross reactivity.

Bioorthogonal Bioconjugation Reactions

The CuAAC reaction employs an azide moiety and an alkyne moiety and yields 1,4-linked 1,2,3-triazoles (Scheme 1.1).[5] Neither of these chemical groups are present in biological systems, making the CuAAC bioorthogonal. The CuAAC reaction is an optimization of the Huisgen 1,3-cycloadditon which required higher temperatures.[6] Meldal and Sharpless both produced improvements to this reaction in the form of a copper catalyst that resulted in a decrease in the temperature requirements for the reaction.[6]
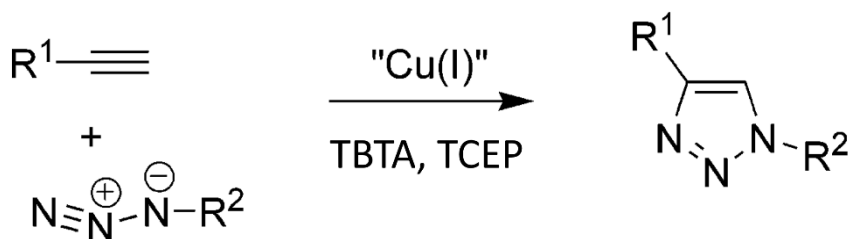
Scheme 1.1: The CuACC "click" reaction between an acetylene and an azide.[5]
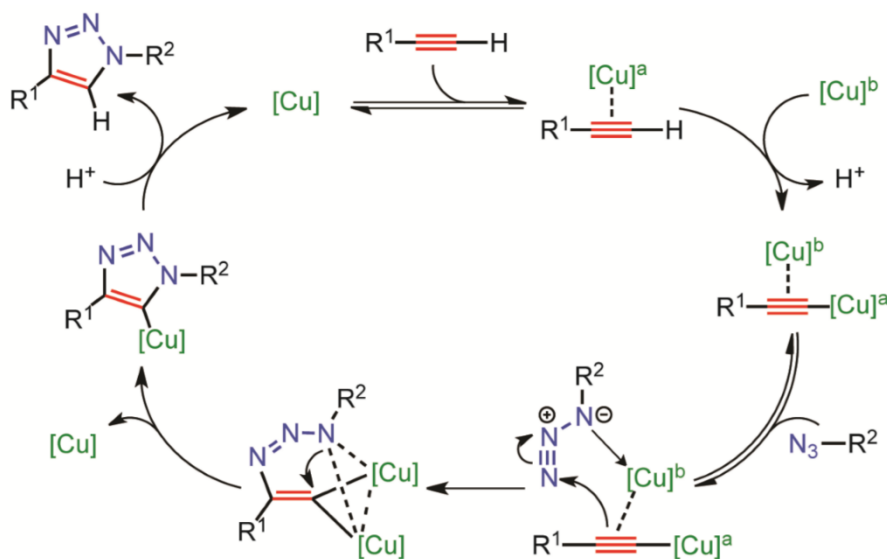
Scheme 1.2: Reaction Mechanism for the CuAAC reaction.[7]

The addition of a tetradentate ligand, such as Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA), also improved the reaction, as it increased the reaction rate and lower the time needed for the reaction.[1] The shorter reaction time is advantageous for biological applications,

as the catalytic copper can degrade proteins during the reaction. Copper I is thermodynamically prone to oxidation into Cu^{2+} and disproportionation into Cu^0 and Cu^{2+}.[5] This prevents efficient catalysis, as the Cu^{1+} concentration is constantly changing. TBTA solves this problem by binding to the Cu atom. This prevents the copper from participating in destabilizing reactions. Additionally, the tertiary nitrogen provides greater electron density to the copper center, increasing the reaction rates. The 1,2,3-triazole moieties are released from the metal to allow for catalytic binding between the copper and the alkyne.[5] TCEP is a reducing agent that also helps prevent the oxidation of the copper. The reaction proceeds through the mechanism illustrated in Scheme 1.2. This click reaction can be utilized with proteins containing either the *para*-propargyloxyphenylalanine (*p*PrF) or *para*-Azidophenylalanine (*p*AzF) UAAs to conjugate proteins to a fluorophore (Chapter 4), a solid support (Chapter 2) or other secondary reaction partners.

The other main bioconjugation reaction utilized in this thesis is the Glaser-Hay reaction, which involves the coupling of two terminal alkynes to form a diyne (Scheme 1.3).
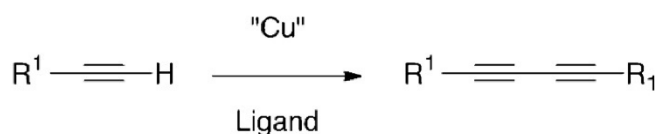
Scheme 1.3: Overall Reaction of two acetylenes through the Glaser-Hay reaction.[8]

This diyne linkage is useful in bioconjugation reactions as it allows for bioorthogonal reactions and provides an electron dense linkage that can be utilized in other reactions to create a multivalent complex (Chapter 5). The organic version of

this reaction was first described by Glaser as an optimization of the Eglinton reaction, using only catalytic copper.[9] This reaction was further improved by the addition of TMEDA by Hay. The TMEDA functions as a bidentate ligand that stabilizes the Cu I, similar to the TBTA function in the CuAAC reaction.[9] The
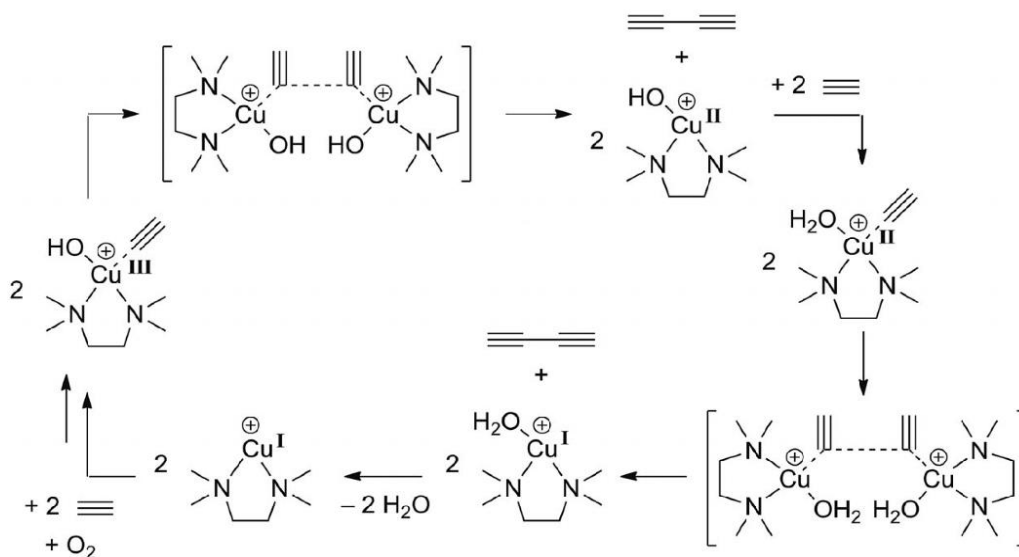
Scheme 1.4: Computationally generated reaction mechanism for the Glaser-Hay reaction.[10]

reaction mechanism was studied computationally using the DFT methodology, and the resulting proposed reaction mechanism is shown in Scheme 1.4.[10] This reaction was not utilized in bioconjugation until it was optimized for biological conditions by Lampkowski and co-workers in 2015. Lampkowski et al. found that the reaction proceeds best at 4 °C for 6 h. This results in 93% coupling between the fluorophore and the protein without undue amounts of protein degradation.[8] Minimal homodimerization between the fluorophore and the GFP was observed, likely due to the bulk of the protein[1,8] This reaction can be utilized in proteins containing *p*PrF or *p*EtF to immobilize the protein (Chapter 2), to label the protein with a fluorophore (Chapter 4), or to conjugate the protein with

an aromatic partner to change the function of the protein (Chapter 6). These UAAs are incorporated into proteins via the Schultz methodology, which exploits the central dogma of molecular biology.

Central Dogma of Molecular Biology

Incorporation of unnatural amino acids takes advantage of the natural protein synthesis mechanisms in *Escherichia coli*. The two steps of protein synthesis are transcription and translation, which are linked together in the central dogma of molecular biology (Figure 1.3).[11,12] An organism's DNA stores
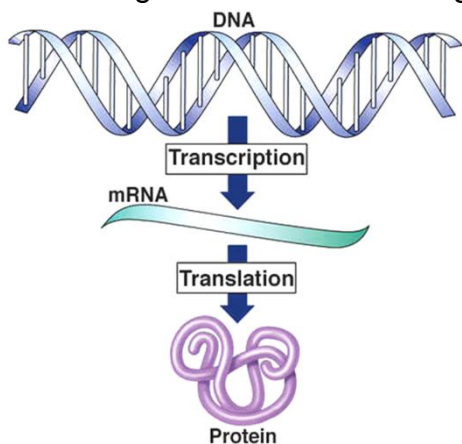
Figure 1.3: The Central Dogma of molecular biology.[11]
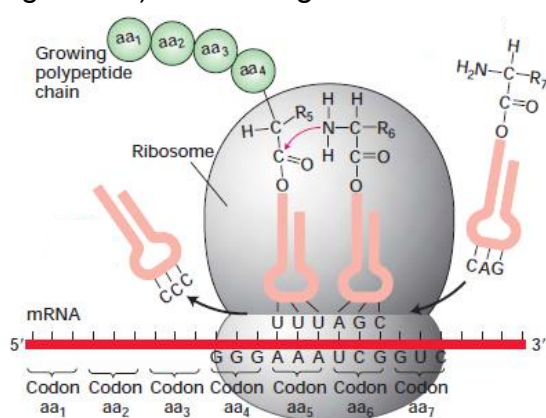
Figure 1.4: Translation of mRNA into a Polypeptide via charged tRNAs.[13]

the information needed for the synthesis of a particular protein as a gene, a unique sequence made up of four nucleotide bases.[12] These nucleotides are arranged into different groups of three to form codons, which each code for an amino acid.[12] During transcription, the DNA encoding the protein is used as a template for RNA Polymerase to synthesize mRNA. Once the gene has been transcribed, the mRNA is then transported out of the nucleus.[12] These mRNAs enter the cytoplasm, and a ribosome is assembled around them. The ribosome

uses tRNAs charged with amino acids to translate the codons in the mRNA sequence into a chain of amino acids, known as a polypeptide (Figure 1.4). The polypeptide will then fold to create secondary and tertiary structures, resulting in the final form of the protein.[12]

During translation, the mRNA sequence is converted to amino acids using a tRNA charged with a specific amino acid. (Figure 1.4). The tRNA is charged with its amino acid by a specific aminoacyl tRNA synthetase (aaRS) in a reaction known as aminoacylation (Figure 1.5). The tRNA and corresponding aaRS are known as a tRNA/aaRs pair. The aaRS contains highly selective binding sites for the amino acid and tRNA, ensuring that they are paired together
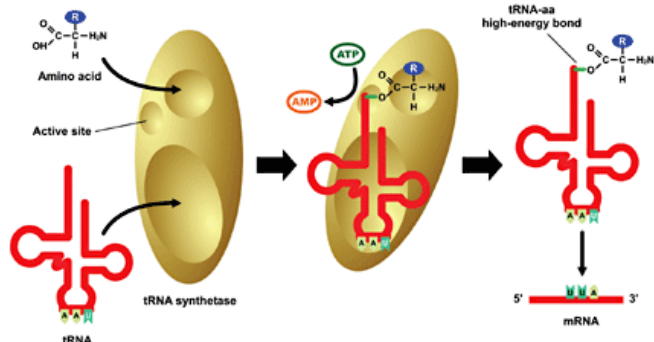
Figure 1.5: Charging a tRNA with an amino acid.[14]

correctly.[12] Once both the tRNA and the proper amino acid are bound to the aaRS, it facilitates the acylation reaction that charges the tRNA with the amino acid.[12]

Current Methodology for the Site-specific Incorporation of UAAs

These natural tRNA/aaRS pairs can be engineered to incorporate UAAs into proteins via a methodology developed by the Schultz lab for use in *E. coli*. The tRNA/aaRS pairs are developed to participate in amber suppression. Amber suppression occurs when the little-used amber stop codon (TAG) is

8

employed to replace the codon of a natural amino acid.[15] This codon is then used to site-specifically insert an unnatural amino acid via specific tRNA/aaRS pairs that correspond to the TAG codon.[15] An aaRS is selected after a double sieve selection (Figure 1.6). Mutant synthetases are evolved by randomly mutating four sites, thought to be involved in tyrosine bonding, in a natural tyrosine aaRS from *Methanococcus jannaschii* (*Mj*).[16] A library of mutant aaRS is developed by generating primers containing a with a random mutation at one of the 4 or 5 sites. This library of primers is then used in a mutagenic PCR with template plasmid of natural *Mj* aaRS.[17] This library of mutant synthetase plasmids is then transformed into *E. coli* cells via electroporation. Electroporation is used instead of heat shock because it affords the higher transformation efficiencies needed to ensure that each plasmid in the library is incorporated into a cell. The synthetase library is co-transformed with a plasmid
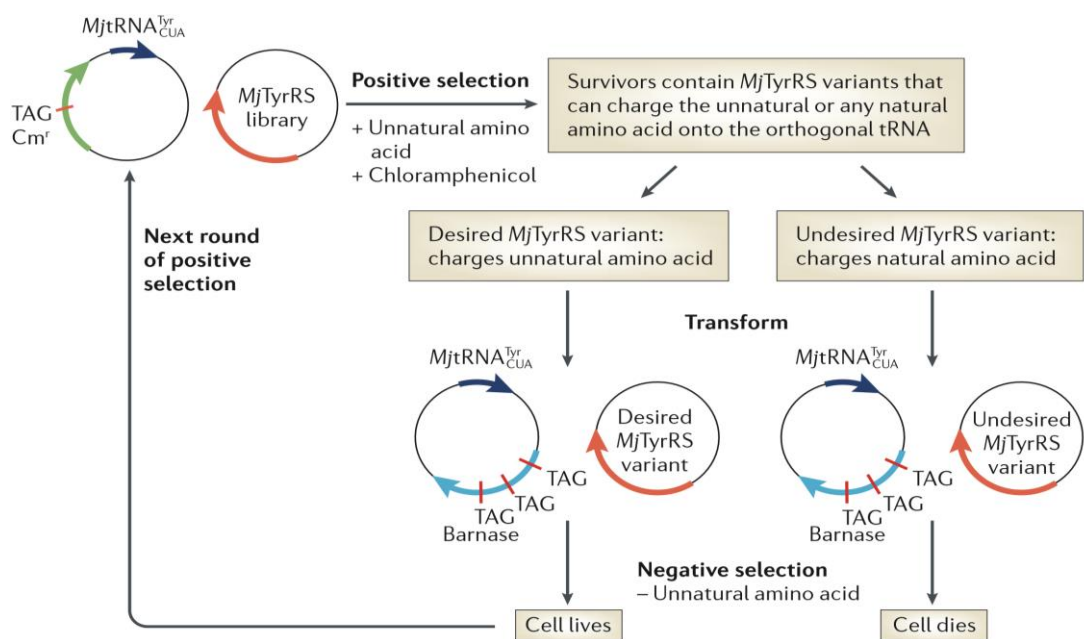
Figure 1.6: Positive and negative selection utilized to screen for a highly selective aaRS.[17]

containing a chloramphenicol resistance.[17] The gene contains a TAG codon in place of a tyrosine, so the cell must be able to read through the TAG to utilize the chloramphenicol resistance. This feature is utilized in the positive selection that takes place immediately after the transformation. The cells are plated on media containing the UAA of interest and chloramphenicol.[16,17] Any cells containing synthetase mutants who can charge either natural tyrosine or the desired UAA to the tRNA survive, and the synthetase plasmids are then put through a negative selection. In the negative selection, the mutant synthetase plasmid is transformed into new cells along with a plasmid containing a gene for barnase, a cytotoxin.[17] The barnase gene contains three TAG mutations. The cells are grown without the unnatural amino acid, so only the synthetases that charge the tRNA with natural tyrosine will read through the TAG mutations in the barnase gene.[17] These cells will produce barnase and die, leaving only the cells containing synthetases that only charge the unnatural. Several rounds of both the positive and negative selection occur to ensure that only the synthetases that are most selective for the unnatural survive.[16,17] To develop the tRNA, a tyrosine tRNA from *M. jannaschii* is mutated to correspond to the amber stop codon and 11 nucleotides that don't bind to the *Mj* aaRS were randomly mutated, then put through several negative selections using the barnase gene to ensure its selectivity for the evolved aaRS.[15,17] The resulting tRNA/aaRS pair is highly selective and therefore orthogonal to the native protein machinery of *E. coli*, ensuring that the UAA is reliably inserted into the protein.

To apply this tRNA/aaRS set, two pET plasmids can be transformed into electrically competent *E. coli* cells, one containing the TAG mutant protein gene, and the other containing the tRNA/aaRS genes (Figure 1.6).[1,15] The protein gene is typically under an isopropyl β-D-1-thiogalactopyranoside (IPTG) promoter with ampicillin (amp) resistance and the tRNA/aaRS genes are under an arabinose promoter with a chloramphenicol (chlor) resistance gene.[15] When the log-phase cells are grown in the presence of arabinose, IPTG, amp, chlor and the desired UAA, the orthogonal tRNA/aaRS pair charges the tRNA with the UAA of interest (Figure 1.7). This charged tRNA is then incorporated

Figure 1.7: Expression of UAA-containing GFP using the Schultz methodology.[1]

into the protein at a specific site in response to the amber stop codon via the *M. jannaschii* protein machinery.[15] As such, mutated protein can be expressed in mg/L yields.[18] This technology is extremely useful in determining the

11

structure-function relationships in proteins, and in conferring new functions to proteins, such as the ability to participate in bioorthogonal bioconjugations.

Unnatural amino acids have provided a new tool to optimize bioconjugation reactions. They allow for greater control of the reaction by offering biorthogonality and site specific incorporation, which increases the efficiency and utility of many applications of bioconjugates. Additionally, the Glaser-Hay reaction that can be accessed using UAAs creates the possibility for multivalent bioconjugation reactions, yielding the possibility of advances in medical, medicinal and biochemistry fields. Some applications of UAAs to bioconjugates are explored in this thesis, including protein immobilization, enzyme labeling, multivalent reactions and protein functionalization. UAAs are also explored in the context of chemical probes to examine the radical pathway of a multifunctional hemoglobin.

References

1. Maza, J. C.; Jacobs, T. H.; Uthappa, D. M.; Young, D. D. Employing Unnatural Amino Acids in the Preparation of Biooconjugates. *Synlett* **2016***, 27*, 805-813.

2. Allen, T. Ligand-targeted therapeutics in anticancer therapy. *Nature Reviews Cancer* **2002***, 2*, 750-763.

3. Sauer, U. Analytical Protein Microarrays: Advancements Towards Clinical Applications. *Sensors* **2017***, 17*, 256.

4. Linkoping University. Department of physics, Chemistry and Biology (IFM): Research. https://www.ifm.liu.se/applphys/molphys/research/contr_prot_orient/ (accessed Feb 28, 2016).

5. Chan, T.; Hilgraf, R.; Sharpless, K.; Fokin, V. Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.* **2004***, 6*, 2853-2855.

6. Hvilsted, S. Facile design of biomaterials by click' chemistry. *Polym. Int.* **2012***, 61*, 485-494.

7. Worrell, B. T.; Malik, J. A.; Fokin, V. V. Direct Evidence of a Dinuclear Copper Intermediate in Cu(I)-Catalyzed Azide-Alkyne Cycloadditions. *Science* **2013***, 340*, 457-460.

8. Lampkowski, J. S.; Villa, J. K.; Young, T. S.; Young, D. D. Development and Optimization of Glaser-Hay Bioconjugations. *Angew. Chem. -Int. Edit.* **2015***, 54*, 9343-9346.

9. Hay, A. Oxidative Coupling of Acetylenes. *J. Org. Chem.* **1960***, 25*, 1275-1276.

10. Vilhelmsen, M. H.; Jensen, J.; Tortzen, C. G.; Nielsen, M. B. The Glaser-Hay Reaction: Optimization and Scope Based on C-13 NMR Kinetics Experiments. *European Journal of Organic Chemistry* **2013**, 701-711.

11. Southwest Tennessee Community College. Burkett General Biology-Final Exam Chapters 14, 15 &18.

http://faculty.southwest.tn.edu/rburkett/DNA,_RNA,_&_Protein_Synthesis.html (accessed 7/28/2015)

12. Lodish, H.; Berk, A.; Kaiser, C. A.; Krieger, M.; Bretscher, A.; Ploegh, H.; Amon, A.; Scott, M. P. *Molecular Cell Biology,* 7th ed.; W. H. Freeman and Company: New York, 2013; pp 7-9, 131-135, 139-144.

13. Protein Biochemistry- Protein Synthesis. http://biochemistrycourse.blogspot.com/2012/12/protein-synthesis.html (accessed Mar 24, 2016).

14. Study Blue-University of Vermont CMB Week 2- Wallace. https://www.studyblue.com/notes/note/n/cmb-week-2-wallace-/deck/3509455 (accessed Mar 24, 2016).

15. Wang, L.; Schultz, P. A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* **2001***, 8*, 883-890.

16. Santoro, S.; Wang, L.; Herberich, B.; King, D.; Schultz, P. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol.* **2002***, 20*, 1044-1048.

17. Xie, J.; Schultz, P. G. Innovation: A chemical toolkit for proteins - an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006***, 7*, 775-782.

18. Noren, C.; Anthonycahill, S.; Griffith, M.; Schultz, P. A General-Method for Site-Specific Incorporation of Unnatural Amino-Acids into Proteins. *Science* **1989***, 244*, 182-188.

Chapter 2: Immobilization of GFP to Stabilizing Resin

Proteins are capable of catalyzing a wide range of chemical transformations, and their utility transcends the realm of biology, finding application in organic synthesis, materials chemistry, and alternative energy technologies.[1–3] However, a major limitation to exploiting the utility of proteins in a more chemical setting is their lack of stability and activity in nonaqueous environments.[2,4–10] The immobilization of proteins has proven to be an effective mechanism to expand their utility into nonbiological contexts. Moreover, immobilization confers the ability to recover and reuse relatively expensive enzymatic catalysts, and a better means of reaction control as catalysis can be halted by a simple filtration to remove the immobilized enzyme.[11] Immobilized enzymes have a wide range of applications including organic catalysis, microarrays, imaging probes, and biosensors.[12–17] While many general protocols already exist for protein immobilization (encapsulation, adsorption, and covalent attachment), there remains room for significant improvement.[18,19] For many uses of proteins, covalent immobilization provides a number of advantages including increased thermal stability, better scale up potential, and decreased probability of protein desorption, contaminating the reaction solution. However, many covalent modification approaches rely upon reactive residues (e.g., lysine and cysteine) on protein surfaces for attachment.[20] This

can lead to alteration in protein selectivity and activity, and may lead to complete inactivation as the active site may be blocked by the solid support.[19,21] Additionally, there is a general lack of control over how many residues react with the surface and remaining unreacted functionality on the surface, which may impede the catalysis via reaction with substrates and produce undesired heterogeneous mixtures on the solid support.[11] Many of the above problems can be solved by unnatural amino acid (UAA) directed immobilization of the protein of interest.

We set out to address issues associated with nonspecific protein immobilization through the use of unnatural amino acids. The site-specific incorporation of UAAs represents a novel mechanism to control protein immobilization via the introduction of a single residue that displays bioorthogonal reactivity. [22,23] The Schultz method for incorporating UAAs is advantageous over other mechanisms to incorporate UAAs, as the technology abrogates issues associated with protein size, scale, and delivery, as well as establishes a permanent, self-sustained system for protein production (not requiring the synthesis of precursors for each protein production).[24−27] The Schultz methodology has been utilized in the reaction of calmodulin onto carbon nanotubes and nanoparticles via the Staudinger-Bertozzi ligation,[31,32] as well as the immobilization of other proteins to M-20 Dynabeads in a cell-free system and indirectly via streptavidin derivatized gold surfaces with a click reaction.[16,17] Finally, a similar approach was performed using an agarose resin and azido

16

UAA to "catch" proteins for secretome MS analysis; however, in this case the UAA was not site-specifically incorporated into the pool of proteins.[33] The combination of these studies demonstrates the utility of site-specific immobilization, and the stabilizing benefits of the solid support; however, a more extensive investigation of immobilization methods and variables has the potential to advance this promising field.

Results and Discussion
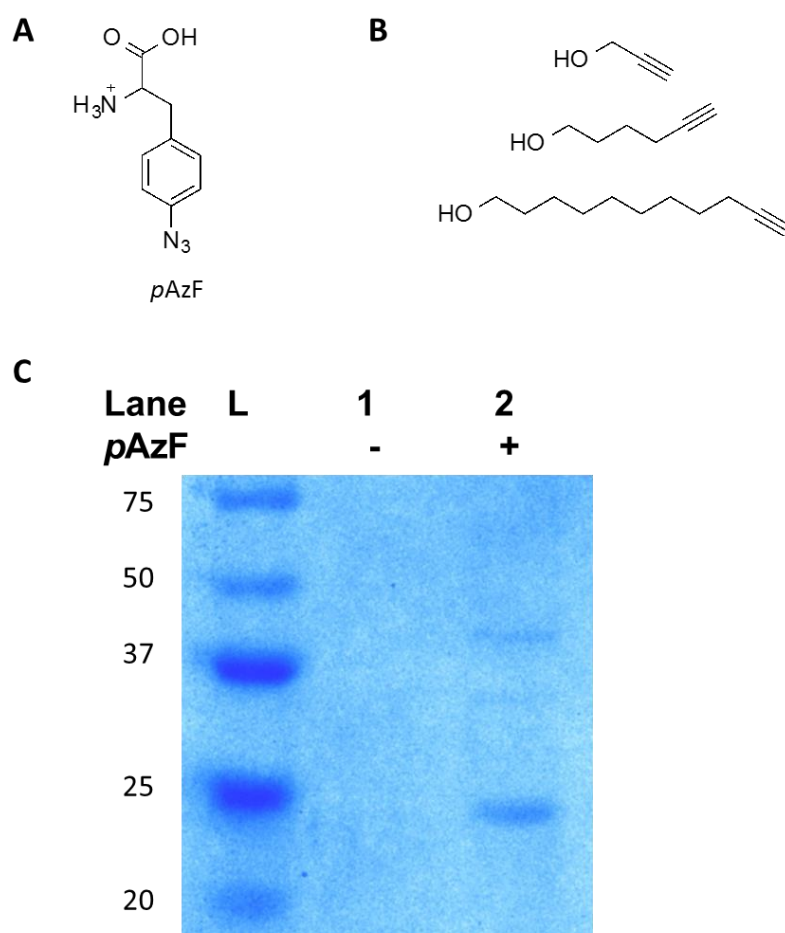
Figure 2.1: Protein and resin preparation. (A) p-Azidophenylalanine (pAzF) unnatural amino acid possessing an azide functionality for click conjugations. (B) Series of alkynols used for resin derivitization, varied methylene units confer differential flexibility for protein attachment. (C) SDS-PAGE analysis of pAzF-GFP-151 expression. No protein is present in the absence of the UAA (lane 1) suggesting successful incorporation of the UAA (lane 2).

Moreover, development of a generalized protocol and better understanding of immobilization conditions is needed to adapt unnatural amino acid technologies to protein immobilizations, and fully exploit this powerful system. Specifically, the p-

azidophenylalanine (*p*AzF) UAA was selected due to its ability to participate in bioorthogonal 1,3-dipolar cylcoaddition "click" reactions with an alkyne (Figure 2.1A). This reaction has found wide use in biological chemistry due to its unique reactivity (relative to the natural 20 amino acids), water compatibility, and overall efficiency.[34,35] Fortuitously, an aaRS has already been evolved to facilitate the incorporation of this UAA.[36,37] In order to rapidly and easily assess the protein immobilization, green fluorescent protein (GFP) was selected as a model protein, and expressed with *p*AzF at surface exposed residue 151
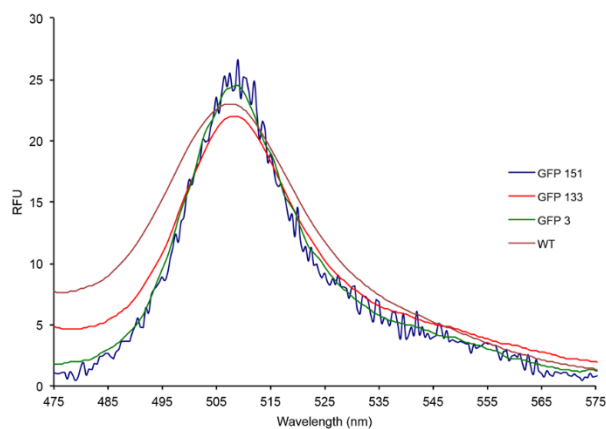
Figure 2.2: Fluorescence spectra of *p*AzF-GFP mutants and wild type protein (Ex. 395 nm).

using a pET-GFPY151TAG plasmid (Figure 2.1C). The mutated GFP containing the UAA retained its fluorescence relative to wild-type protein, and served as a functional reporter for the immobilization. Furthermore, as the fluorescence of GFP is not dependent on the orientation of immobilization (Figure 2.2), GFP was able to serve as a good model protein for the quick assessment of immobilization efficiency when different orientations were considered. It is important to note that this orientational independence does not necessarily translate to enzymes, and additional considerations must be made when adapting the approach.

Initially, attempts were made to ascertain the best material for the solid support. A range of supports was assessed including polystyrene, Tenta-Gel, and Sepharose resins. The solid support was evaluated based on immobilization efficiency, and retention of GFP fluorescence. As a proof-of concept, a representative of each class of support was examined. A tritylchloride modified polystyrene resin, an epoxide derivatized Sepharose 6B, and a carboxy-modifed Tenta-Gel resin were all derivatized with an alkynyl linker to facilitate click chemistry with the azide on the protein surface. Conveniently, each resin is reactive with alcohol functionalities under appropriate conditions, and each was subjected to immobilization with propargyl alcohol, 1-hexynol, or 1-undecynol to vary the length of the carbon tether separating the reactive alkyne from the solid-support (Figure 2.1B). These linkers were selected due to their low cost and commercial availability. Moreover, the linkers provided a mechanism for the assessment of linker length and immobilization efficiency. Alkyne immobilization and approximate resin loading was determined by cleavage of 15 mg of resin followed by GC/MS and TLC analysis, affording loadings of 0.6−0.9 mmol/g of immobilized alkyne. Resin loadings at these concentrations should ideally be low enough to adequately provide enough spatial protein separation to retain function; however, lower loadings may be necessary if protein function is inhibited. Additionally, the immobilization capability of propargyl amine instead of an alkyn-ol was tested by immobilizing propargyl amine to the sepharose resin and subsequently reacting the loaded resin with *p*Azf-GFP- 151. Similarly, the ability

of an azide- derivatized sepharose to immobilize *p*-propargyloxyphenyalanine (*p*PrF)-GFP-151 was tested. Both variations of loading the sepharose resin resulted in GFP immobilization, but were never compared to the propargyl alcohol loading method.

After confirmation of immobilization the different resins were subjected to click conditions for 16 h at 4 °C. The amount of resin utilized was normalized to the corresponding loading in order to ensure 12 μmol of immobilized alkyne reacted with a normalized 35 μg of GFP protein. In order to prevent aggregation of both the polystyrene and Tenta-Gel resins 10% DMSO was added. To verify that DMSO did not affect protein stability fluorescence spectra were taken in both the presence and absence of 10% DMSO (Figure 2.3). Resins were then transferred to spin columns and washed 5 times with PBS buffer. The presence of GFP was assessed both visually on a
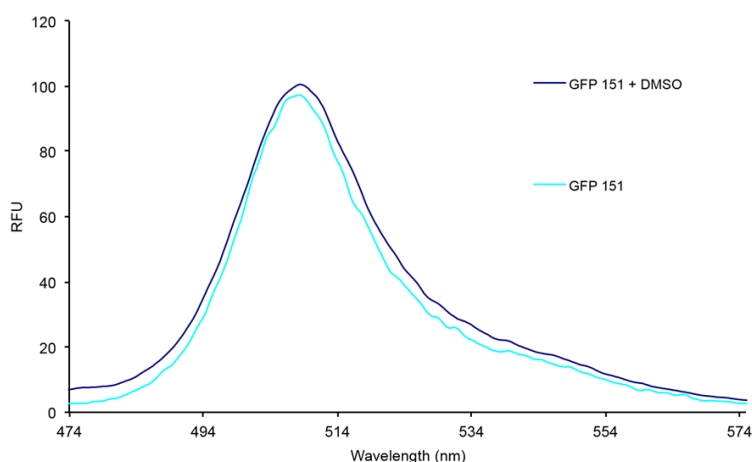
Figure 2.3: Fluorescence spectrum of *p*AzF-GFP-151 in the prescence and absence of 10% DMSO.

transilluminator and with a fluorescence plate reader (Ex. 482/Em. 520). While little to no fluorescence was observed on the polystyrene resin, definitive immobilization was noted on both the Tenta-Gel and sepharose resins, with the

20

fluorescence intensity being much greater on the sepharose resin. The decreased immobilization efficiency is most likely due to polystyrene resin aggregation in aqueous solvents. The immobilization of the GFP on the polystyrene and Tenta-Gel resins may be further improved via optimization of organic solvent ratios to accommodate the resin while not denaturing the protein of interest. However, these results do demonstrate that the GFP is capable of being immobilized on a solid support without affecting protein function, as fluorescence is maintained.
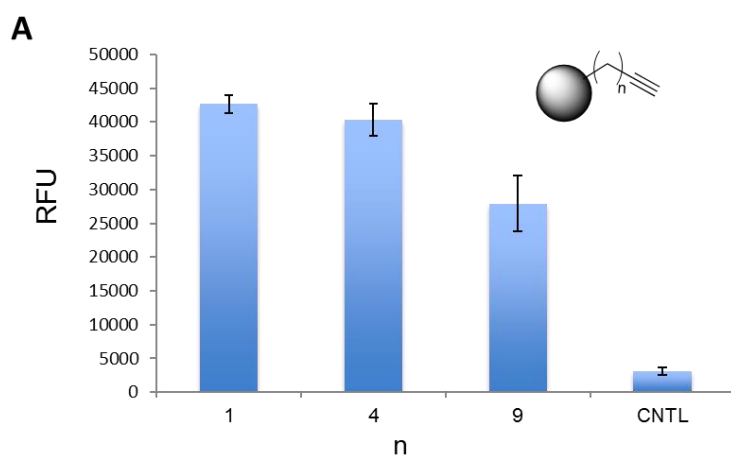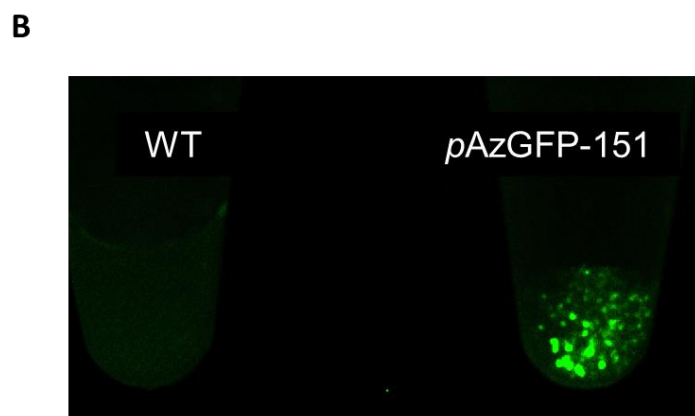
While all resins with all three alkynes were analyzed, the resin immobilized with propargyl alcohol reproducibly yielded the highest levels of GFP fluorescence with the sepharose resins

Figure 2.4: Resin and linker optimization. (A) Fluorescence data from alkyne derivatized Sepharose 6B resins conjugated with the pAzF GFP-151 mutant. Methylene units were varied with the alkyne substrate, and maximal fluorescence was observed with propargyl alcohol. A control resin using WT GFP (no azide functionality) demonstrated very little background fluorescence. All error bars represent standard deviations from three independent experiments. (B) Visualization of resin immobilization, as no fluorescence is observed under identical immobilization conditions using WT-GFP. Resin reacted with the pAzF-GFP-151 is notably fluorescent.

(Figure 2.4). Gratifyingly, resin incubated with wild-type GFP (containing no azide functionality) displayed little to no fluorescence, and conditions involving *p*AzF-GFP-151 in the absence of copper catalyst had similar results. This confirms that the click reaction was the site of immobilization and noncovalent interactions were not responsible for the immobilization event. Also, the absence of fluorescence indicates that the sepharose resin is not prone to noncovalent "stickiness" and thus only the desired protein can be immobilized. This confirms that the protein is only immobilized via the UAA in a controlled orientation, not in a random orientation via nonspecific interactions. Additionally, the effect of linker length could also be examined in the initial experiments. A trend was observed as decreased immobilization efficiency was correlated to increased methylene units, with undecynol affording the lowest GFP immobilization (Figure 2.4A). Similar results were observed even with increased immobilization times. It is possible that the increased flexibility of the linker leads to more conformational freedom that inhibits the proper orientation for the conjugation of the polymeric resin with the biopolymer.

22

With optimized immobilization conditions in hand, the propargyl alcohol derivatized sepharose resin was utilized to ascertain if UAA context is important for protein immobilization. A TAG mutation was introduced into the pET-GFP plasmid at residues 3,133, and 151, and the corresponding GFP mutants were expressed containing the pAzF UAA and diluted to identical concentrations (Figures 2.2 and 2.5). These residues were selected due to their differing protein context and variable rotational flexibility, while maintaining solvent exposure. To ensure that changes in fluorescence were due to immobilization and not altered spectra, the fluorescence spectra of all mutants were taken (Figure 2.2). The mutants were subjected to click conditions for 16 h at 4 °C as previously described. Visualization of the immobilization indicated that selection of UAA residue is important in the immobilization strategy (Figure 2.6). These results were confirmed by analysis of resin fluorescence on a plate reader.

Figure 2.5: SDS-PAGE analysis of expressed and purified GFP mutants containing pAzF at the indicated residue.
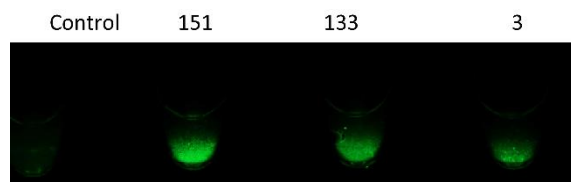
Figure 2.6: GFP immobilization on propargyl alcohol derivatized Sepharose 6B resin with various pAzF-GFP mutants. No detectable immobilization occurred under control conditions using a WT GFP protein without the azide functionality.

Residues 3 and 133 are contained within loop motifs at opposite ends of the β-barrel, while residue 151 is at the terminus of a β-sheet member of the β-barrel (Figure 2.7A). Due to the differential degree of immobilization despite identical conditions and GFP concentrations, there does seem to be a context effect for the placement of the UAA (Figure 2.7B). The highest degree of immobilization was observed at residue 151, followed by residues 133 and 3, respectively. These trends mimic those observed when selecting optimal linker flexibility with the most rigid residue possessing the highest immobilization efficiency. Residue 3 is at the N-terminus of the protein and in a very unstructured region, possessing a significant degree of conformational freedom. Conversely, residue 151 is extremely structured and very little rotational freedom can be achieved. This rigidity may be useful in facilitating the click reaction to the resin by holding the azide in the proper orientation. Residue 133 has intermediate conformational flexibility, as

Figure 2.7: UAA residue context dependence on immobilization. (A) Various GFP pAzF mutants were expressed with the UAA at different positions (highlighted in blue), adapted from PDB: 1EMA.39 (B) Fluorescence data demonstrating maximal immobilization with the pAzF-GFP-151 mutant, suggesting some importance of UAA context within the protein for the immobilization. All error bars represent standard deviations from three independent experiments.

well as intermediate immobilization yield. Again, little to no fluorescence was observed when the copper catalyst was removed from the reaction as a control. These results indicate that when utilizing UAA bioorthogonal handles for protein immobilization, some consideration must be given to the site of UAA incorporation to maximize immobilization efficiency as well as retain protein function.
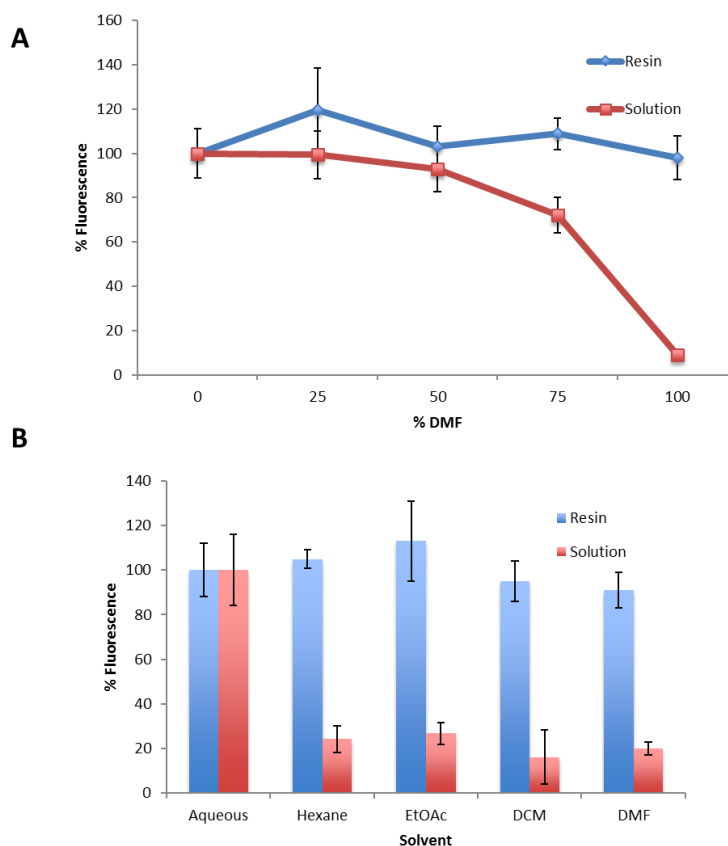
Figure 2.8: Resin stabilization of GFP in organic solutions. (A) Incubation of both immobilized GFP and GFP in solution with increasing DMF concentration demonstrates a decrease in fluorescence for only the solution phase GFP in increased organic solvent. (B) Incubation of both resin-bound GFP and solution phase GFP with different organic solvents leads to a marked difference in fluorescence values. All error bars represent standard deviations from three independent experiments.

In order to demonstrate the utility of resin immobilization, the retention of GFP fluorescence was examined when exposed to different organic solvents. Due to its miscibility, we first exposed the resin and free GFP-151 mutant to varied percentages of DMF (25−100%), and incubated the proteins for 2 h, the fluorescence was then measured on a plate reader as an indicator of functional protein. In agreement with literature precedence, GFP fluorescence decreases dramatically in solutions above 50% DMF; however, immobilized GFP fluorescence was retained, even at 100% DMF (Figure 2.8A).[38] Moreover,
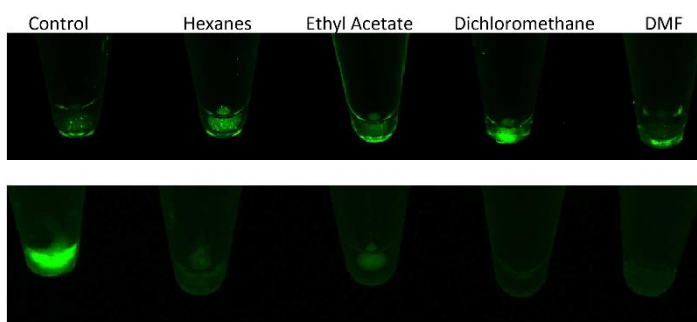
Figure 2.9: Non-aqueous solvent studies of immobilized GFP. Top set of tubes contained GFP immobilized on Sepharose 6B resin, resulting in retention of some fluorescence/protein function in all solvents. Bottom set of tubes contain solution phase GFP in the same solvents, demonstrating a loss of function in the non-aqueous media.

when incubated in other pure organic solvents including dichloromethane, ethyl acetate, and hexanes, a significant differential in fluorescence was observed and quantified on a BioRad gel imaging system (Figures 2.8B and 2.9). It should be noted that these observed increases in stability are not limited to the site-specifically immobilized GFP; however, the success of this method does demonstrate its potential usefulness with regard to proteins that are more sensitive to the orientation of immobilization. In addition, while some basic solvation sphere around the protein may exist, the protein remained stable on the resin even when subjected to 48 h incubation times and sonication, while solution based protein readily denatured. In summary, this increased protein

stability suggests this methodology can prove useful in the immobilization of a variety of proteins for use in nonaqueous catalysis, and coupled with the specificity of the immobilization technique, could afford heightened enzymatic efficiencies to current technologies.

These results will be extended to a more industrially relevant protein and the immobilization procedure will be further optimized using microwave technology. The enzyme used for these studies is a hyperthermophilic carboxylesterase of *Sulfolobus solfataricus*, a thermoacidophile found in volcanic muds.[40] This protein is abbreviated SSO. SSO provides a spatially constrained binding pocket and easily quantifiable chemical enzymatic activity. The position of the UAA will be optimized to yield maximum enzymatic function and stability in non-aqueous solvents. The high temperature tolerance of SSO carboxylesterase will also allow for experimentation with high temperature immobilization conditions that could shorten the immobilization reaction time. SSO is a more industrially relevant enzyme than GFP that will be used to optimize immobilization conditions for proteins that perform chemical transformations.

In conclusion, the utilization of unnatural amino acids has proven to be an effective mechanism for the immobilization of proteins. This has far-reaching implications toward the extension of protein function to nonbiological conditions. Moreover, this methodology is advantageous when compared to

existing techniques due to the high level of immobilization control, and its ability to generate covalently linked proteins in a homogeneous fashion.

Materials and Methods

Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Tritylchloride resin, 100200 mesh, 1% DVB crosslinking, and Tenta-Gel-COOH (130 μm beads) were purchased from Chem-Impex International. Epoxy-activated Sepharose 6B was obtained from GE Healthcare. Reactions were conducted under ambient atmosphere with solvents directly from the manufacturer. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit.

Immobilization of Alkynes onto Trityl Chloride Resin: To flame dried vial, trityl chloride resin (200 mg, 0.36 mmol, 1 equiv.) and dichloromethane (5 mL) were added. The resin was swelled at room temperature with gentle stirring for 15 min. Alkyn-ol (1.2 equiv.) was added to reaction, followed by triethylamine (0.2 equiv) for the propargyl alcohol reaction or 4 equiv of pyridine for hexyn-ol and undecyn-ol. The mixture was stirred at room temperature for 16 h. The resin was transferred to a syringe filter and washed with DCM and MeOH (5 alternating rinses with 5 mL each). The resin was swelled in CH2Cl2 and dried under vacuum for 45 min before further use. Resin (15 mg) was added to an Eppendorf tube followed by trifluoroacetic acid (2%, 200 μL). The mixture was shaken at room temperature for 1 h, then subjected to thin layer

chromatography in 1:1 Hexanes: Ethyl Acetate, GC/MS, and mass difference analysis to confirm resin loading.

Immobilization of Alkynes onto Tenta-Gel Carboxy Resin: To a vial, Tenta-gel carboxy resin (200 mg, 0.046 mmol, 1 equiv.) and dichloromethane (3 mL) were added. The resin was swelled at 30 °C for 15 minutes. Alkyn-ol (5 equiv.) was added to the swelled resin, followed by 4dimethylaminopyridine (DMAP, 0.2 equiv.) and 1-Ethyl-3-(3-dimethylaminopropyl)3-ethylcarbodiimide hydrochloride (ECDI, 5 equiv.). The reaction was stirred at room temperature for 16 h. The resin was transferred to a syringe filter and washed with dichloromethane and methanol (5 alternating washes with 3 mL each). The resin was swelled in DCM and dried under a vacuum for 45 minutes before further use. A tetrahydrofuran/methanol solution (4:1, 200 μL) was added to an Eppendorf tube along with potassium carbonate (25 mg) followed by resin (15 mg). The mixture was shaken for 16 h at 37 °C. The product was then subjected to thin layer chromatography, GC/MS, and mass difference analysis to confirm resin loading.

Immobilization of Alkynes onto Epoxy Sepharose: 6B Resin Epoxy-activated 6B Sepharose (GE Healthcare, 200 mg) was added to a filter syringe and washed with distilled water (5 washes, 3 mL). Alkyn-ol (700 μmol) and coupling buffer (3.5 mL, PBS at pH 13.0) was added to a 15 mL tube followed by the resin. The mixture was shaken at room temperature for 16 h. The resin was transferred to a filter syringe and washed with coupling buffer (4mL). The

sepharose was transferred to a 15 mL tube and capped with ethanolamine (3.5 mL). The resin was incubated at 30℃ for four hours then washed in a filter syringe with acetate buffer (10 mM) and tris-HCl buffer (pH 4 and pH 8 respectively, 3 alternating washes with 3 mL each). The same procedure was followed to immobilize propargyl amine and 3-azido-1-propargyl amine, except the coupling buffer consisted of PBS at pH 10. Additionally, 20 µL of 3-azido-1-propargyl amine was used instead of 700 µmol of alkyn-amine.

General GFP Expression: A pET-GFP-TAG variant plasmid (0.5 µL) was co-transformed with a pEVOL-*p*AzF plasmid (0.5 µL) into Escherichia coli BL21(DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (34 mg/mL) and ampicillin (50 mg/mL) at 37 ̊C overnight. One colony was then used to inoculate LB media (4 mL) containing both ampicillin and chloramphenicol. The culture was incubated at 37 ̊C overnight and used to inoculate an expression culture (10 mL LB media, 50 mg/mL Amp, 34 mg/mL Chlor) at an OD_{600} 0.1. The cultures were incubated at 37 ̊C to an OD_{600} between 0.6 and 0.8 at 600 nm, and protein expression was induced by addition of *p*AzF (100 L, 100 mM) and 20 % arabinose (10 L) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 10 L). The cultures were allowed to shake at 30 ̊C for 16-20 h then centrifuged at 5,000 rpm for 10 minutes and stored at -80 ̊C for 3 hours. The cell pellet was re-suspended using 500 µL of Bugbuster (Novagen) containing lysozyme, and incubated at 37 ̊C for 20 minutes. The

solution was transferred to an Eppendorf tube and centrifuged at 15,000 rpm for 10 minutes, then the supernatant was poured into an equilibrated His- pur Ni-NTA spin (Qiagen) column with of nickel resin (200 μL) and GFP was purified according to manufacturer's protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10% precast gels, 150V, 1.5h), and employed without further purification. Protein concentrations were determined both by a BCA assay and fluorescence measurements. The same protocol was used to express *p*PrF-GFP-151, with the substitution of a promiscuous pEVOL-Ambryx plasmid for the pEVOL-*p*AzF plasmid and the *p*PrF unnatural amino acid for the *p*AzfFunnatural amino acid.

General GFP Immobilization Conditions: Tris[(1-benzyl-1H-1,2,3-triazol-4yl)methyl]amine (5 mM, TBTA; 5 μL) was added to an Eppendorf tube along containing 4 mM tris(2-carboxyethyl)phosphine (TCEP; 5 μL) and 1 mM CuSO4 (6 μL). The catalyst system was diluted in phosphate buffered saline solution (PBS; 15 μL buffer) and 2.3 mg/mL of GFP-*p*AzF mutant (15 μL) was added, followed by 20 mg of derivatized Sepharose 6B resin (GE Healthcare). This mixture was incubated at 4 °C for 16 h then transferred to an empty spin column and washed with 1X PBS Buffer (5 x 200 μL) and centrifuged at 3,700 rpm for 60 s. The reactions were then imaged using a Bio-Rad gel imager (BioRad Molecular Imager Gel Doc XR+) and the fluorescence was taken by transferring resin to a 96-well plate and measuring relative fluorescence on a plate reader (Biotek Synergy HT Microplate Reader).

Non-Aqueous Solvent Studies: Dimethylformamide (DMF) solutions were prepared via dilution with PBS buffer, yielding 0%, 25%, 50%, 75%, and 100% solutions. To an eppendorf tube either 10 µL of *p*AzF-GFP 151 (2.3 mg/mL) or *p*AzF-GFP-151-Sepharose 6B resin was added followed by the appropriate DMF solution in triplicate. The tubes were mixed and incubated at room temperature for 2 h then transferred to a 96-well plate and recorded for fluorescence (Ex. 482/Em. 520). Fluorescence values were normalized to the solution phase or resin condition containing no DMF. Due to solvent incompatibility, incubations with dicholormethane, hexanes, and ethyl acetate were conducted with pure solvent under similar conditions; however, fluorescence was measured by densiometry using a BioRad gel imaging system.

References

1. Schmid, A., Dordick, J., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B. (2001) Industrial biocatalysis today and tomorrow. Nature 409, 258−268.

2. Comfort, D., Chhabra, S., Conners, S., Chou, C., Epting, K., Johnson, M., Jones, K., Sehgal, A., and Kelly, R. (2004) Strategic biocatalysis with hyperthermophilic enzymes. Green Chem. 6, 459− 465.

3. Montiel, C., Quintero, R., and Aburto, J. (2009) Petroleum biotechnology: Technology trends for the future. Afr. J. Biotechnol. 8, 7228−7240.

4. Sellek, G., and Chaudhuri, J. (1999) Biocatalysis in organic media using enzymes from extremophiles. Enzyme Microb. Technol. 25, 471− 482.

5. Cowan, D. (1997) Thermophilic proteins: Stability and function in aqueous and organic solvents. Comp. Biochem. Physiol. 118, 429− 438.

6. Ayala, M., Verdin, J., and Vazquez-Duhalt, R. (2007) The prospects for peroxidase-based biorefining of petroleum fuels. Biocatal. Biotransform. 25, 114−129.

7. Lee, M. Y., and Dordick, J. S. (2002) Enzyme activation for nonaqueous media. Curr. Opin. Biotechnol. 13, 376−84.

8. Cruz, J. C., Würges, K., Kramer, M., Pfromm, P. H., Rezac, M. E., and Czermak, P. (2011) Immobilization of enzymes on fumed silica nanoparticles for applications in nonaqueous media. Methods Mol. Biol. 743, 147−60.

9. Zaks, A., and Klibanov, A. M. (1988) Enzymatic catalysis in nonaqueous solvents. J. Biol. Chem. 263, 3194−201.

10. Krishna, S. H. (2002) Developments and trends in enzyme catalysis in nonconventional media. Biotechnol. Adv. 20, 239−67.

11. Brady, D., and Jordaan, J. (2009) Advances in enzyme immobilisation. Biotechnol. Lett. 31, 1639−1650.

12. Steen Redeker, E., Ta, D. T., Cortens, D., Billen, B., Guedens, W., and Adriaensens, P. (2013) Protein engineering for directed immobilization. Bioconjugate Chem. 24, 1761−77.

13. Jiang, X., Li, D., Xu, X., Ying, Y., Li, Y., Ye, Z., and Wang, J. (2008) Immunosensors for detection of pesticide residues. Biosens. Bioelectron. 23, 1577−87.

14. Vo-Dinh, T., and Cullum, B. (2000) Biosensors and biochips: advances in biological and medical diagnostics. Fresenius J. Anal. Chem. 366, 540−51.

15. Hernandez, K., and Fernandez-Lafuente, R. (2011) Control of protein immobilization: coupling immobilization and site-directed mutagenesis to improve biocatalyst or biosensor performance. Enzyme Microb. Technol. 48, 107−22.

16. Smith, M. T., Wu, J. C., Varner, C. T., and Bundy, B. C. (2013) Enhanced protein stability through minimally invasive, direct, covalent, and site-specific immobilization. Biotechnol. Prog. 29, 247−54.

17. Seo, M. H., Han, J., Jin, Z., Lee, D. W., Park, H. S., and Kim, H. S. (2011) Controlled and oriented immobilization of protein by sitespecific incorporation of unnatural amino acid. Anal. Chem. 83, 2841− 5.

18. Elgren, T., Zadvorny, O., Brecht, E., Douglas, T., Zorin, N., Maroney, M., and Peters, J. (2005) Immobilization of active hydrogenases by encapsulation in polymeric porous gels. Nano Lett. 5, 2085−2087.

19. Mateo, C., Palomo, J., Fernandez-Lorente, G., Guisan, J., and Fernandez-Lafuente, R. (2007) Improvement of enzyme activity, stability and selectivity via immobilization techniques. Enzyme Microb. Technol. 40, 1451−1463.

20. Wilchek, M., and Miron, T. (2003) Oriented versus random protein immobilization. J. Biochem. Biophys. Methods 55, 67−70.

21. Ulbrich, R., Schellenberger, A., and Damerau, W. (1986) Studies on the thermal inactivation of immobilized enzymes. Biotechnol. Bioeng. 28, 511−522.

22. Xie, J., and Schultz, P. G. (2006) A chemical toolkit for proteins–an expanded genetic code. Nat. Rev. Mol. Cell Biol. 7, 775− 82.

23. Wang, L., and Schultz, P. G. (2004) Expanding the genetic code. Angew. Chem., Int. Ed. Engl. 44, 34−66.

24. Yee, C. S., Seyedsayamdost, M. R., Chang, M. C. Y., Nocera, D. G., and Stubbe, J. (2003) Generation of the R2 subunit of ribonucleotide reductase by intein chemistry: insertion of 3-nitrotyrosine at residue 356 as a probe of the radical initiation process. Biochemistry 42, 14541−14552.

25. Seyedsayamdost, M. R., and Stubbe, J. (2006) Site-specific replacement of Y356 with 3,4-dihydroxyphenylalanine in the β2 subunit of E. coli ribonucleotide reductase. J. Am. Chem. Soc. 128, 2522−2523.

26. Seyedsayamdost, M. R., Yee, C. S., and Stubbe, J. (2007) Sitespecific incorporation of fluorotyrosines into the R2 subunit of E. coli ribonucleotide reductase by expressed protein ligation. Nat. Protoc. 2, 1225−1235.

27. Duffy, N., and Dougherty, D. (2010) Preparation of translationally competent tRNA by direct chemical acylation. Org. Lett. 12, 3776−3779.

28. Young, T. S., and Schultz, P. G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. J. Biol. Chem. 285, 11039− 11044.

29. Liu, C. C., Mack, A. V., Tsao, M. L., Mills, J. H., Lee, H. S., Choe, H., Farzan, M., Schultz, P. G., and Smider, V. V. (2008) Protein evolution with an expanded genetic code. Proc. Natl. Acad. Sci. U. S. A. 105, 17688−93.

30. Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering. Curr. Opin. Biotechnol. 14, 603− 609.

31. Yoshimura, S. H., Khan, S., Ohno, S., Yokogawa, T., Nishikawa, K., Hosoya, T., Maruyama, H., Nakayama, Y., and Takeyasu, K. (2012) Site-specific attachment of a protein to a carbon nanotube end without loss of protein function. Bioconjugate Chem. 23, 1488−93.

32. Ikeda-Boku, A., Kondo, K., Ohno, S., Yoshida, E., Yokogawa, T., Hayashi, N., and Nishikawa, K. (2013) Protein fishing using magnetic nanobeads containing calmodulin site-specifically immobilized via an azido group. J. Biochem. 154, 159−65.

33. Eichelbaum, K., Winter, M., Diaz, M., Herzig, S., and Krijgsveld, J. (2012) Selective enrichment of newly synthesized proteins for quantitative secretome analysis. Nat. Biotechnol. 30, 984.

34. Best, M. D. (2009) Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. Biochemistry 48, 6571−84.

35. Jewett, J. C., and Bertozzi, C. R. (2010) Cu-free click cycloaddition reactions in chemical biology. Chem. Soc. Rev. 39, 1272−9.

36. Deiters, A., Cropp, T. A., Summerer, D., Mukherji, M., and Schultz, P. G. (2004) Site-specific PEGylation of proteins containing unnatural amino acids. Bioorg. Med. Chem. Lett. 14, 5743−5.

37. Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. J. Am. Chem. Soc. 124, 9026−7.

38. Raghunathan, G., Sokalingam, S., Soundrarajan, N., Munussami, G., Madan, B., and Lee, S. (2013) A comparative study on the stability and structure of two different green fluorescent proteins in organic cosolvent systems. Biotechnol. Bioprocess Eng. 18, 342−349.

39. Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392−5.

40. Sehgal, A. C., Callen, W., Mathur, E. J., Short, J. M., and Kelly, R. M. (2001) Carboxylesterae from Sulfolobus solfataricus P1. Methods Enzym. 330, 461-471.

Chapter 3: Initial Steps Towards Carboxylesterase Immobilization

Enzymes are widely used in industrial applications to catalyze complex reactions in a highly efficient fashion, allowing manufacturers to streamline synthetic protocols. While enzymes can facilitate efficient syntheses, they also pose a unique set of challenges.[1,2] Enzymes are expensive and fragile reagents, requiring a specific pH, salt concentration and temperature to remain functional. They are also incompatible with organic solvents and detergents as these chemicals cause denaturation, which results in a loss of functionality. Enzyme immobilization offers a solution to these challenges, as it stabilizes proteins under denaturing conditions.[3-5] Additionally, immobilization offers the possibility of recovering expensive enzymes via simple filtration, which is both cost-efficient and environmentally friendly.

There are many established methodologies for immobilizing enzymes, including adsorption and covalent reaction with reactive natural amino acids.[3] The main weakness of these methods is that they afford no control over the orientation of the protein relative to the solid support. This can result in a heterogeneous layer of proteins on the solid support, where the active sites of some of the proteins are blocked from the solvent by the solid support (Figure 3.1).[6-8] This steric hindrance can prevent proteins from binding their substrate
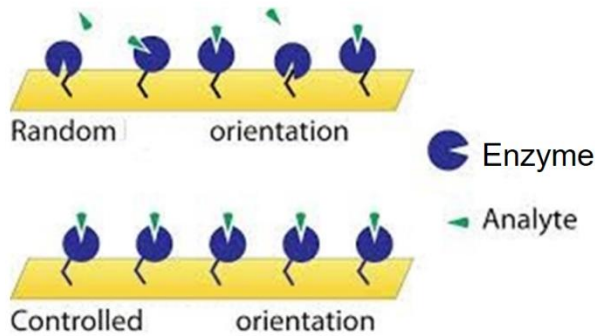
Figure 3.1: Two methodologies for enzyme immobilization. Utilization of natural reactive groups results in a heterogenous layer of randomly oriented enzymes, while immobilization using a site specifically incorporated UAA results in a homogenous layer of protein in a controlled orientation.[6]

and performing their enzymatic function, decreasing the overall catalytic capacity of the enzyme preparation. Unnatural amino acids (UAAs) can be used to solve this problem by offering precise control over the orientation of the protein relative to the solid support, as well as provide a homogenous immobilization. A surface exposed UAA can be site-specifically inserted into an enzyme with minimal to no impact on catalytic function. This UAA then provides a chemical handle that can be employed in bioorthogonal reactions to immobilize the enzyme onto a solid support in an orientation-controlled manner.

This immobilization method has been optimized in the previous chapter using GFP. In that research, we proved that it is feasible to utilize UAAs to immobilize a protein without interfering with its function. However, GFP is not
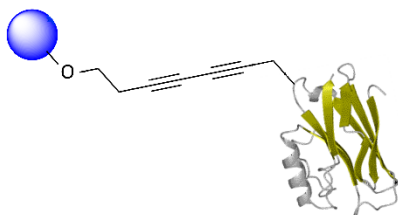
Figure 3.2: Immobilization of SSO Carboxylesterase P1 via a diyne linkage from the Glaser-Hay reaction.

an "enzyme," and thus we aim to expand the scope of this methodology to more interesting proteins. SSO Carboxylesterase P1 (SSO Est) was selected to serve as a model protein to extend the technology to an enzyme with measurable catalytic function. SSO Est is a hyperthermophilic enzyme that

39

cleaves an ester into an alcohol and an aldehyde.[9,10] This protein will provide a platform from which to extend the UAA-mediated immobilization technique to industrially relevant mesophilic and hyperthermophilic enzymes, as well as to optimize activation conditions for hyperthermophilic enzymes (Figure 3.2).

This enzyme affords several advantages over GFP and mesophilic enzymes for this stage of the immobilization optimization. Unlike GFP, SSO Est catalyzes an organic reaction, so the kinetics of the immobilized and wild type enzyme can be compared. This will allow for further optimization of the immobilization strategy to prevent perturbation of the catalytic function of the enzyme. Additionally, SSO Est is unusually stable and performs optimally at high temperatures.[9] This stability of this enzyme will make it easier to optimize immobilization and catalytic conditions as the enzyme is likely to maintain its activity in situations where more fragile mesophilic enzymes might not, such as microwave-assisted applications and reactions in organic solvents. These reaction conditions can then be extended to mesophilic proteins or other industrially useful hyperthermophilic proteins to yield solid supports with the maximum amount of loaded and active protein.

The added stability afforded by immobilization to SSO Est will also allow for optimization of activation conditions. Hyperthermophilic enzymes exhibit a low degree of conformational flexibility at normal physiological temperatures, which prevents destabilization of the non-covalent interactions holding the enzyme in its tightly folded state.[9] This rigidity also prevents hyperthermophilic

enzymes from performing their intended enzymatic activity at relatively low temperatures. To activate these enzymes for catalysis, the conformational flexibility must be increased. This is usually accomplished by increasing the temperature, but can also be accomplished by the addition of small amounts of denaturants or co-solvents to the solution at lower temperatures.[9] Microwave irradiation can also be used to activate hyperthermophilic enzymes at temperatures below their optimum temperature. The microwaves are thought to cause a rapid dipole alignment of the peptide bonds, which generates molecular motion.[11] This increase in molecular motion mimics the effect of high temperatures, causing the protein to become more flexible. Unfortunately, chemical or microwave activation can lead to denaturation of the enzymes.[11] UAA mediated immobilization will offer further stability to hyperthermophilic enzymes, preventing denaturation and expanding the range of activation conditions that can be utilized.

SSO Est will be used to further optimize a UAA mediated immobilization methodology that has been proven to stabilize GFP under denaturing conditions. Once the immobilization strategy is optimized for SSO Est, the methodology can be extended to other hyperthermophilic proteins as well as more sensitive mesophilic proteins that are utilized in industrial reactions. Immobilized SSO Est will also be used to optimize activation conditions for hyperthermophilic enzymes, which can them be extended to other

hyperthermophilic enzymes to maximize their catalytic efficiency at lower temperatures.

Results and Discussion

The wild type (WT) SSO Est plasmid was obtained from the laboratory of Robert Kelly at North Carolina State University. Three surface exposed tyrosine residues were selected for mutation based on the crystal structure. SSO Est Y90TAG (SSO 90), SSO Est Y116TAG (SSO 116) and SSO Est Y191TAG (SSO 191) mutants were prepared using Quik-Change PCR site directed mutagenesis. SSO 90 and SSO 191 reactions exhibited the most differential growth, compared to a negative control, under conditions using template plasmid concentrations of 24.7 ng/μL and 12.3 ng/μL. SSO 116 exhibited the most differential growth under conditions using template plasmid concentrations of 24.7 ng/μL and 6.2 ng/μL. Successful insertion of the TAG mutation was confirmed by sequencing at Genewiz using primers designed for the T5 promoter present in the pQE-30 backbone of the SSO Est plasmid. The insertion of the TAG mutation was successful in all the reactions containing 24.7 ng/μL of template plasmid, as well as in the SSO 116 reaction containing 6.2 ng/μL of template plasmid.

The mutant and wild-type SSO genes were encoded on a pQE-30 plasmid, which is typically not compatible for use in BL21 (DE3) cells. The pQE-30 plasmid contains a T5 promoter, which is recognized the general *E. coli* RNA polymerase.[12] The pET plasmids typically used within our labratory contain a T7 promoter, and the BL21 (DE3) cells contain a plasmid harboring an IPTG-
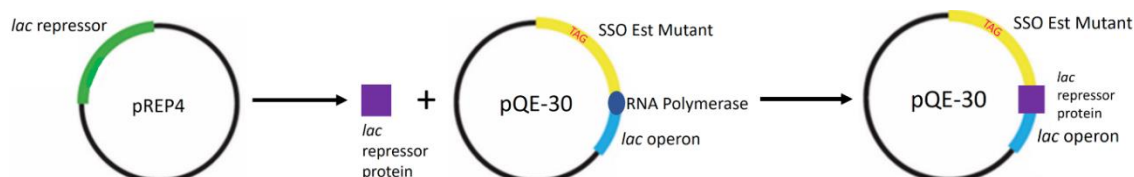
Figure 3.3: Cartoon of the pREP4 and pQE-30 system. The lac repressor protein synthesized from pREP4 binds to the lac operon on the pQE-30 plasmid to prevent expression of the mutant SSO protein.

inducible gene for the T7 RNA polymerase. These two plasmids create an inducible system that is orthogonal to the host protein expression machinery.[12,13] Because the pQE-30 plasmid does not contain an inducible RNA polymerase, but relies on the host RNA polymerase, constitutive expression of the gene of interest occurs. This system is not properly inducible since there are no repressors preventing expression of the gene of interest. To prevent this leaky gene expression and create an inducible system, a pREP4 plasmid can be co-transformed with the pQE-30 plasmid. Multiple pREP4 plasmids in each cell produce large amounts of *lac* repressor protein which prevent expression of the gene of interest (Figure 3).[13]

Initially, SSO WT was expressed with pREP4 by mixing the two plasmids together and transforming 1 μL of the solution into BL21 (DE3) cells. The SSO WT protein was expressed in good yield from this preparation (data not shown). Next, the SSO mutants were expressed containing the unnatural amino acid

*p*AzF without pREP4. The mutant proteins required the *Mj* synthetase and *Mj* tRNA to incorporate the UAA, so the SSO mutants could not be fully expressed before the synthetase system had been induced. Theoretically, the synthetase repressor/inducer system could act as repressors and inducers for the SSO mutants, preventing the expression of full length SSO until the arabinose was added. This theory was not supported by the SDS-PAGE of the expression. The yield of the mutant proteins was much lower when compared to the WT yield, confirming the need for the pREP4 plasmid in the mutant expressions.

Out of a desire to maintain pure stocks of the SSO plasmids while accommodating the need for the pREP4, SSO WT plasmid was co-transformed with pREP4 plasmid. This preparation was then expressed alongside a negative control to determine if this method would result in similar protein yield to the previous expression of SSO WT with pREP4. Transforming the SSO WT
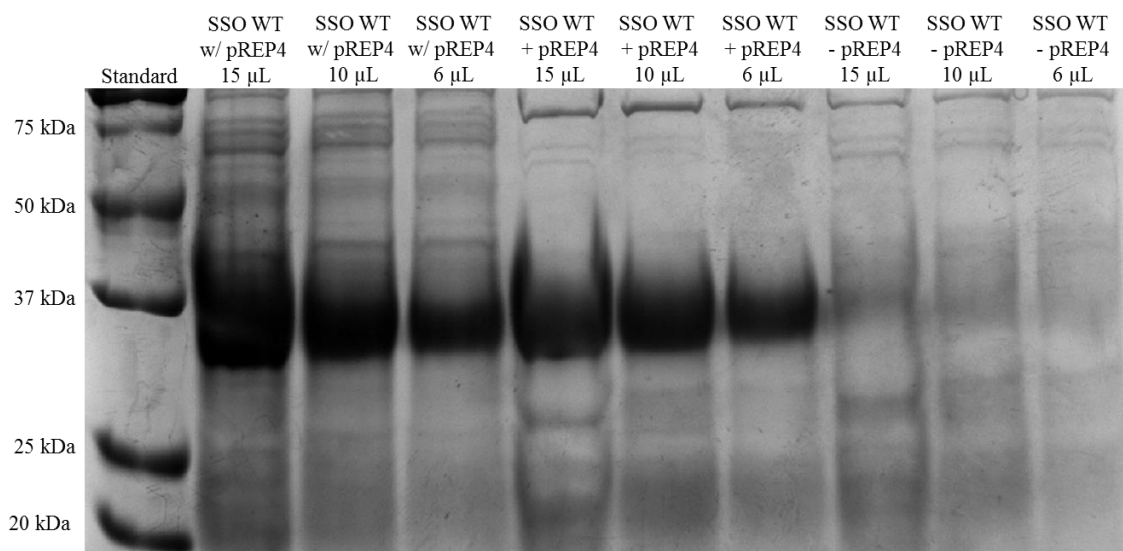
Figure 3.4: SSO WT expressed with pREP4 or without pREP4. The pREP4 was either mixed with the SSO WT before transformation into cells (w/ pREP4), co-transformed with the SSO WT (+ pREP4) or not transformed into the cells at all (-pREP4). Each sample was loaded at three different volumes (15, 10 or 6 μL) to obtain a clear band of protein.
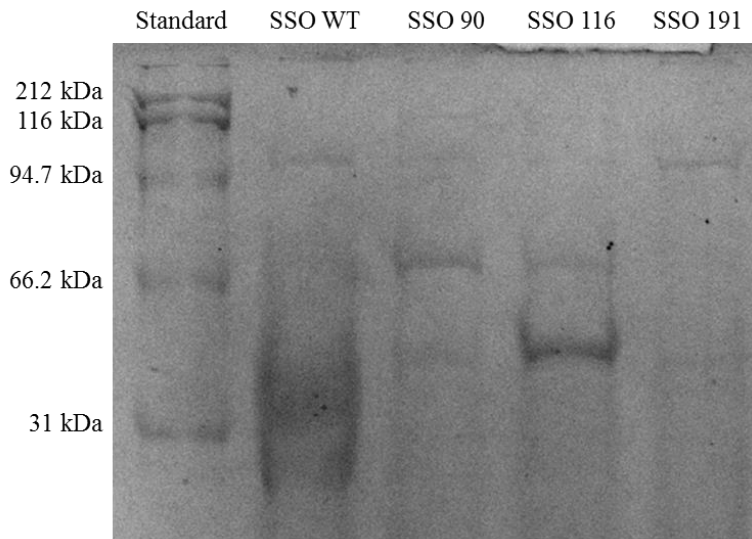
Figure 3.5: SSO WT and SSO mutants expressed with pREP4.

with the pREP4 resulted in identical yields to the previous preparation of SSO WT, while the expression of SSO WT without the pREP4 plasmid resulted in very little protein (Figure 3.4). This indicates that transforming the SSO plasmid with the pREP4 plasmid is a viable alternative to mixing the SSO and pREP4 plasmids prior to transforming. This result was then examined on the three SSO mutants, which were expressed by triple transformation with the pREP4, synthetase plasmid and the SSO mutant plasmid (Figure 3.5). Only the SSO 116 mutant seemed to exhibit an increase in yield, but it is difficult to tell without direct SDS-PAGE analysis of the two expressions. The SSO WT and mutant proteins were then subjected to a carboxylase assay as described by Park et al. to determine their kinetic constants. To date the reproducibility of these assays has been questionable, and current work is working on optimizing the assay using SSO WT.

Thus, we have expressed SSO Est containing surface exposed *p*AzF in three different locations. In the future, this assay will be optimized and kinetic

parameters for the wild type and each mutant will be obtained. Provided the SSO mutants kinetic constants are similar to those of the SSO WT, the mutants will be immobilized onto sepharose resin using a CuAAC or Glaser-Hay reaction as described in the previous chapter. The immobilized protein will then be assayed to determine its kinetic parameters. Provided the kinetic parameters are similar enough to ensure that the enzyme is still functioning the same way as the WT, further optimization of the immobilization reaction and activation conditions will occur using microwave technology.

Materials and Methods

General: The SSO P1 Carboxylesterase plasmid was obtain from the lab of Robert Kelley at NCSU. All PCR was performed a BioRad icycler, with the Quik-Change II kit. All assays were performed on a synergy HT microplate reader.

Quik-Change PCR Protocol: A pQE-30 plasmid harboring SSO Est P1 carboxyl esterase was diluted to 24.7 ng/µL, 12.3 ng/µL, and 6.2 ng/µL. The PCR reaction mixture contained one of the dilutions of the plasmids (1 µL), 10 mM DNTPs (1 µL), KAPA Hi-Fi Polymerase (1 unit/ µL, 1 µL), forward primer (10 µM, 1 µL), reverse primer (10 µM, 1 µL), KAPA buffer (5 µL), and Milliq water (40 µL). The reaction mixture was subjected to the following heating protocol: 95 ˚C (1 min); eighteen cycles of melting (95 ˚C, 30 s), annealing (55 ˚C, 30 s), and extension (68 ˚C, 6 min); a final extension period (68 ˚C, 6 min) and then held at 4 ˚C. Next the parent plasmid was digested and the mutant plasmid was ligated by adding 2 µL of DPN1 [20,000 units/µL], T4 Ligase buffer

(4 µL, 5x) and T4 Ligase (1 µL). The mixtures were then subjected to the following heating protocol: 37 ˚C 2 h and heat deactivation (80 ˚C, 15 min). The reaction was transformed (2 µL) into BL21 DE3 *Escherichia coli* cells by heat shock. The transformed cells were plated (500 µL) onto agar containing ampicillin and incubated at 37 ˚C overnight. The resulting cells were miniprepped using IBI High Speed Plasmid Mini Kit. The resulting plasmids were analyzed for successful insertion of the TAG mutation by sequencing at Genewiz. The primers were obtained from IDT DNA Technologies Inc. and are as follows: Y90TAG 5'-TGT AAT AGG CGA TGT GGA ATC TTA GGA CCC ATT ATG TAG AG-3' (forward) and 5'-CTC TAC ATA ATG GGT CCT AAG ATT CCA CAT CGC CTA TTA CA-3' (reverse), Y116TAG: 5'- CTA TAG GTT AGC TCC AGA ATA GAA GTT TCC TTC TGC AGT-3' (forward) and 5'-ACT GCA GAA GGA AAC TTC TAT TCT GGA GCT AAC CTA TAG-3' (reverse), Y191TAG 5'-CAA GAT CCA TGA TAG AGT AGT CTG ATG GGT TCT TCC T-3' (forward) and 5'-AGG AAG AAC CCA TCA GAC TAC TCT ATC ATG GAT CTT G-3' (reverse). The sequencing primer was 5'-TTC TGC TGA GCG GAT AAC-3'.

Expression and Purification of SSO: A pQE-30 plasmid harboring a variable SSO Est TAG mutant (0.33 µL) was co-transformed with a M15 pREP4 plasmid (0.33 µL) and a pEVOL-pCNf plasmid (0.33 µL) into BL21 (DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (0.034 mg/mL), ampicillin (0.05

mg/mL) and kanamycin (0.01 mg/mL) at 37 °C overnight. One colony was then used to inoculate LB media (4 mL) containing ampillicin, chloramphenicol and kanamycin. The culture was incubated at 37 °C overnight and used to inoculate an expression culture (10 mL LB media, 0.05 mg/mL Amp, 0.034 mg/mL Chlor, 0.01 mg/mL Kan) at an OD_{600} 0.1. The cultures were incubated at 37 °C to an OD_{600} between 0.6 and 0.8 at 600 nm, and protein expression was induced by addition of *p*AzF (100 L, 100 mM) and 20 % arabinose (10 L) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 10 L). The cultures were allowed to shake at 30 °C for 16-20 h then centrifuged at 5,000 rpm for 10 minutes and stored at -80 °C for 3 hours. The cell pellet was re-suspended using 500 μL of Bugbuster (Novagen) containing lysozyme and 200 μL Lysis buffer. The mixture was incubated at 37 °C for 20 minutes. The solution was transferred to an Eppendorf tube and centrifuged at 15,000 rpm for 10 minutes. The supernatant was then incubated at 80 °C for 10 minutes and centrifuged at 15,000 rpm for 10 minutes. Purified SSO Est was analyzed by SDS-PAGE (10%) and employed without further purification. Protein concentrations were determined by UV-Vis absorbance at 280 nm.

Arylesterase Assay: The reaction mixture was prepared in triplicate by adding 2 μL of SSO Est TAG variant protein to a mixture of 76 μL of 100 mM Acetate buffer, 4 μL of 4-nitrophenol-hexanoate (0.14 g/mL), and 18 μL of 1x PBS (pH 7.2). A control mixture was also made in triplicate, which did not contain protein. The control and reaction mixtures were incubated at 70 °C for 10, 20 or 30

minutes, then an absorbance reading was taken at 405 nm to measure the initial product formation. The reaction was quenched with 100 µL of distilled water saturated with Na_2CO_3 and another reading was taken at 405 nm.[4]
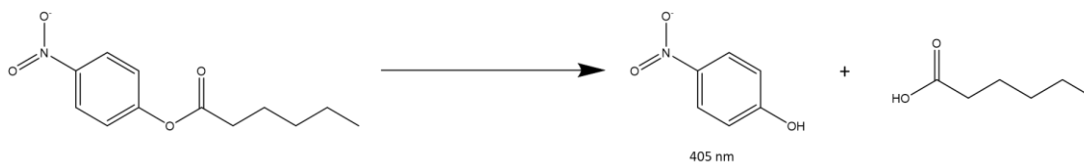
Scheme 3.1: Carboxylesterase assay reaction.

References

1. Madhu, A.; Chakraborty, J. N. Developments in application of enzymes for textile processing. *J. Clean. Prod.* **2017**, *145*, 114-133.

2. Porter, J. L.; Rusli, R. A.; Ollis, D. L. Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBioChem* **2016**, *17*, 197-203.

3. Brady, D.; Jordaan, J. Advances in enzyme immobilisation. *Biotechnol. Lett.* **2009**, *31*, 1639-1650.

4. Jiang, Y.; Zheng, P.; Li, D.; Zhou, L.; Tian, L.; Wang, J.; Yang, B.; Wang, X.; Zhang, X.; Gao, J. Enzyme-containing silica inverse opals prepared by using water-soluble colloidal crystal templates: Characterization and application. *Biochem. Eng. J.* **2016**, *112*, 123-129.

5. Lee, J.; Ko, J. H.; Lin, E.; Wallace, P.; Ruch, F.; Maynard, H. D. Trehalose hydrogels for stabilization of enzymes to heat. *Polym. Chem.* **2015**, *6*, 3443-3448.

6. Linkoping University. Department of physics, Chemistry and Biology (IFM): Research. https://www.ifm.liu.se/applphys/molphys/research/contr_prot_orient/ (accessed Feb 28, 2016).

7. Mateo, C., Palomo, J., Fernandez-Lorente, G., Guisan, J., and Fernandez-Lafuente, R. (2007) Improvement of enzyme activity, stability and selectivity via immobilization techniques. Enzyme Microb. Technol. 40, 1451−1463.

8. Ulbrich, R., Schellenberger, A., and Damerau, W. (1986) Studies on the thermal inactivation of immobilized enzymes. Biotechnol. Bioeng. 28, 511−522.

9. Sehgal, A.; Tompson, R.; Cavanagh, J.; Kelly, R. Structural and catalytic response to temperature and cosolvents of carboxylesterase EST1 from the extremely thermoacidophilic archaeon Sulfolobus solfataricus P1. *Biotechnol. Bioeng.* **2002***, 80*, 784-793.

10. Park, Y.; Yoon, S.; Lee, H. A Novel Thermostable Arylesterase from the Archaeon Sulfolobus solfataricus P1: Purification, Characterization, and Expression. *J. Bacteriol.* **2008***, 190*, 8086-8095.

11. Young, D. D.; Nichols, J.; Kelly, R. M.; Deiters, A. Microwave activation of enzymatic catalysis. *J. Am. Chem. Soc.* **2008***, 130*, 10048-+.

12. New England Biolabs Inc. FAQ: What is the difference between BL21 and BL21(DE3) competent *E.coli?*

https://www.neb.com/faqs/2016/01/11/what-is-the-difference-between-bl21-and-bl21-de3-competent-e-coli (accessed April 13, 2017).

13. Quiagen. The QIAexpressionist- (EN). https://www.qiagen.com/us/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en (accessed April 13, 2017).

Chapter 4: Initial Steps Towards a Novel Prostate Cancer Test

Prostate cancer (PCa) is an extremely common cancer in the male population. The most common method of detecting cancerous changes in the prostate is an assay that quantitates the concentration prostate-specific antigen (PSA) present in the blood stream. This test is somewhat unreliable, as it often yields false positives which necessitate the use of invasive biopsies to ultimately diagnose cancer.[1] A modified version of Ubiquitin like protease 1 (Ulp1) can be used to develop a more accurate and less invasive test for PCa. Ulp1 is a yeast protease that cleaves small ubiquitin-like modifier (SUMO) from other cellular proteins.[2,3] Cancerous cells are characterized by abnormal gene expression and unusual levels of post translational modifications, such as abnormally SUMOylated proteins.[4] For instance, a tumor suppressor gene known as reptin is highly SUMOylated in metastatic prostate cancer cells.[5] This difference in SUMOylation between healthy and cancerous cells provides a unique opportunity to develop a non-invasive method of screening patient samples for cancerous changes. This SUMO based assay would use cell-penetrating peptides (CPPs) to deliver a fluorescently labeled, non-catalytic form of Ulp1 (Utag) into patient cells, where the fluorescent Utag would bind to SUMO-protein conjugates. The patient cells could then be compared to control cells to assess the relative levels of protein SUMOylation. High amounts of

SUMOylation, indicated by localized fluorescence, would indicate a cancerous sample. This test could be performed on cells that are normally secreted from the prostate, allowing doctors a chance to screen men who received a false positive from the PSA test.

SUMO, like Ubiquitin (Ub), is a human protein that is involved in post-translational modification of other cellular proteins to regulate cellular pathways. Both Ub and SUMO are conjugated to other proteins via an isopeptide linkage between a lysine on the target protein and the C terminus of Ub or SUMO. Unlike Ub, SUMO is not primarily utilized to degrade proteins. For instance, the yeast analog of human SUMO (Smt3, Figure 4.1A and 4.1B, blue) is used to modulate cellular localization and transcription factor stability.[3]

Like Ub, SUMO is regulated by other cellular proteins. Ulp1 is one such protein, it cleaves the target protein lysine-SUMO/Smt3 C terminus isopeptide bond following the SUMO Gly-Gly motif.[2,3] The C-terminal region of Ulp1 is used to bind and cleave SUMO linked peptide bonds, and the N-terminus is utilized by the cell to regulate Ulp1 activity. The C-terminal region was cocrystallized by
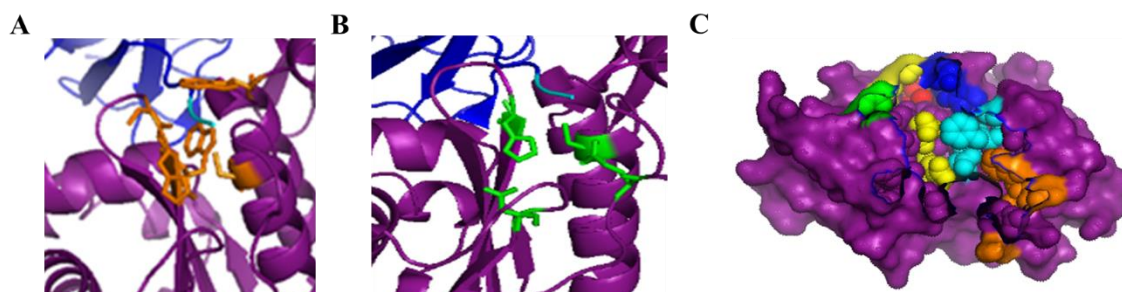
Figure 3.1: Ulp1 (purple) and Smt3 (blue) with the GG motif highlighted in cyan (A) the hydrophobic tunnel highlighted in orange (B) the catalytic triad highlighted in green (C) SUMO/Smt3 interacting motifs of Ulp1 highlighted in various colors.

Mossesova and Lima with Smt3, where the details of the Smt3-Ulp1 interface where revealed. Ulp1 maintains its selectivity for SUMO by recognizing the -GGX C terminal sequence (Figure 4.1A, cyan). This sequence must be threaded through a hydrophobic tunnel to access the active site. This tunnel is created by residues Trp-448 and Trp-515, with the catalytic triad at the end (Figure 4.1A, orange).[3] This tunnel interacts with the GGX sequence via van Der Waals (VDW) contacts between the α carbon of the glycines and the amino acids in the tunnel, except for the cysteine. Trp-448 acts as a clamp, holding down the GG motif. This GG recognition domain is relatively small when compared to those of Ub processing enzymes, which allows Ulp1 to cleave SUMO from large protein conjugates without affecting their structure.[3] Multiple non-covalent interactions, such as salt bridges and VDW contacts are employed in the six SUMO interaction motifs to keep SUMO and Ulp1 spatially close in order to facilitate recognition of the GG motif.

The catalytic function of Ulp1 is accomplished by a catalytic triad (Figure 4.1B, green) where Cys-580 is coordinated to a basic His-514 which is stabilized by Asp-531. Gln-574 Nε and Cys-580 N (Figure 4.1B, green) stabilize the transition state by providing an oxyanion hole.[3] A truncated version of Ulp1, known as Utag, lacks C-580 and is not capable of cleaving SUMO from other proteins. This lack of catalytic activity makes Utag the ideal protein for the development of a SUMO-based cancer diagnostic.

There are two main challenges in developing this novel diagnostic. One, the Utag must be easily detectable *in vivo*, and two, the Utag must be able to enter cells. The first challenge can be overcome by expressing Utag containing a surface-exposed UAA. The UAA containing Utag could then be subjected to bioconjugation reactions to couple it to a fluorophore. This chapter will focus on the optimization of the UAA containing Utag expression and subsequent fluorophore conjugation. Two isoforms of Utag, one from *S. cerevisiae* and one form *Kluyveromyces marxianus* (*Km*), will be utilized in this optimization. The *Km* Utag is a heat stable mutant, and so it is a few amino acids smaller and folded more tightly that the *Sc* Utag. Both the *Sc* Utag and the *Km* Utag will be expressed as fusion proteins with maltose- binding protein (MBP), as this fusion helps to stabilize the expressed Utag. The mechanism of importing the Utag complex will be discussed briefly in the conclusion, as this mechanism will be studied and optimized in the future.

Results and Discussion

Surface exposed tyrosine residues on *S. cerevisiae* Utag were selected for mutation. *S. cerevisiae* Utag Y147TAG (*Sc* 147), *S. cerevisiae* Utag Y241TAG (*Sc* 241), *S. cerevisiae* Utag Y253TAG (*Sc* 253) mutants were prepared using Quik-Change PCR site directed mutagenesis. *K. marxianus* Utag Y576TAG (*Km* 576) and *K. marxianus* Utag Y147TAG (Km 147) were prepared using Q5 PCR. *Sc* 147 exhibited the most differential growth relative to a negative control under conditions using a template plasmid concentration of 12.3 ng/μL and an

annealing temperature of 68 °C. *Sc* 241 and *Sc* 253 did not exhibit differential growth when subjected to the concentration gradient QuikChange, but did exhibit differential growth after the annealing temperature gradient QuikChange. *Sc* 241 exhibited the most differential growth at 53 °C and 55 °C and *Sc* 253 exhibited the best differential growth at 53 °C and 61 °C. The *Sc* 241 and *Sc* 253 plasmids will be sequenced in the future for successful insertion of the TAG mutation.

Wild type (WT) Utag containing the TAG mutation was expressed in duplicate using pEVOL WT at 30 °C. This pEVOL WT plasmid incorporates a standard tyrosine residue in place of the TAG mutation. Several conditions were examined, including one where the log-phase cells were pelleted and resuspended in one-fifth the original volume of media prior to adding IPTG to induce Utag expression. Another expression was treated normally, without
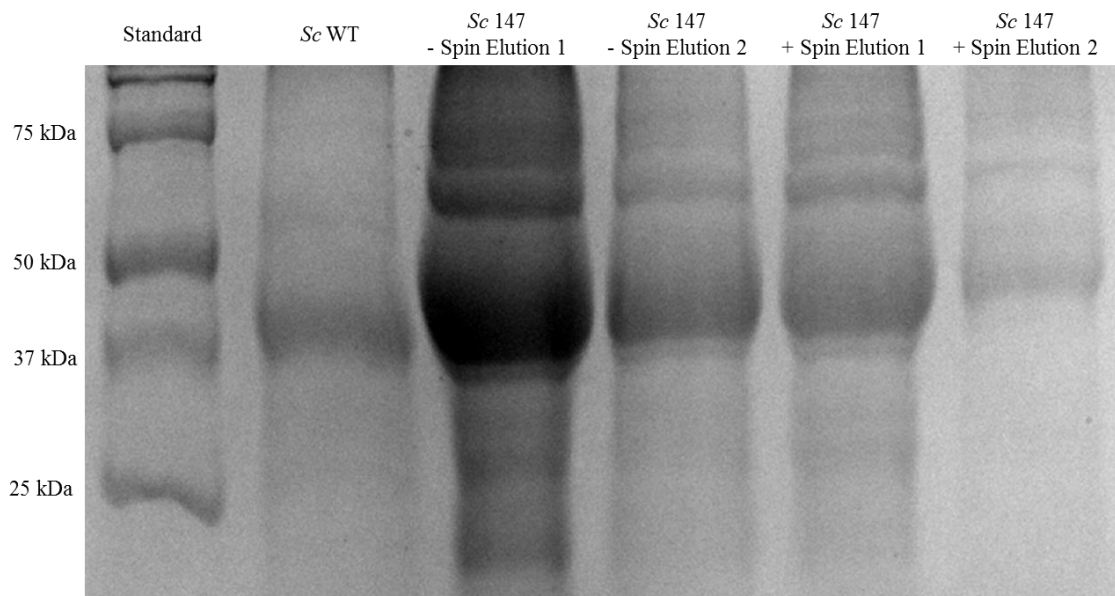
Figure 4.2: Expression levels of Sc Utag 147 using pEVOL WT as the synthetase when the induction media is concentrated (+ Spin) or not concentrated (- Spin).

pelleting the cells. This was performed to optimize the expression and assess if the concentration protocol could be employed to minimize the amount of UAA needed for each expression. The pelleted expression yielded less protein than the non-concentrated expression, so the induction solution was not concentrated during expression of Utag mutant proteins from this point on (Figure 4.2). Having optimized the expression conditions, *Sc* Utag 147 and *Km* Utag 147 was expressed containing *para*-azidophenylalanine (*p*AzF, **1**) using pEVOL-pCNF. The resulting protein was subjected to a click reaction with an alkyne-linked fluorophore, alongside a GFP positive control and a Utag WT negative control (Scheme 4.1). The reaction was analyzed by SDS-PAGE. No fluorescence was observed, but the proteins did stain with coomassie blue,
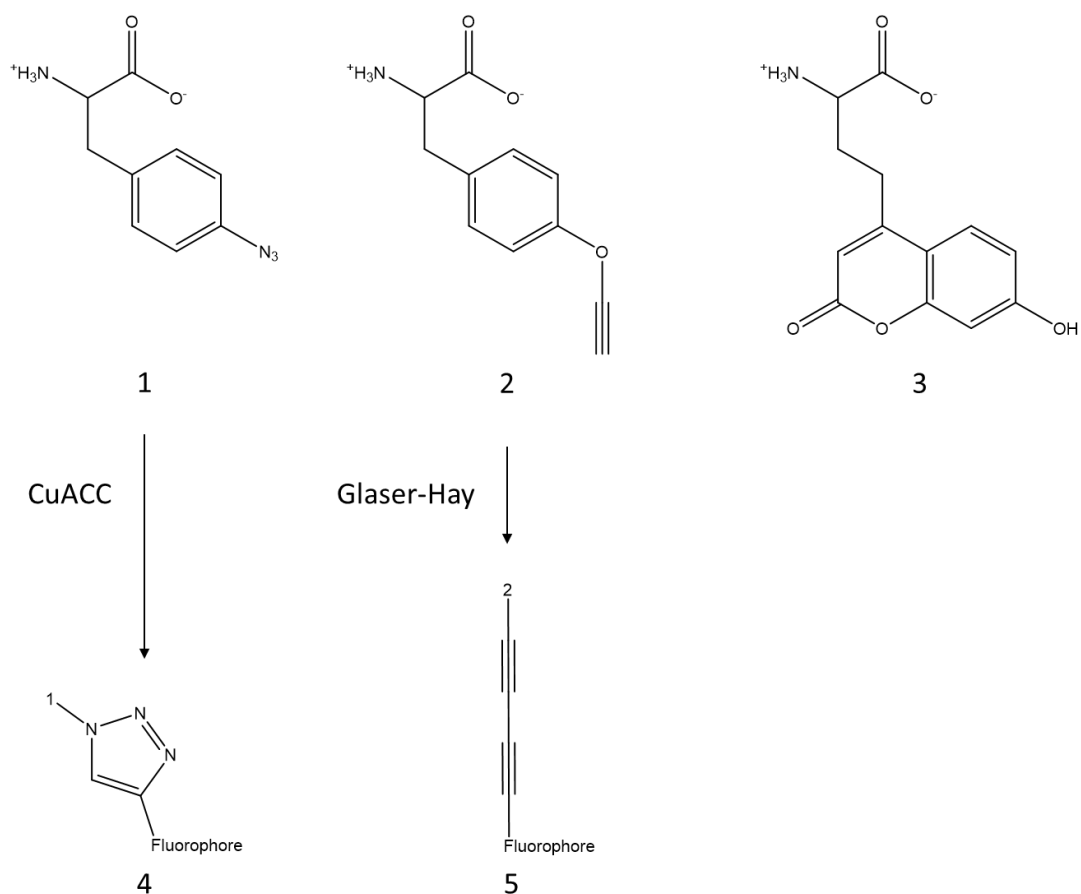
Scheme 4.1: Structure of the unnatural amino acids *p*AzF (**1**), *p*PrF (**2**) and coumarin (**3**) along with the products of the reactions they are utilized in. **1** is utilized in the CuAAC reaction to yield a 1,4 linked 1,2,3-triazole (**4**) and **2** is utilized in the Glaser-Hay reaction to yield a diyne (**5**).

indicating that the reaction did not work. After a several unsuccessful attempts to perform the click reaction, the *Sc* 147 and *Km* 147 were then expressed with *para*-propargyloxyphenylalanine (*p*PrF, **2**) instead of *p*AzF. The Utag + **2** mutants were then successfully reacted with the alkyne fluorophore via a Glaser-Hay reaction (Figure 4.3). The presence of fluorescent bands at the weight of the protein indicate that the fluorophore is covalently bound to the protein at the *p*PrF residue, confirming the successful incorporation of the UAA. The slight fluorescence in the Utag WT negative control lane is likely due to non-covalent interactions that have trapped a small amount of the fluorophore

with the protein. The lack of any fluorescence or protein in the GFP positive

control lane is unusual, as this Glaser-Hay reaction was optimized for GFP, but
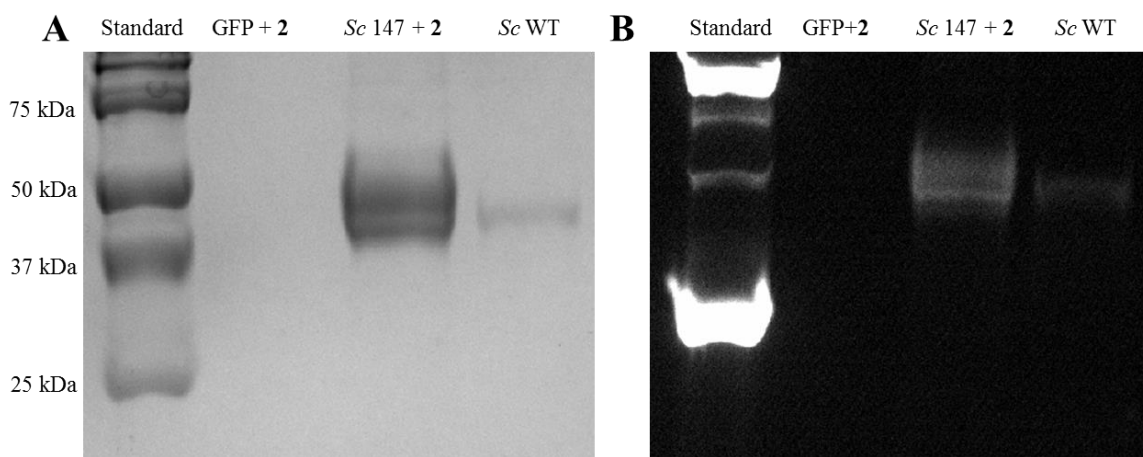
Figure 4.3: Glaser-Hay reaction between alkyne fluorophore and either GFP + **2**, Sc 147 + **2**, or Sc WT. (A) SDS-PAGE of the Glaser-Hay stained with coomassie and (B) the fluorescent scan of the SDS-PAGE, where the presence of a fluorescent band indicates successful conjugation between the fluorophore and the protein.

it doesn't necessarily indicate that the Glaser-Hay reaction was unsuccessful.

Instead it appears based on the coomassie stained gel that there was no

appreciable GFP in the reaction to be labeled.

This fluorescently labeled Utag mutant was then assayed for SUMO

binding ability via three types of pull-down assay. First, the fluorescently labeled

Sc Utag 147 was incubated with a previously prepared Western blot membrane

that had SUMO bound. The membrane was then analyzed by UV-irradiation to

detect the fluorophore; however, no detectable fluorescence was observed.

This may be due to the low concentration of the fluorescently labeled *Sc* Utag

147 in the loading solution.

The second pull-down assay utilized agarose beads with SUMO

immobilized on the surface. These beads were incubated with the WT Utag and

fluorescently labeled *Sc* Utag 147 to allow the Utag to bind to the SUMO. The beads were washed and imaged on a Biorad gel imager in eppendorf tubes with 20 µL of PBS and in a 96-well plate in a Synergy plate reader with 70 µL of PBS. Neither imaging system showed a difference in fluorescence between the wild type and the fluorescently labeled *Sc* Utag 147. Once again, this is likely because there wasn't enough fluorescently labeled *Sc* Utag 147 bound to visualize or detect in such a dilute solution.

Finally, a plate-based pull-down assay was performed by immobilizing SUMO onto the bottom of a 96-well plate using poly-L-lysine, then allowing the negative control Utag WT, the fluorescently labeled *Sc* Utag 147 and fluorescently labeled *Km* fluorescently labeled Utag 147 to bind to the SUMO. The plate was then read at 485 nm to determine the presence of the fluorophore. The WT Utag and the fluorescently labelled mutants did not exhibit any difference in fluorescence, indicating that this assay was not successful in binding the either mutant of the fluorescently labeled Utag 147.

The bead based assay was then adapted to be used with SDS-PAGE, as the SDS-PAGE would make it possible to visualize both the wild type and the mutant protein, giving us positive control for the binding assay. The assay was performed as before with WT Utag, fluorescently labeled *Sc* 147 and fluorescently labeled *Km* 147. Instead of analyzing the assay via fluorescence, the Utag was cleaved from the bead by denaturing it with 0.2% SDS. The resulting SDS-protein solution was then analyzed for fluorescence in a 96-well

plate. This did not yield any difference in fluorescence, so the protein was further analyzed by SDS-PAGE. Neither the WT nor either of the mutants bound to the SUMO, possibly because the fluorophore's bulk or Glaser-Hay conditions were interfering with the folding of the protein. Since the WT protein was also subjected to the Glaser-Hay, the reaction conditions may have interfered with its ability to bind as well. To alleviate any decrease in binding ability originating from the Glaser-Hay, the *Km* 147 was expressed containing a fluorescent amino acid (coumarin, **3**) with pEVOL-Cou (Figure 4.4). This mutant protein

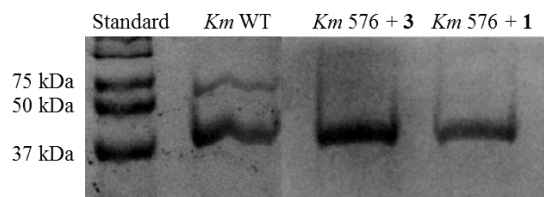Figure 4.4: Expression of *Km* 576 + **3** and *Km* 576 + **1** along with a *Km* WT control. One lane was omitted from this gel for clarity.

also didn't bind to SUMO, leading to the conclusion that UAA incorporation at this position interfered with the SUMO binding ability of the mutants. In light of this conclusion, the Y147TAG mutants were not used again and further experiments focused on the *Km* Y576TAG mutant.

61

Km 576 was expressed with **2** and subjected to a Glaser-Hay coupling with an alkyne fluorophore. Unfortunately, this Utag mutant did not bind SUMO either, as evidenced by the lack of protein in the elution lane. The fluorescently labeled *Km* 576 didn't seem to bind to SUMO at all, as the bead loading solution contained a substantial amount of protein (Figure 4.5). To test if the Glaser-Hay conditions or addition of the fluorophore was interfering with the ability of
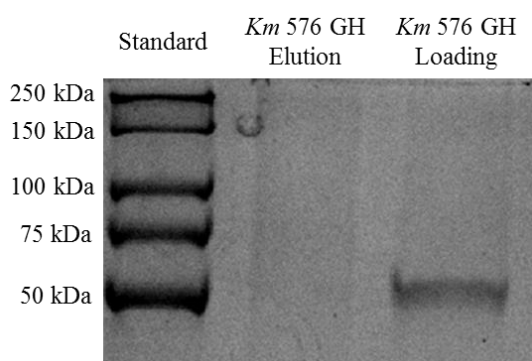
Figure 4.5: SUMO binding assay of Km 576 + **2** conjugated to fluorophore via a Glaser-Hay (GH) reaction. No protein eluted from the SUMO beads, and a significant amount of protein was left in the bead loading solution. This indicates that the Utag did not bind to the immobilized SUMO.
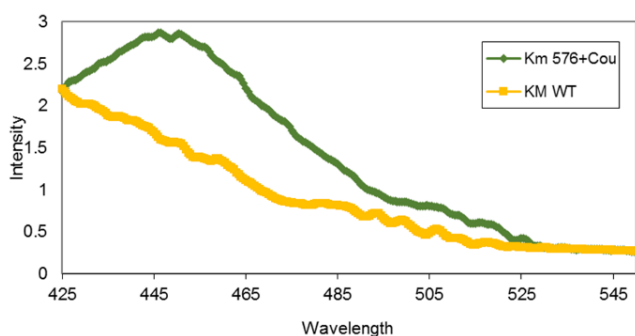
Figure 4.6: Fluorescence Spectra of *Km* WT and *Km* 576 + **3**.

Km 576 to bind SUMO, *Km* 576 was expressed with **3** and **1**. Wild type *Km* Utag was also expressed as a positive control. The coumarin incorporation was confirmed by fluorescence analysis in a fluorimeter. The presence of coumarin is detectable by a peak at 450 nm, as is seen in the Km 576 + **3** mutant, indicating that the UAA incorporation was successful (Figure 4.6).[6] The expressed proteins were then assayed for their ability to bind SUMO (Figure 4.7). Only the Utag WT

bound SUMO, as indicated by the presence of a band in the elution lane. No protein was detected in the elution lanes of the mutants, meaning that the mutants are incapable of binding SUMO. Different volumes of each protein

were also run on the gel fluorescently labeled to quantitate the amount of protein that bound to the SUMO. The 25% lanes were loaded with one-quarter the volume of each protein added to the bead loading solutions, and the 50% lanes were loaded with half the volume of the specific protein added to the bead loading solution. The presence of two bands in the wild type lanes revealed that two sizes of the Utag WT were being expressed. The larger protein was the full-length Utag which bound SUMO. The smaller protein was a truncated version of Utag which did not bind SUMO. This smaller size was the same size as the mutant proteins, indicating that most of the mutant protein that had been expressed was a truncated form of Utag. This truncation was likely the reason none of the mutant proteins were binding SUMO.

It is unlikely that this truncation happened at the amber stop codon, since the non-mutated wild type protein and the mutant protein were both the same
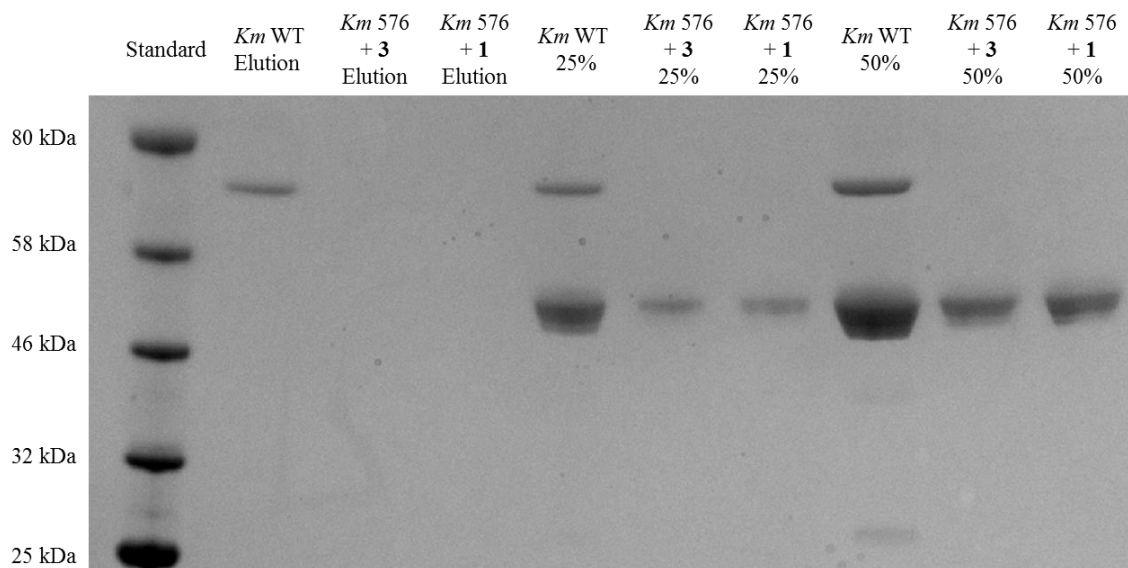

Figure 4.7: Binding Assay of *Km* 576 mutant proteins. Protein in the elution lanes indicate that the *Km* mutant bound to the bead immobilized SUMO and was eluted via SDS denaturation. The indicated percentage on the lanes indicates the percentage of the protein loaded onto the beads.

size. The successful Glaser-Hay reactions and the confirmation of coumarin incorporation also support the hypothesis that this truncation did not happen at the amber stop codon. If the translation machinery had read the amber stop codon as a regular stop codon and truncated the protein at position 576, the UAAs would not be incorporated. Since Utag is a yeast protein, it is likely that the difference in codon bias between the two organisms is responsible for the truncated Utag. Codon bias is the tendency of an organism to preferentially use one of the codons for an amino acid over other codons for the same amino acid. For instance, yeast typically use AGA for arginine, while *E. coli* use other codons for arginine.[7,8] Because of this tendency, yeast cells synthesize more tRNAs containing AGA than bacteria do. These tRNAs that are not usually synthesized in bacteria are called rare tRNAs in the context of bacterial protein expression. When Utag expression is induced in *E. coli*, the bacteria use up all the available rare tRNAs before the protein is fully synthesized, resulting in the truncated protein we observe. To correct for this codon bias, a pRIL plasmid was introduced to the expression system. The pRIL plasmid encodes the genes for rare tRNA synthesis, allowing the bacteria to make the tRNAs necessary to synthesize a heterologous protein. This pRIL plasmid is under chloramphenicol selection, so a different synthetase plasmid was also introduced to the expression system. The new synthetase plasmid contains the same pCNF synthetase in a different plasmid backbone that is under spectinomycin selection. The pRIL plasmid was transformed into *E. coli* cells which were made

electrocompetent to facilitate facile transformation with the protein and synthetase plasmids. Using this new three plasmid expression system, *Km* 576
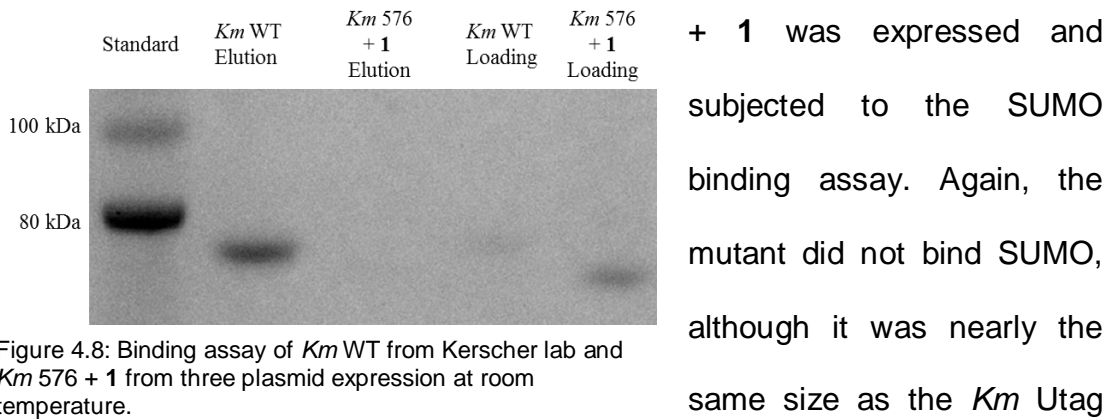
Figure 4.8: Binding assay of *Km* WT from Kerscher lab and *Km* 576 + **1** from three plasmid expression at room temperature.

+ **1** was expressed and subjected to the SUMO binding assay. Again, the mutant did not bind SUMO, although it was nearly the same size as the *Km* Utag

WT (Figure 4.8). This indicates that the bacteria were able to synthesize nearly all of the mutant Utag before running out of tRNAs, and that more optimization of the expression protocol is needed.

To encourage expression of full-length Utag, the three-plasmid expression system was used to express *Km* 576 + **2** and *Km* Utag WT at 18 ˚C
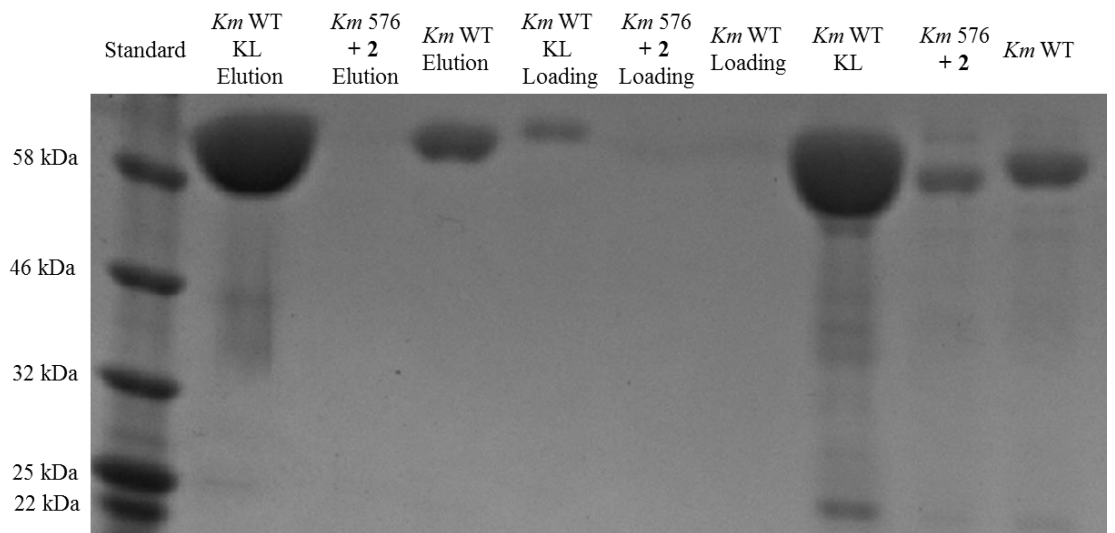

Figure 4.9: Binding assay of *Km* WT obtained from Kerscher lab (KL) as well as *Km* WT and *Km* 576 + **2** from three plasmid expression at 18 ˚C in 1% glycerol.

65

in induction media containing 1% (v/v) glycerol. The cooler temperature and IPTG-glycerol combination slow the growth rate of the *E. coli.*[9] Slowing the growth of the cells means that the cells can devote more resources to synthesis of a full-length protein instead of to reproducing themselves. These proteins were assayed for their SUMO binding ability alongside a positive control of full-length *Km* WT obtained from the lab of Oliver Kerscher. The *Km* WT from Kerscher lab and *Km* WT the triple plasmid expression system bound SUMO, but the *Km* 576 + **2** did not bind SUMO (Figure 4.9). The protein expressed with the new method is full length, so *Km* 576 is likely not a good position for amber suppression, since UAA incorporation appears to interfere with the folding or another mechanism that is responsible for SUMO binding.

Considering, the difficulties present in expressing suitable Utag protein containing a UAA, the amber stop codon was moved to sites in the maltose binding protein (MBP) that is fused with both the *Sc* Utag and *Km* Utag genes. Q5 PCR was used to generate two MBP mutants, MBP Y99TAG and MBP Y341TAG. These plasmids, along with the previously generated *Sc* 241 and *Sc* 253, will be sequenced to confirm the TAG mutation, then expressed with a UAA and assessed for SUMO binding ability. Additionally, another Km Utag mutant, *Km* S406TAG, will be generated and assayed for SUMO binding ability.

We have shown that it is possible to site specifically incorporate an unnatural amino acid into full length *Km* Utag mutants and into truncated *Sc* Utag mutants using a three-plasmid system induced at 18 ˚C in 1 % glycerol.

66

In the future, these findings will be extended to produce a full length Utag protein with an unnatural amino acid incorporated either directly into the Utag or into its fusion protein, MBP. This construct will then be expressed in a fusion with a CPP to facilitate its uptake into cells. This complex will be subjected to a Glaser-Hay reaction to conjugate it to a fluorophore so the complex can be tracked though the cell. This protein complex will then be incubated with mammalian cells, where the CPP will facilitate uptake of the complex and the complex will bind to SUMO *in vivo*. Further optimization will allow this complex to be developed into a robust diagnostic tool to identify cancerous cells in prostate cell samples.

Materials and Methods

General: All reagents were obtained from Acros, USB, Biotium, Alfa Aesar, Sigma Aldrich or Amresco and used without further purification unless noted. Unnatural amino acids were prepared according to literature protocols.[10] Protein and DNA gels were analyzed using a Bio-Rad gel imager (Bio-Rad Molecular Imager Gel Doc XR+).

Quik-Change PCR: This protocol was carried out on a Biorad icycler, with the Quik-Change II kit. A pET plasmid harboring a MBP-Utag fusion was diluted to 24.7 ng/μL, 12.3 ng/μL, and 6.2 ng/μL. The PCR reaction mixture contained one of the dilutions of the plasmids (1 μL), 10 mM DNTPs (1 μL), KAPA Hi-Fi Polymerase (1 unit/ μL, 1 μL), forward primer (10 μM, 1 μL), reverse primer (10 μM, 1 μL), KAPA buffer (5 μL), and Milliq water (40 μL). The reaction mixture

was subjected to the following heating protocol: 95 ˚C (1 min); eighteen cycles of melting (95 ˚C, 30 s), annealing (55 ˚C, 30 s), and extension (68 ˚C, 6 min); a final extension period (68 ˚C, 6 min) and then held at 4 ˚C. Next the parent plasmid was digested and the mutant plasmid was ligated by adding 2 µL of DPN1 [20,000 units/µL], T4 Ligase buffer (4 µL, 5x) and T4 Ligase (1 µL). The mixtures were then subjected to the following heating protocol: 37 ˚C 2 h and heat deactivation (80 ˚C, 15 min). The reaction was transformed (2 µL) into BL21 DE3 *Escherichia coli* cells by heat shock. The transformed cells were plated (250 µL) onto agar containing ampicillin and incubated at 37 ˚C overnight. The resulting cells were miniprepped using IBI High Speed Plasmid Mini Kit. Alternately, the pET-Utag plasmid was used at 12.3 ng/µL and a heat gradient was used instead of the concentration gradient. In the heat gradient version of the protocol, all procedures are the same except the annealing temperature was 53 ˚C, 55.1 ˚C, 61.3 ˚C, or 63 ˚C and all reactions contained 12.3 ng/µL of parent plasmid. The resulting plasmids were analyzed for successful insertion of the TAG mutation by sequencing at Genewiz. The primers were obtained from IDT DNA Technologies Inc. and are as follows: *Sc* Y147TAG 5'-CCG GAC GCC TTG CTA ACC CCT TTC TGA TAA ATT G-3' (forward), 5'-CAA TTT ATC AGA AGG GGT T<u>TA G</u>CA AGG CGT CCG G-3' (reverse); *Sc* Y241TAG 5'-ACA TAT ATT CCA CTG TCC <u>TAG</u> CCA TTT GGT TGC TGC-3' (forward), 5'-GCA GCA ACC AAA TGG CTA GGA CAG TGG AAT ATA TGT-3' (reverse); *Sc* Y253TAG 5'-GGC GCA TCT GCA CTT CCC <u>TAG</u>

AGA GTA TTC ATA CAA AC-3' (forward), 5'-GTT TGT ATG AAT ACT CTC TAG GGA AGT GCA GAT GCG CC-3' (reverse).

Q5 PCR: This protocol was carried out on a thermocyler (Applied Biosystems 2720 Thermocycler) with the Q5 Site Directed Mutagenesis Kit (New England Biolabs). A pET plasmid harboring a MBP-Utag fusion was diluted 1:10, 1:100, 1:1000. The PCR reaction mixture contained one of the dilutions of the plasmids (1 μL), forward primer (10 μM, 1.25 μL), reverse primer (10 μM, 1.25 μL), Q5 Hot Start Master Mix (12.5 μL), and Milliq water (9 μL). The reaction mixture was subjected to the following heating protocol: 98˚C (30 s); eighteen cycles of melting (98˚C, 10 s), annealing (60-70 ˚C, 30 s), and extension (72 ˚C, 8 min); a final extension period (72 ˚C, 10 min) and then held at 4 ˚C. The annealing temperature was changed to match the annealing temperature of the primers for each reaction. Next the parent plasmid was digested and the mutant DNA was circularized using the Q5 Site Directed Mutagenesis Kit (New England Biolabs). The reaction was transformed (5 μL) by heat shock into NEB 5-α competent *Escherichia coli*. The transformed cells were plated (15, 60, 120 μL) onto agar containing ampicillin and incubated at 37 ˚C overnight. The resulting cells were miniprepped using IBI High Speed Plasmid Mini Kit. The resulting plasmids were analyzed for successful insertion of the TAG mutation by sequencing at the William and Mary Molecular Core Facility. The primers were obtained from IDT DNA Technologies Inc. and are as follows: *Km* Y576TAG 5'-TAG AGC CTA GAT TTC AAT GCA CAA GAT GCG GTT AAT ATG-

69

3'(forward), 3'-AT TTT TGC TCA TAT AAA GAG TAT TCA AGC AAA CGT ATA TTC-5' (reverse); *Km* S406TAG 5'-<u>TAG </u>GTC AGT TTC TTA ATA ATG AAG AAC ATT CAG TCA TAT TTG-3' (forward), 3'-A TTT GCT CCC GAA GAG AGC GAA TCA AGA TAT AAT ATC CTT TTA TTT TTG-5' (reverse); MBP Y341TAG 5'- <u>TAG</u> GCC GTG CGT ACT GTG ATC AAC GCC GCC AGC-3' (forward), 3'- CCA GAA AGC GGA CAT CTG CGG GAT GTT CGG CAT G-5' (reverse); MBP Y99TAG 5'- AAC GGC AAG CTG ATT ATT GCT TAC CCG ATC GCT-3' (forward), 3'- <u>CTA</u> ACG TAC GGC ATC ATC CCA GGT AAA CGG ATA C-5' (reverse).

Expression and Purification of Protein: A pET plasmid harboring a variable MBP-Utag fusion mutant (0.5 µL) was co-transformed with a synthetase plasmid (pEVOL-pCNF unless stated, 0.5 µL) into BL21 (DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (0.034 mg/mL), ampicillin (0.050 mg/mL) at 37 °C overnight. One colony was then used to inoculate an aliquot of LB media (5 mL) containing ampicillin and chloramphenicol. The cultures were incubated at 37 °C overnight and used to inoculate an expression culture (100 mL LB media, 0.050 mg/mL Amp, 0.034 mg/mL Chlor) at an OD_{600} 0.1. The cultures were incubated at 37 °C to an OD_{600} between 0.6 and 0.8 at 600 nm, and protein expression was induced by addition of an unnatural amino acid (UAA; 1 mL, 100 mM) and 20 % arabinose (1.5 mL) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 1.5 mL). The cultures were allowed to shake at

room temperature for 5 h then centrifuged at 5,000 rpm for 10 minutes and stored at -80˚C for 20 min. The cell pellet was re-suspended in SPB with 5mM TCEP (2 mL). The cell suspension was sonicated on ice (3 x 10 s, 1 min rests, 30 % amplitude). The cellular debris was pelleted by centrifugation (15,000 rpm, 8 min) and the resulting supernatant was diluted to 5 mL with SPB containing 5mM TCEP and loaded onto Talon Metal Affinity resin (200 µL). The resin was washed with SPB (3 x 3 mL) and the protein was eluted with 250 mM imidazole in SPB (250 µL). TCEP was added to each elution and the protein was frozen at -80˚C in 10 % glycerol.

CuACC Reaction: The Utag + **1** mutant was thawed and buffer exchanged into PBS containing 1mM DTT using a molecular weight cutoff column. Tris[(1-benzyl-1H-1,2,3-triazol-4yl)methyl]amine (TBTA; 5mM, 20 µL) was added to an Eppendorf tube containing 50 mM tris(2-carboxyethyl)phosphine (TCEP; 2 µL) and 50 mM $CuSO_4$ (2 µL). The catalyst system was diluted in phosphate buffered saline solution (PBS; 6 µL) and 0.335 mg/mL of Utag-*p*AzF mutant (20 µL) was added along with 1 mM Alkyne fluorophore (Fluor-488- Alkyne, 10 µL). This mixture was incubated at 4˚C for 14 h then transferred to a molecular weight cutoff column and washed with PBS Buffer (8 x 80 µL) and centrifuged at 13,200 rpm for 2 min. The reactions were then analyzed via SDS-PAGE (10 %).

Glaser-Hay Reaction: The Utag + **2** mutant was thawed and buffer exchanged into PBS containing 1mM DTT using a molecular weight cutoff column. The

reactants were added to a microcentrifuge tube in the following order. Copper iodide (5 µL, 500mM) was added to 0.335 mg/mL Utag-*p*Prf (15 µL). Next N,N,N′,N′-Tetramethylethylenediamine (TMEDA; 500 mM, 2 µL) and 1 mM alkyne fluorophore (Flur-488-Alkyne, 10 µL) and PBS (8 µL) were added to the reaction. The reaction was incubated at 4 ˚C for 4 h then transferred to a molecular weight cutoff column and washed with PBS Buffer (8 x 80 µL) and centrifuged at 13,200 rpm for 2 min. The reactions were then analyzed via SDS-PAGE (10%).

Pull Down Assay on Fluorescent Labeled Utag: Bead containing immobilized SUMO (Boston Biochemical, rhSUMO-1 beads; 25 µL) were washed with SPB (1 mL). SPB (900 µL), TCEP (50 mM, 100 µL) and Utag variant protein (30 µL) were added to the washed beads and rotated for 30 min at room temperature. The beads were washed with SPB (3 x 1 mL) by pelleting the beads and pipetting the supernatant SPB off. The beads were resuspended in PBS (20 µL) and analyzed using a Bio-Rad gel imager (BioRad Molecular Imager Gel Doc XR+) on the SPYRO Ruby setting. The resin was also imaged by transferring the resin to a 96-well plate and adding PBS (50 µL). The resin was then analyzed on a plate reader (Biotek Synergy HT Microplate Reader). The protocol was set to measure the fluorescence end point after 5 s of shaking, at an excitation wavelength of 485 nm and an emission wavelength of 528 nm.

Plate Pull Down Assay: Poly-L-lysine (200 µL, 0.1% w/v in water) was added to two wells on a 96-well plate and incubated at room temperature for 5 min.

The wells were emptied and allowed to dry for 5 minutes, then washed once with SPB (300 µL). SUMO-1 was added to the wells and the plate was rocked for 30 minutes. The wells were washed three times with SPB (300 µL) and 1% (w/v) BSA in SPB (200 µL) was added to the wells. The plate was rocked at 4 ˚C for 16 h, then the wells were washed with SPB. The variant Utag Glaser-Hay (10 µL) was added to 90 µL to SPB containing 5 mM TCEP. 50 µL of the diluted variant Utag Glaser Hay was added to the treated wells and 50 µL were added to untreated wells. The plate was rocked at 4 ˚C for 1 h. The wells were washed three times with SPB containing 5 mM TCEP (300µL) and analyzed on a plate reader (Biotek Synergy HT Microplate Reader). The protocol was set to measure the fluorescence end point excitation wavelength was 485 nm and the emission wavelength was 528 nm.

Pull Down Assay: Bead containing immobilized SUMO (25 µL) were washed with SPB (1 mL). SPB (900 µL), TCEP (50 mM, 100 µL) and Utag variant protein (30 µL) were added to the washed beads and rotated for 30 min at room temperature. The beads were washed with SPB (3 x 1 mL) by pelleting the beads and pipetting the supernatant SPB off. The protein was eluted by adding of 2x SDS-PAGE loading dye (20 µL). The reactions were then analyzed via SDS-PAGE (10%).

Fluorescence Measurement to Confirm Coumarin Incorporation: After purification 10 µL of the *Km* 576 Utag + **3** was added to a quartz cuvette and diluted up to 3 mL with PBS. This was then analyzed on a fluorimeter (Perkin

Elmer LS 55 Luminescence Spectrometer). The mixture was excited at 340 nm with a 10 nm slit width for the excitation wavelength and 5 nm slit width for the emission wavelengths. The scan speed was set at 300 nm/min.

References

1. Schipper, M.; Wang, G.; Giles, N.; Ohrnberger, J. Novel Prostate Cancer Biomarkers Derived from Autoantibody Signatures. *Transl. Oncol.* **2015***, 8*, 106-111.

2. Li, S.; Hochstrasser, M. A new protease required for cell-cycle progression in yeast. *Nature* **1999***, 398*, 246-251.

3. Mossessova, E.; Lima, C. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* **2000***, 5*, 865-876.

4. Elmore, Z. C.; Donaher, M.; Matson, B. C.; Murphy, H.; Westerbeck, J. W.; Kerscher, O. Sumo-dependent substrate targeting of the SUMO protease Ulp1. *BMC Biol.* **2011***, 9*, 74.

5. Baek, S. H. A novel link between SUMO modification and cancer metastasis. *Cell Cycle* **2006***, 5*, 1492-1495.

6. Wang, J.; Xie, J.; Schultz, P. G. A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* **2006***, 128*, 8738-8739.

7. Letzring, D. P.; Dean, K. M.; Grayhack, E. J. Control of translation efficiency in yeast by codon-anticodon interactions. *RNA* **2010***, 16*, 2516-2528.

8. Gustafsson, C.; Govindarajan, S.; Minshull, J. Codon bias and heterologous protein expression. *Trends Biotechnol.* **2004***, 22*, 346-353.

9. Law, J.; Lee, S.; Tseng, A.; Tsui, K. W.; Yu, N. The role of glycerol and isopropyl thiogalactoside in Escherichia coli growth and lactose induction of β-galactosidase. *Journal of Experimental Microbiology and Immunology (JEMI) Vol* **2002***, 2*, 97-102.

10. Maza J.C.; Nimmo Z.M.; Young D.D. Expanding the scope of alkyne-mediated bioconjugations utilizing unnatural amino acids. *Chem Commun.* **2016**, *52*, 88–91.

Chapter 5: Functionalizing Divalent Bioconjugates to Multivalent Complexes

Bioconjugations have found a wide range of applications in modern science, from antibody-drug conjugations to enzyme immobilization. One limitation of these applications is that only two molecules can be linked together, limiting the complex performing to two functions. This creates a problem when more than two functionalities need to be incorporated into a conjugate, as in molecular assembly lines where a degree of multivalency is required. Molecular assembly lines also require controlled enzyme orientation for maximum efficiency. UAAs can represent a precise mechanism for the introduction of more complex bioconjugates to address these problems. UAAs can be site-specifically inserted to minimize interference with the enzyme's function and
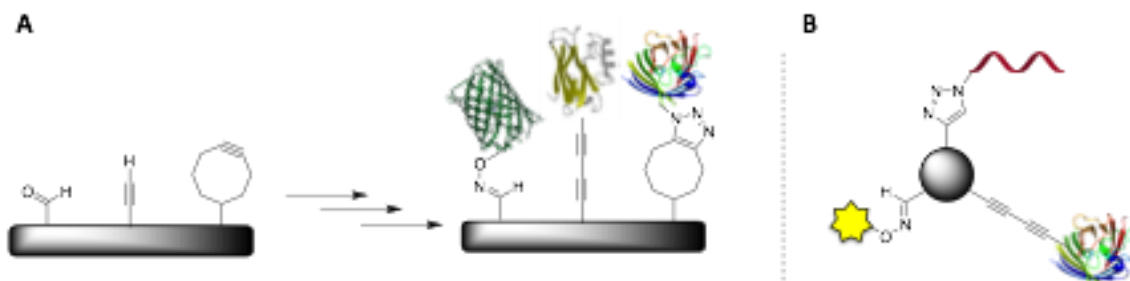

Figure 5.1: Potential applications of multivalent complexes (A) molecular assembly line with proteins immobilized by UAAs (B) a multivalent complex with a fluorophore (yellow), a therapeutic peptide (red) and a targeting molecule (multicolored) conjugated together via UAAs.

optimize the orientation of the protein in reference to the solid support (Figure 5.1A). Multivalent complexes would also be useful in medical applications, for instance in developing ligand-targeted therapeutics. In this system a

therapeutic molecule, a fluorophore and an antibody could be complexed together to make simultaneous cellular targeting and tracking of the drug possible (Figure 5.1B). Multivalent conjugation could also be used as an alternate to hapten-carrier complexes in vaccines. Several small molecules could be conjugated together, instead of to a carrier protein, to yield an appropriately immunogenic molecule.

Multivalent reactions have been developed in organic solvents at high temperatures, but they have not been explored in a biological context due to a lack of appropriate bioconjugates. M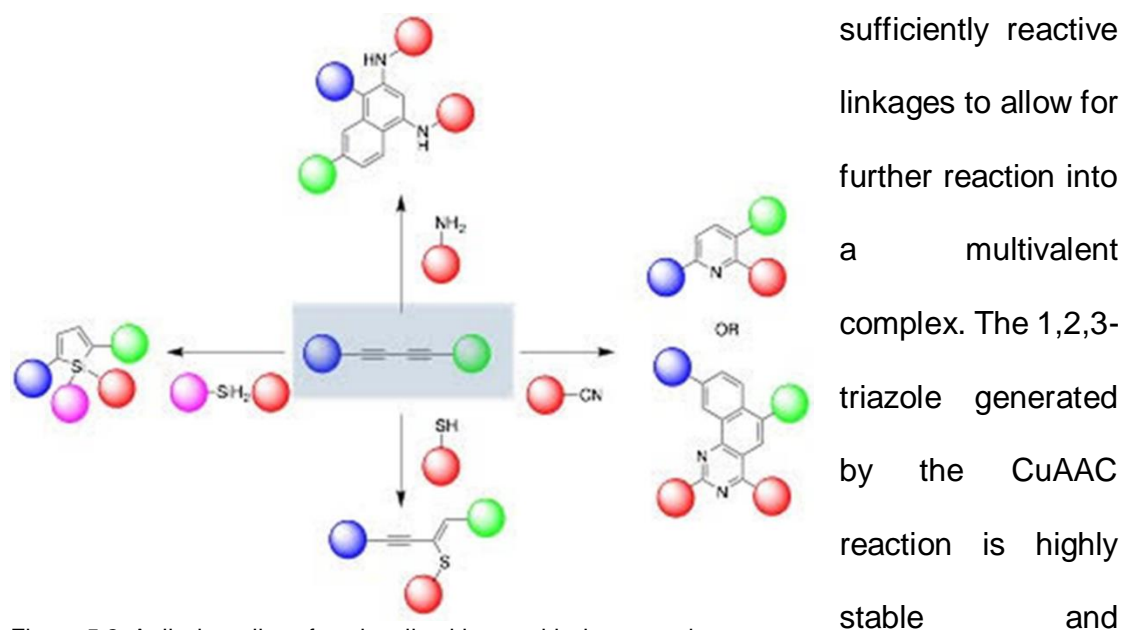ost bioconjugations do not install sufficiently reactive linkages to allow for further reaction into a multivalent complex. The 1,2,3-triazole generated by the CuAAC reaction is highly stable and unreactive, making bioconjugates containing this chemistry unsuitable for further reaction to create a multivalent complex. The biological Glaser-Hay reaction developed by Lampkowski et al. is unique in this respect, as it results in an electron dense diyne linkage that affords further reaction to create

Figure 5.2: A divalent diyne functionalized into multivalent complexes.

multivalent complexes. Figure 5.2 illustrates some organic reactions that result in multivalent complexes, which theoretically can be optimized to perform under physiological conditions. One such reaction is a reaction between a diyne and an amine. This reaction was developed in an organic context by Sun and colleagues using cyclic amines such as pyrrolidine (Scheme 1A). This reaction will be optimized for biological conditions in this chapter.

Results and Discussion

To replicate the literature protocol, 1,4-diphenylbuta-1,3-diyne (**1**) and hexa-2,4-diyne-1,6-diol (**2**) were synthesized using a Glaser-Hay reaction. **1** was successfully reacted at 80 °C using pyrrolidine as a solvent to create **3** and analyzed by ^{1}H NMR (Scheme 5.1A). To test the protocol at physiological

A

B
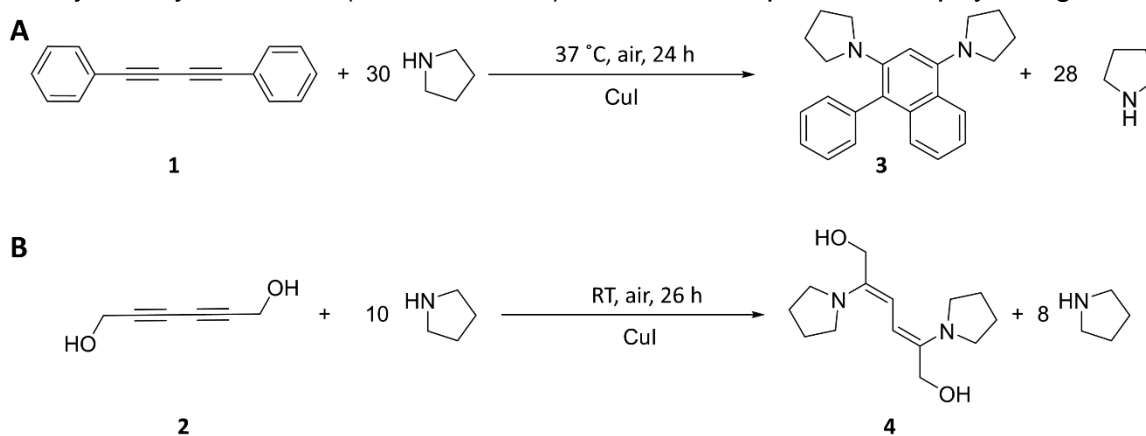
Scheme 5.1: Cyclic amine mediated multivalent reaction on a diyne (A) as described by Sun et al. (B) under biological conditions.[1]
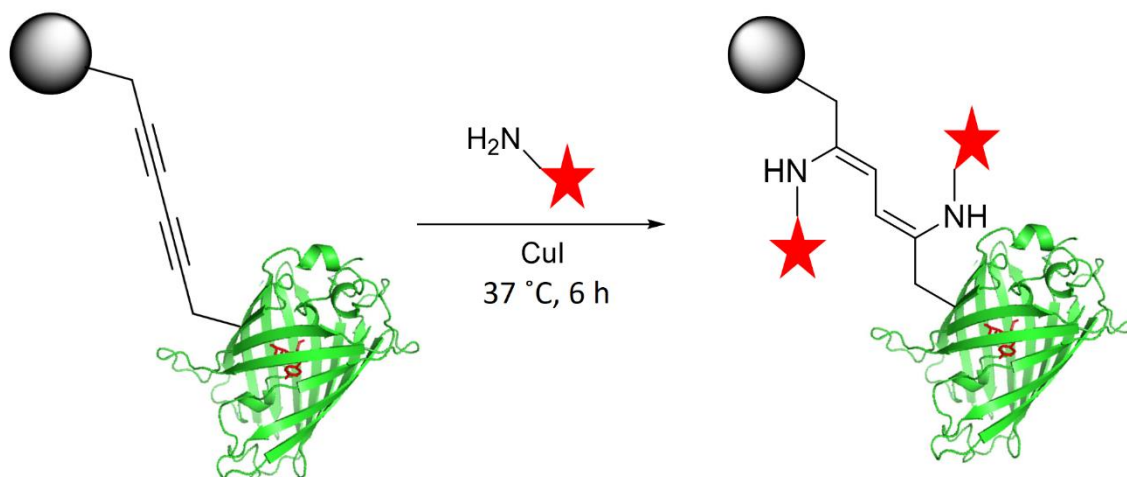
Figure 5.3: Multivalent reaction with immobilized GFP and 5-Aminofluorescein (red star).

temperatures, **1** was reacted, still using pyrrolidine as a solvent, at 37 ˚C for 24 h. This reaction was also successful, so the reaction was tested in aqueous solvent. For this reaction, **2** was synthesized to overcome **1**'s lack of aqueous solubility. **2** was reacted at room temperature in 10% DMSO in water, a more physiologically appropriate solvent than pyrrolidine. The product was analyzed by mass spec to confirm the presence of **4**, indicating that the reaction was successful (see experimental). Since this reaction was working under physiological conditions, this methodology was extended to a more biologically relevant context using GFP (Figure 5.3). GFP 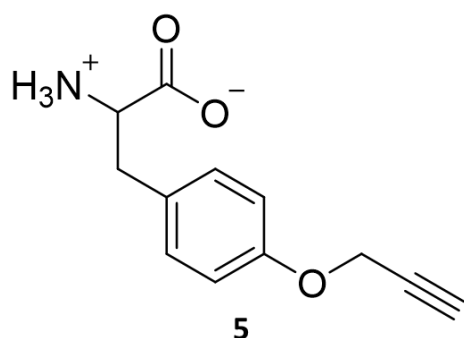was expressed containing *para*-propargyloxyphenylalanine (*p*PrF, **5)**, and reacted with the immobilized alkyne moiety by a biological Glaser-Hay (Figure 5.4).[3] Aminofluorescein was incubated with the resin and copper for 6 h. The resin was then washed and analyzed by UV-irradiation on

Figure 5.4: The unnatural amino acid incorporated into GFP, pPrF, which contains the alkyne functionality utilized in the Glaser-Hay reaction.

79

a Bio-Rad gel imager. Unfortunately, the resin did not exhibit any increase in fluorescence compared to a negative control, indicating that further optimization of the is reaction is necessary.

Thus far, we have successfully translated the multivalent reaction devised by Sun and colleagues to biological conditions. In the future, this methodology will be optimized for GFP immobilized on a sepharose bead, as well as more relevant enzymes such as SSO Carboxylesterase P1. The multivalent enzyme complex can then be assayed for catabolic activity to ensure the conjugation did not affect the function of the enzymes. In the future, this technology could be applied to create a molecular assembly line, enhanced tools for drug delivery or alternative vaccines.

Materials and Methods

General: Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Reactions were conducted under ambient atmosphere with non-distilled solvents. Unnatural amino acids were prepared according to literature protocols.[4,5,?] NMR data was acquired on an Agilent 400 MHz. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit. Mass data was acquired on a Finnigan LCQ Linear Ion Trap mass spectrometer. UV irradiation was performed on a Bio-Rad gel imager (Bio-Rad Molecular Imager Gel Doc XR+) on the SPYRO Ruby setting.

Synthesis of 1: Phenylacetylene (0.93 g/mol, 2.15 mL) was added to THF (24 mL) in a flame dried round-bottom flask. Copper iodide (CuI, 160 mg) was dissolved in TMEDA (0.775 g/mol, 240 μL) and THF (8 mL). This mixture was then combined with the Phenylacetylene and stirred overnight (16 h) at 60 ˚C. The reaction was then analyzed by Thin Layer Chromatograpgy (TLC) in Hexanes and extracted in DCM and brine. The reaction was further analyzed by ^{1}H NMR after evaporation on a vacuum line for 2 h (Scheme 2).

Synthesis of 2: Propargyl alcohol (0.949 g/mol, 2.11 mL) was added to THF (24 mL) in a flame dried round-bottom flask. CuI (160 mg) was dissolved in TMEDA (0.775 g/mol, 240 μL) and THF (8 mL). This mixture was then combined with the Propargyl alcohol and stirred for 48 h while bubbling air through at 60 ˚C. The reaction was then purified via column chromatography starting with a mixture of 1:3 Hexanes: Ethyl Acetate and ending with a 1:1 mixture of Hexanes: Ethyl Acetate. The reaction was further analyzed by ^{1}H NMR after evaporation on a vacuum line for 2 h (Scheme 2).
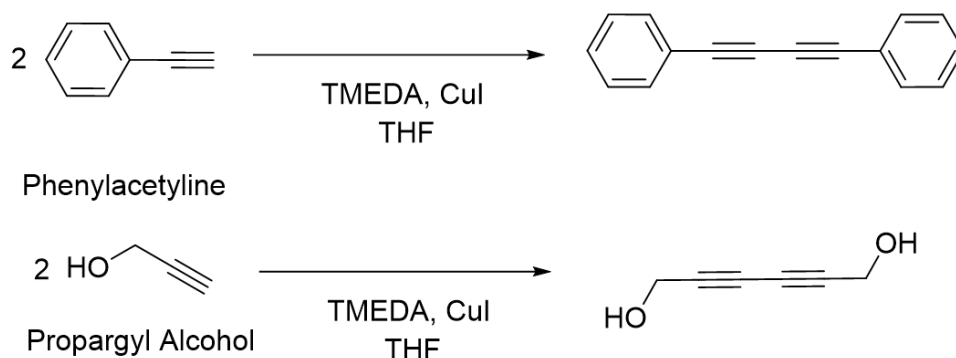
Scheme 5.2: Synthesis of the starting materials **1** and **2** via a Glaser-Hay reaction.

Multivalent Reaction with 1: Diphenylacetylene (20 mg) and CuI (5 mg) were added to Pyrrolidine (0.866 g/mol, 240 µL) in a vial. The mixture was stirred at 80 ˚C for 6 h, then 15 mL of DCM was added to the vial. The reaction was filtered and analyzed by thin layer chromatography (TLC) in Hexanes. The crude reaction was subjected to column chromatography using 10:1 Petroleum Ether: Ethyl Acetate mixture.[2] The fractions were analyzed by TLC and ^{1}H NMR in Chloroform after evaporating on a vacuum line for 1 h. This reaction was also done at 37 ˚C for 24 h.
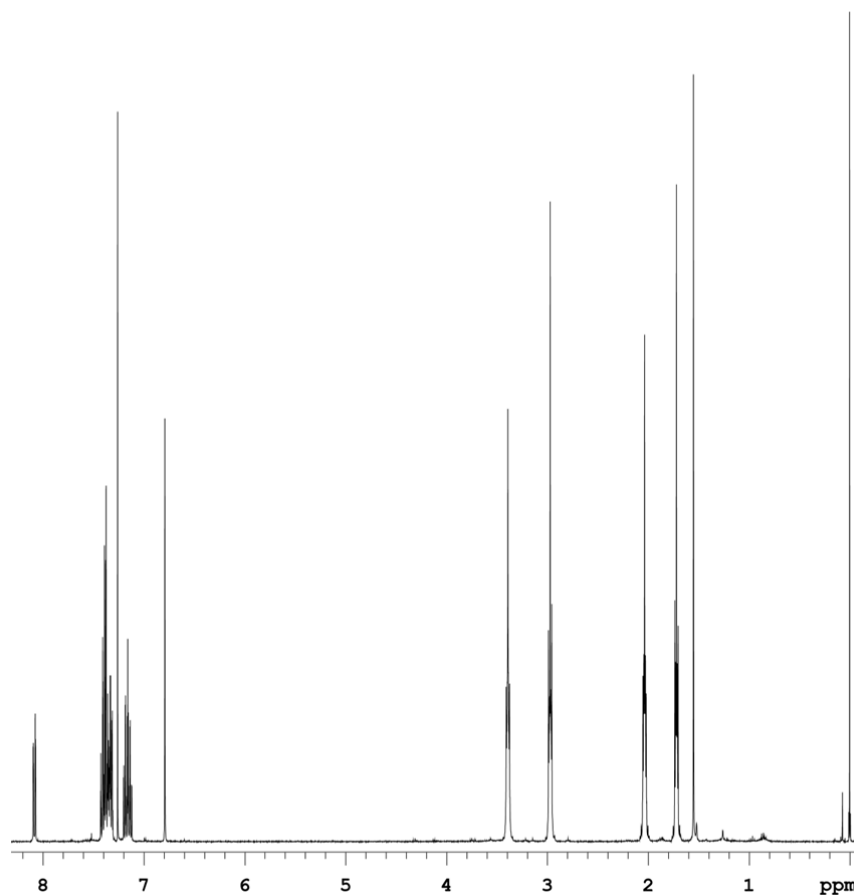
Figure 5.5: ^{1}H NMR of **3**, from the 80 ˚C multivalent reaction using **1**. All peaks agree with the literature, except the CDCl$_3$ peak at 7.26 ppm, the water peak at 1.55 ppm, the small peak at 0.08 ppm, and the TMS peak at 0 ppm.[1]

Multivalent Reaction with 2: Dipropargyl alcohol (20 mg) and CuI (5 mg) was added to pyrrolidine (0.866 g/mol, 112 µL) and 10% DMSO in distilled water (128 µL) in a vial.[2] The mixture was stirred for 26 h at room temperature then subjected to column chromatography using 10% Methanol in DCM. The fractions were allowed to evaporate for 16 h then analyzed by TLC in 10% Methanol in DCM. Fractions 12-15 were sepaated and allowed to evaporate on a vacuum line for 3 h, then analyzed by ^{1}H NMR. The presence of aromatic

83

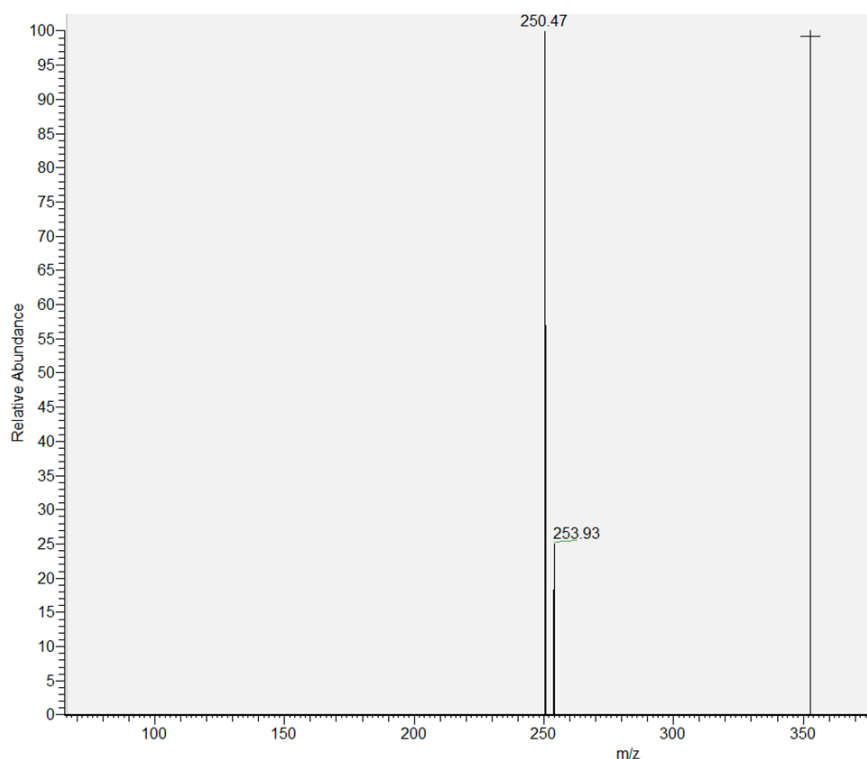peaks prompted further analysis by LCMS, where the product was observed at 253 m/z.

Figure 5.6: Mass Spectra of **4**, showing a peak at 253.93 m/z, the weight of 4 (252.18 Da) plus one hydrogen within the error of the instrument.

Incorporation of *p*PrF: A pET-GFP-TAG-151 plasmid (0.5 µL) was cotransformed with a pEVOL-*p*CNF plasmid (0.5 µL) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (0.034 mg/mL) and ampicillin (0.050 mg/mL) at 37 ˚C overnight. One colony was then used to inoculate LB media (4 mL) containing both ampicillin and chloramphenicol. The culture was incubated at 37 ˚C overnight and used to inoculate an expression culture (10 mL LB media, 0.050 mg/mL Amp, 0.034

mg/mL Chlor) at an OD_{600} 0.1. The cultures were incubated at $37\,^{\circ}$C to an OD_{600} between 0.6 and 0.8 at 600 nm, pelleted and resuspended in 2 mL of media containing ampicillin and chloramphenicol. Protein expression was induced by addition of pPrF (100 mM, 20 µL) and 20 % arabinose (2 µL) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 2 µL). The cultures were allowed to shake at $30\,^{\circ}$C for 16-20 h then centrifuged at 5,000 rpm for 10 minutes and stored at $-80\,^{\circ}$C for 20 minutes. The cell pellet was re-suspended using 500 µL of Bugbuster (Novagen) containing lysozyme, and incubated at $37\,^{\circ}$C for 20 minutes. The solution was transferred to an Eppendorf tube and centrifuged at 15,000 rpm for 10 minutes, then the supernatant was poured into an equilibrated His- pur Ni-NTA spin (Qiagen) column with nickel resin (200 µL) and GFP was purified according to manufacturer's protocol. Purified GFP was analyzed by UV irradiation.

Immobilization of Alkyne onto Epoxy Sepharose 6B Resin Epoxy-activated 6B Sepharose (GE Healthcare, 200 mg) was added to a filter syringe and washed with distilled water (5 x 3 mL). Propargyl alcohol (700 µmol) and coupling buffer (3.5 mL, PBS at pH 13.0) was added to a 15 mL tube followed by the resin. The mixture was shaken at room temperature for 16 h. The resin was transferred to a filter syringe and washed with coupling buffer (4 mL). The resin was transferred to a 15 mL tube and capped with ethanolamine (3.5 mL). The resin was incubated at 30 ℃ for 4 h then washed (3 x 3 mL) in a filter syringe with alternating acetate buffer (10 mM, pH 4) and tris-HCl buffer (pH 8).

Glaser-Hay to Alkyne Resin: To an Eppendorf tube, CuI (500mM, 5 µL) was added along with TMEDA (500mM, 5 µL). The mixture was equilibrated by shaking for 10 min at 37 ˚C. Propargyl Sepharose resin (20 mg) was added, and the mixture was equilibrated for 10 min at 37 ˚C. PBS (pH 7.8, 10 µL) and GFP-151-*p*PrF (2.3 mg/mL, 20 µL) were added. The reaction was incubated at 4 ˚C for 6 h, then washed with PBS (4 x 200 µL). The reaction was analyzed via UV irradiation.

Multivalent Conjugation to GFP Immobilized Resin: Sepharose resin with GFP-151-*p*Prf (20 mg) immobilized was added to a mixture of CuI (1 M, 5.5 µL), 5-Aminofluorescein (50 mM, 5 µL), and distilled water (45 µL). The negative control contained sepharose resin with the propargyl linker immobilized (20 mg), 5-Aminofluorescein (50 mM, 5 µL), and distilled water (45 µL). Both mixtures were incubated with shaking for 6 h at 37 ˚C while protected from light. The resin was then transferred to a spin column and washed with PBS (20 x 100 µL). The resins were compared to each other by UV-irradiation on a Bio-Rad gel imager.

References

1. Sun, H.; Wu, X.; Hua, R. Copper(I)-catalyzed reaction of diaryl buta-1,3-diynes with cyclic amines: an atom-economic approach to amino-substituted naphthalene derivatives. *Tetrahedron Lett.* **2011***, 52*, 4408-4411.

2. Raliski, B. K.; Howard, C. A.; Young, D. D. Site-Specific Protein Immobilization Using Unnatural Amino Acids. *Bioconjug. Chem.* **2014**, *25*, 1916-1920.

3. Lampkowski, J. S.; Villa, J. K.; Young, T. S.; Young, D. D. Development and Optimization of Glaser-Hay Bioconjugations. *Angew. Chem. -Int. Edit.* **2015**, *54*, 9343-9346.

4. Maza J.C.; Nimmo Z.M.; Young D.D. Expanding the scope of alkyne-mediated bioconjugations utilizing unnatural amino acids. *Chem Commun.* **2016**, *52*, 88–91.

Chapter 6: Utilization of alkyne bioconjugations to modulate protein function

Protein engineering is a powerful tool for the development of new therapeutics, catalysts, and biosensors.[1–8] While many advances in the field have been made, designing novel protein functionality is still a challenge, as it requires an intricate understanding of the subtle interplay between protein structure and function. Current engineering techniques often focus on using selections and screens to optimize or enhance existing protein functionality.[2] While this method has proved useful for optimizing existing function, generating new function where it does not exist is still a hurdle.

While UAAs have allowed for the development of unique protein function, the evolved proteins are often limited to a single new function, depending on the UAA incorporated.[2,10] In addition, the functionality is limited to the UAA itself, which suffers from constraints, such as the requisite for an aminoacyl tRNA synthetase capable of recognizing the UAA, the synthetic accessibility of the UAA, and the size of the UAA which may preclude its uptake by a biological system.[9] A more appealing strategy would allow for the generation of a UAA-containing protein "template" upon which researchers could synthetically introduce different chemical moieties that would in turn lead to altered protein function depending on the moiety employed.

Bioorthogonal chemistry, which employs reactions that proceed to completion under physiological conditions (pH 7, 37 °C), offers a unique mechanism to add new chemical functionality to proteins.[11,12] Indeed, a variety of reactions have been developed that can add novel chemistry to living systems. In particular, the cycloaddition between azides and alkynes that is either copper(I) mediated or strain promoted has become a widespread technique to introduce unique chemistry to proteins.[13–16] More recently, our group has developed a bioorthogonal variant of the Glaser-Hay reaction,[17,18] which brings together two terminal alkynes to form a diyne in the presence of copper(I) under physiological conditions.[19–21] The resulting stable diyne linkage has a well-defined linear geometry, and due to the abundance of commercially available terminal alkynes, a variety of chemical moieties can be reacted onto a protein using this technique. As such, we sought to utilize the power of this new chemistry to generate new and different protein function dependent upon the alkyne reaction partner and not the UAA alone.

Specifically, we designed a proof-of-concept experiment to alter the function of green fluorescent protein (GFP) via reaction of different terminal alkynes onto the GFP chromophore. GFP is a 27 kDa protein isolated from Aequorea victoria with photochemical properties arising from an internal chromophore composed of Ser65-Tyr66-Gly67.[22–24] New chemical properties afforded by UAA introduction in place of Tyr66 have already been documented to alter GFP's fluorescence profile.[25] All UAAs incorporated were found to blue-shift the fluorescence profile of GFP, with more highly conjugated UAAs exhibiting a

greater degree of spectral shifting. Based on these results, the ability to modulate the conjugation of GFP's fluorophore using UAA mutagenesis is apparent, and represents a convenient means to rationally design new protein function.[26] However, this mutagenesis approach is limited by the size and complexity of the UAA. An alternative approach involves exploiting the chemical functionality in pre-existing UAAs to serve as functional handles for bioorthogonal reactions, acting as a "template" for the chemical derivatization of new protein function.

This chemically templated protein function could be achieved via the genetic incorporation of *p*-propargyloxyphenylalanine (*p*PrF, **1**) or *p*-ethynylphenylalanine (*p*EtF, **2**) into residue 66 of GFP (Figure 6.1).[27] This provides a terminal alkynehandle for reaction with different chemical moieties via the bioorthogonal Glaser-Hay reaction. The resulting diyne linkage is highly conjugated, and a prime candidate to introduce new photochemical properties into GFP's fluorophore without the need to evolve a new aaRS or express multiple versions of the protein containing different UAAs. Moreover, the *p*EtF (**2**) is directly conjugated with the aromatic ring of the UAA, allowing for a comparison of
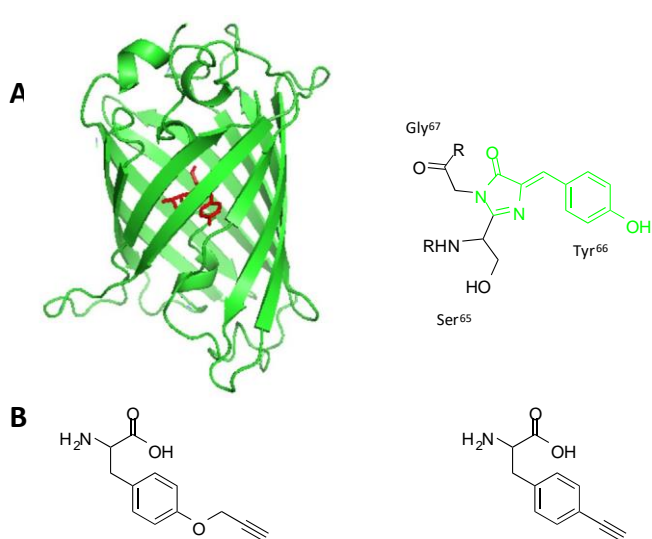
Figure 6.1: Modification of GFP function via UAA mutagenesis. (A) GFP b-barrel (green) with the chromophore (red) and chromophore structure. (B) Unnatural amino acids employed as bioconjugation handles.

90

the conjugation between the different UAAs. We hypothesized that the altered conjugation and chemical properties around the fluorophore would lead to new photophysical properties, demonstrating the utility of a chemically programmable protein engineering strategy. Herein we report our findings on utilizing the Glaser-Hay reaction on GFP's fluorophore to alter its fluorescent properties.

In order to obtain protein possessing an alkynyl moiety, a GFP plasmid harboring a TAG mutation at position 66 was co-transformed with the polyspecific pCNF-aaRS/tRNA pair.[28] Conveniently, this aaRS is capable of recognizing both **1** and **2** for expression of an alkyne-containing GFP.[20] As the alkynyl UAA is incorporated into position 66 in GFP's fluorophore, the extended conjugation afforded by the UAA alters the spectral properties of GFP. To assess that pPrF-GFP$_{TAG66}$ was successfully produced, spectra for the GFP-
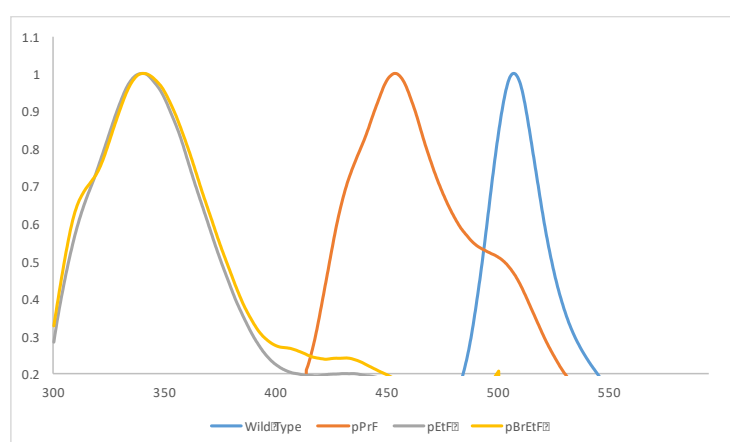

Figure 6.2: Fluorescence spectra for the alkynyl containing GFP mutants.

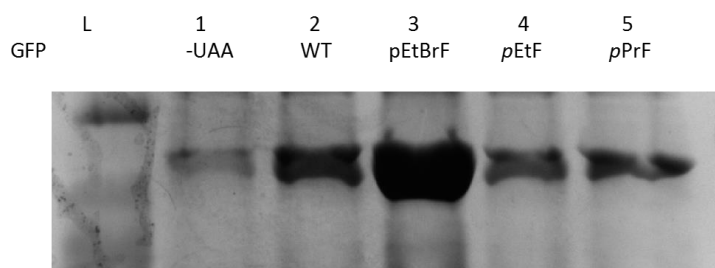Figure 6.3: GFP expression with different alkynyl unnatural amino acids. *E. coli* transformed with a pEVOL-pCNF aaRS plasmid and a pET-GFP-66TAG plasmid were induced in the presence of one of the alkynyl amino acids. Following Ni-NTA purification, GFP was present in all 4 expressions. Small amounts of WT GFP are observed in the absence of UAA as previously documented with the promiscuous pCNF-aaRS.

91

variant were compared to the wild type (Figure 6.2). These spectra exhibited blue-shifts, in agreement with the literature precedent, in the *p*PrF variant relative to the wild type. Incorporation was also confirmed by SDS-PAGE analysis of protein expression in the presence and absence of the UAA (Figure 6.3). A similar expression was performed using the *p*CNF aaRS/tRNA pair and *p*EtF to produce a separate GFP mutant with a bioconjugation handle. Conveniently, due to the modularity of this approach, only a single protein expression is necessary and all functional modification can be achieved synthetically. This is in contrast to previous experiments, which required an individual protein expression for each UAA in order to modify protein function.
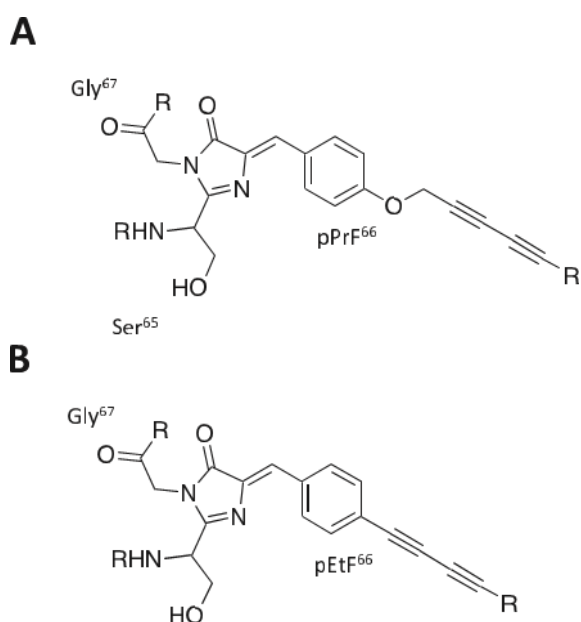

Figure 6.4: Structures of the fluorophores generated when performing a Glaser-Hay reaction with the two UAAs. (A) Product from the coupling of a terminal alkyne to a GFP mutant containing **1**. (B) Product from the coupling of a terminal alkyne to a GFP mutant containing **2**.

With a *p*PrF-GFP$_{TAG66}$ and *p*EtF-GFP$_{TAG66}$ in hand, we then sought to employ our previously reported bioorthogonal Glaser-Hay reaction to install new and varied chemical functionality into the chromophore of GFP (Figure 6.4). To investigate, we performed bioorthogonal Glaser-Hay reactions on the chromophore's alkyne handle to couple terminal alkyne-bearing aliphatic and aromatic compounds with different chemical functionalities. Glaser-Hay reactions were performed by using a working concentration of 500 mM of CuI

A

B

R =
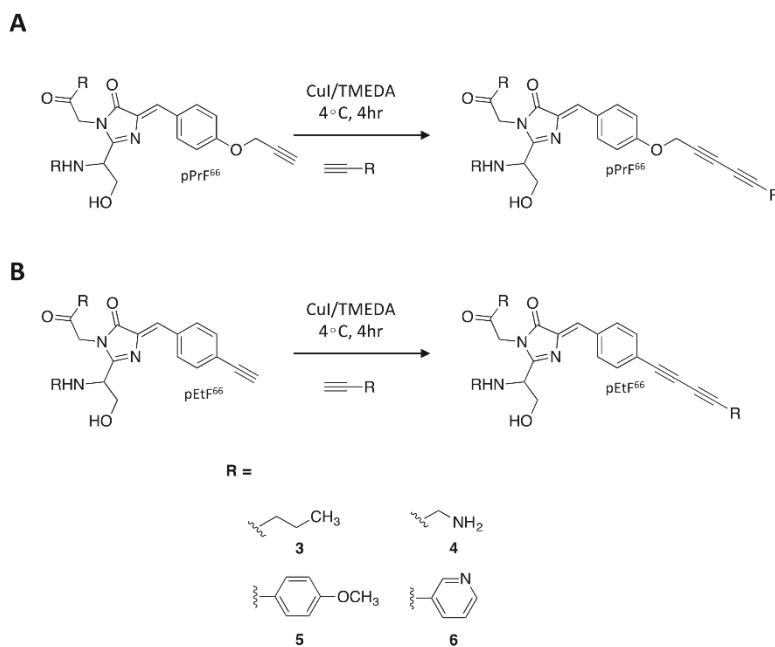
Scheme 6.1: Glaser-Hay bioconjugations to modify the GFP fluorophore.

and TMEDA in the presence of alkyne-UAA bearing GFP$_{TAG66}$ and the cognate alkynyl partner (Scheme 6.1). Reactions proceeded for 4 h at 4 ˚C and then were purified via centrifugation with a molecular weight cut-off column. The protein was placed in phosphate-buffered saline solution (pH 7.2) for analysis using fluorescence spectroscopy. Gratifyingly, our initial attempts to couple the terminal alkynes to the fluorophore were successful. Furthermore, the different characteristics of the installed alkyne moieties successfully shifted the fluorescence profile away from the parental alkynyl-GFPTAG66 spectra, each in a unique way. Control reactions in the presence of the alkyne but the absence of either the CuI/TMEDA or the soluble alkyne afforded no spectral shifts, confirming that the protein modification is indeed due to the coupling reaction. While it might be expected that the requisite for the reaction to occur within the β-barrel of GFP may hinder this reaction from occurring, we hypothesize that the hydrophobic nature of the interior of GFP actually aided in the hydrophobic alkyne localization, thereby facilitating the reaction by increasing the effective concentration. Additionally, SDS-PAGE analysis with Coomassie revealed that

93

the Glaser-Hay reaction only minimally altered protein concentration suggesting only minimal protein degradation (see materials and methods section).

We found that our initial Glaser-Hay reactions on *p*PrF-GFP~TAG66~ had different effects on the fluorescent profile of GFP (Figure 6.5A). Reacting 1-hexyne (**3**) on the chromophore caused a general broadening and quasi-red shift of the fluorescent spectra.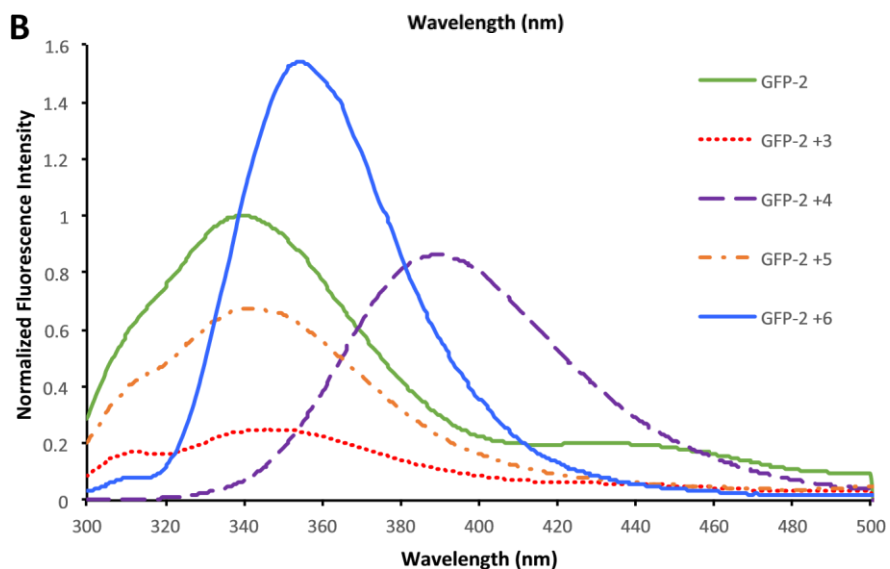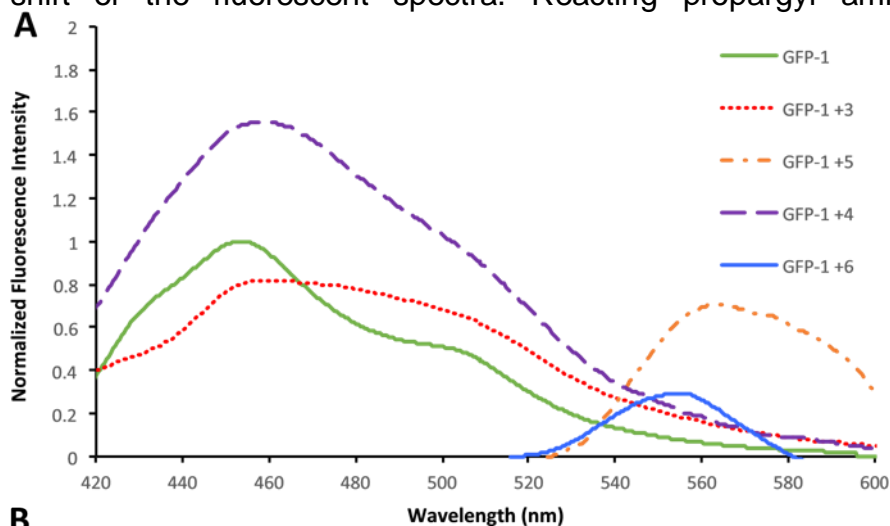 Reacting propargyl amine (**4**) on the chromophore caused a slight band broadening, as well as potential increase in fluorescence intensity, perhaps due to the increased polarity of the introduced amine group. Interestingly, coupling with an aromatic alkyne resulted in a dramatic red-shift of the

Figure 6.5. Fluorescent profile obtained after reacting various alkyne containing partners with *p*Prf or *p*EtF-containing GFPTAG66. (A) Fluorescent spectra for Glaser-Hay modified *p*PrF-GFP chromophore. (B) Fluorescent spectra for Glaser-Hay modified *p*EtF-GFP chromophore

94

fluorescence to above that of wild-type GFP. Both ethynylanisole (**5**) and ethynylaniline (**6**), resulted in excitation spectra maxima above 540 nm, dramatically altering the fluorescence of GFP.

We next sought to explore the effects of reacting terminal alkynes in direct conjugation with the aromatic ring of residue 66. This was feasible with the GFP mutant harboring **2**. Interestingly, this strategy resulted in an even greater blue shift of the pEtF-GFP$_{TAG66}$ compared to both the pPrF and wild-type variants, likely due to the increased conjugation of the direct attachment of that terminal alkyne on the phenyl ring. As a result of this shift, a different excitation wavelength was necessary, as 395 nm was found to no longer excite the pEtF-containing chromophore. Based on absorption experiments, we selected 280 nm as the wavelength to excite the pEtF-GFP$_{TAG66}$ and all its Glaser-Hay derivatives. In the same fashion as the pPrF, the bioorthogonal Glaser-Hay was performed on pEtF-GFP$_{TAG66}$ using the same reaction partners. Once again, 3 was found to broaden the fluorescence spectra. Interestingly, **4** had a drastic red-shift relative to the pEtF parent chromophore. We believe this helps validate our initial speculation that the polarity of the amine has a drastic impact on the fluorescent properties of the chromophore, as in this instance the whole system is in direct conjugation. Interestingly, when employing the aromatic alkynes in the fluorophore modulation, a less dramatic effect was observed than with the pPrF mutants. Reaction with **5** only slightly red-shifted the spectra; however, **6** had a more significant impact both on the intensity and the redshifting of the fluorophore. Additionally, attempts to repeat the experiments using the bromo-

alkyne derivative of **2** under Cadiot-Chodkiewicz coupling conditions resulted in the identical spectra, but were performed under more mild reaction conditions (Figure 6.6). These results were expected as the final products of both the Glaser-Hay or Cadiot-Chodkiewicz reactions are identical.[20] This represents a viable alternative reaction to these protein modification approaches.
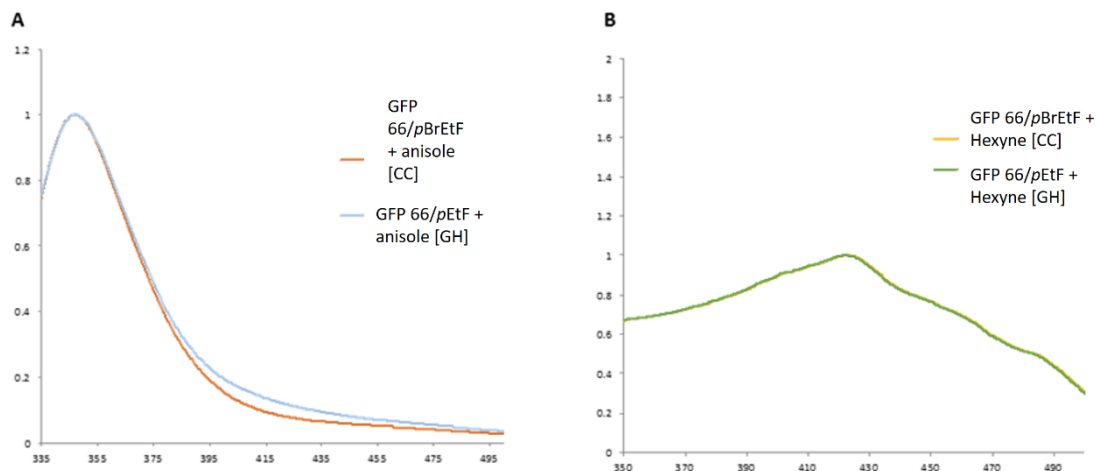

Figure 6.6: Comparison of Glaser Hay couplings and Cadiot-Chodkiewicz couplings with anisole (A) and hexyne (B) resulting in similar spectral shifts due to the synthesis of identical products.

In conclusion, we have extended our work on the biological Glaser-Hay to utilize the bioorthogonal chemistry to modulate protein function. Using two previously reported alkyne-containing UAAs within the chromophore of GFP (position 66), we have successfully performed the Glaser-Hay reaction on the chromophore of GFP. The resulting diyne linkage alters the fluorescence profile of GFP depending on the moiety attached to the terminal alkyne. Our future work seeks to extend the reaction to additional aromatic containing alkynes, which we hope will have a greater impact on GFP fluorescence due to the increased conjugation found in aromatic systems. Our findings highlight the potential of bioorthogonal chemistry, particularly diyne-forming chemistries, to modulate protein function without the need for tedious selections and screens.

Materials and Methods

General: Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Reactions were conducted under ambient atmosphere with non-distilled solvents. Unnatural amino acids were prepared according to literature protocols.[20] NMR data was acquired on an Agilent 400 MHz. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit. Fluorescence data was measured using a PerkinElmer LS 55 Luminescence Spectrometer.

General Biological Glaser-Hay Protocol: To 1.5 mL Eppendorf tube was added 3 μL of a 500 mM CuI solution in water and 3 μL 500 mM TMEDA solution in water. This mixture was then incubated at 4 °C for 10 mins. Following the incubation,

1xPBS (22 μL) was added, followed by the alkyne containing GFP$_{TAG66}$ (10 μL) in PBS (~0.5mg/mL) and a 40 mM solution (4 μL) of the cognate alkyne in DMSO. The mixture was allowed to react at 4°C for 4 hr.

Protocol for Fluorimetry Scans: After reacting for 4 hr, 10 μL of the reaction mixture was added to a quartz cuvette and diluted up to 3 mL with PBS. This was then excited at either 395 nm (for the *p*PrF-containing chromophore) or 280 nm (for the *p*EtF-containing chromophore) with a 10 nm slit width for the excitation and emission wavelengths. Slit widths were increased or decreased as necessary depending on the intensity of the reaction product's signal. The scan speed was set at 500 nm/min.

General Protocol for Biological Glaser-Hay with Aromatic Reaction Partners:
To 1.5 mL Eppendorf tube was added 3 μL of a 500 mM CuI solution in water and 3 μL 500 mM TMEDA solution in water. This mixture was then incubated at 4 °C for 10 mins. Following the incubation, 1xPBS (22 μL) was added, followed by the alkyne containing GFP$_{TAG66}$ (10 μL) in PBS (~0.5mg/mL) and a 40 mM solution (4 μL) of the cognate alkyne in DMSO. The mixture was allowed to react at 4°C for 4 hr. Following this, the unreacted aromatic alkyne was washed away using a 10 MWCO spin column (Corning) and rinsing with 50 μL portions of PBS 8 times The solution was then concentrated to ~25 μL, as indicated on the spin column. Of this cleaned solution, 10 μL was placed into a quartz cuvette and diluted to 2 mL with PBS for fluorescence analysis.

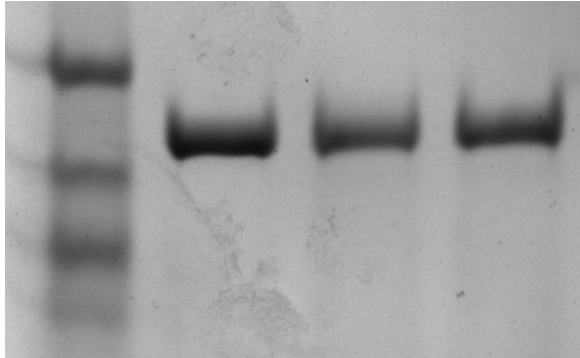|	L GFP Cu/TMEDA	1 WT -	2 *p*PrF -	3 *p*PrF +

Figure 6.7: Sample Glaser-Hay bioconjugation reaction with 1-hexyne. No protein degradation was observed between protein not subjected to reaction conditions (lane 2) and reacted protein (lane 3).

References

1. Wu A, Senter P. Arming antibodies: prospects and challenges for immunoconjugates. Nat Biotechnol. 2005;23:1137–1146.

2. Brustad EM, Arnold FH. Optimizing non-natural protein function with directed evolution. Curr Opin Chem Biol. 2011;15:201–210.

3. Zhu H, Snyder M. Protein chip technology. Curr Opin Chem Biol. 2003;7:55–63.

4. Tan W, Sabet L, Li Y, et al. Optical protein sensor for detecting cancer markers in saliva. Biosens Bioelectron. 2008;24:266–271.

5. Link AJ, Mock ML, Tirrell DA. Non-canonical amino acids in protein engineering. Curr Opin Biotechnol. 2003;14:603–609.

6. Banghart M, Volgraf M, Trauner D. Engineering light-gated ion channels. Biochemistry. 2006;45:15129–15141.

7. Steen E, Redeker DT, Ta D, Cortens B, Billen W, Guedens P. Adriaensens, protein engineering for directed immobilization. Bioconjug Chem. 2013;24:1761–1777.

8. Presta L. Engineering of therapeutic antibodies to minimize immunogenicity and optimize function. Adv Drug Deliv Rev. 2006;58:640–656.

9. Young TS, Schultz PG. Beyond the canonical 20 amino acids: expanding the genetic lexicon. J Biol Chem. 2010;285:11039–11044.

10. Liu C, Schultz P, Kornberg R, Raetz C, Rothman J, Thorner J. Adding new chemistries to the genetic code. Annu Rev Biochem. 2010;79:413–444.

11. Sletten E, Bertozzi C. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed. 2009;48:6974–6998.

12. Maza J, Jacobs T, Uthappa D, Young D. Employing unnatural amino acids in the preparation of bioconjugates (vol. 27, p. 805, 2016). Synlett. 2016;27. e6-e6.

13. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angew Chem Int Ed Engl. 2002;41:2596–2599.

14. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG. Bioconjugation by copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. J Am Chem Soc. 2003;125:3192–3193.

15. Agard NJ, Prescher JA, Bertozzi CR. A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. J Am Chem Soc. 2004;126:15046–15047.

16. Baskin JM, Prescher JA, Laughlin ST, et al. Copper-free click chemistry for dynamic in vivo imaging. Proc Natl Acad Sci USA. 2007;104:16793–16797.

17. Glaser C. Ber Dtsch Chem Ges. 1896;2:422–424.

18. Shi W, Lei A. 1,3-Diyne chemistry: synthesis and derivations. Tetrahedron Lett. 2014;55:2763–2772.

19. Maza JC, McKenna JR, Raliski BK, Freedman MT, Young DD. Synthesis and incorporation of unnatural amino acids to probe and optimize protein bioconjugations. Bioconjug Chem. 2015.

20. Maza JC, Nimmo ZM, Young DD. Expanding the scope of alkyne-mediated bioconjugations utilizing unnatural amino acids. Chem Commun. 2016;52:88–91.

21. Lampkowski JS, Villa JK, Young TS, Young DD. Development and optimization of Glaser-Hay bioconjugations. Angew Chem Int Ed Engl. 2015.

22. Pakhomov A, Martynov V. GFP family: structural insights into spectral tuning. Chem Biol. 2008;15:755–764.

23. Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the Aequorea victoria green fluorescent protein. Science.1996;273:1392–1395.

24. Craggs T. Green fluorescent protein: structure, folding and chromophore maturation. Chem Soc Rev. 2009;38:2865–2875.

25. Young D, Jockush S, Turro N, Schultz P. Synthetase polyspecificity as a tool to modulate protein function. Bioorg Med Chem Lett. 2011;21:7502–7504.

26. Maza J, Villa J, Landino L, Young D. Utilizing unnatural amino acids to illustrate protein structure-function relationships: an experiment designed for an undergraduate biochemistry laboratory. J Chem Edu. 2016;93:767–771.

27. Deiters A, Schultz PG. In vivo incorporation of an alkyne into proteins in Escherichia coli. Bioorg Med Chem Lett. 2005;15:1521–1524.

28. Young D, Young T, Jahnz M, Ahmad I, Spraggon G, Schultz P. An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. Biochemistry. 2011;50:1894–1900.

Chapter 7: Utilizing 3-Fluorotyrosine to Probe the Radical Pathway of Dehaloperoxidase

Dehaloperoxidase (DHP) is a multifunctional hemoglobin from the terebellid polychaete *Amphitrite ornata.* In addition to the standard globin function of oxygen transport, DHP can also preform peroxidase, oxidase and peroxygenase functions.[1,12] These functions are used by *A. ornata* to detoxify a wide range of chemicals released by nearby organisms, making DHP an exciting target for use as a bioremediation enzyme. Additionally, the ability to carry out these four functions presents a unique biochemical situation to researchers. Not only is this an unusual number of abilities for a single protein, but two functions, globin and peroxidase, have very different functional requirements. The globin function utilizes ferrous iron (Fe^{2+}), while peroxidase mechanism relies on ferric (Fe^{3+}) and ferryl (Fe^{4+}) iron.[1] The exact mechanisms responsible for the transitions between these redox states and conflicting functions are still unclear, but are attributed to interactions with product radicals and reactant molecules as well as the generation of protein tyrosyl radicals. [2,3]

The position of these tyrosyl radicals have been studied by Dumarieh and colleagues by comparing computational Electron Paramagnetic Resonance (EPR) spectra simulations and experimentally generated spectra.[2] Further non-computational experiments are needed to confirm these findings, as the nature of computational chemistry leaves room for inaccuracies. One such experimental

approach is to utilize the unnatural amino acid 3-fluorotyrosine (3-FY). Unnatural amino acids (UAAs) can be incorporated site specifically into proteins and subsequently used as probes or to confer a novel chemical ability.[4] 3-FY and similar halogenated phenols have been successfully used to probe the structure-function relationships in myoglobin and bacterial hemoglobin.[5,6] In the case of DHP, the 3-FY will be used to probe structure-function relationships between the tyrosyl radicals and the environment of the protein. 3-FY will create a unique EPR spectrum from the natural tyrosine, providing the ability to distinguish between a tyrosyl and fluorotyrosyl radical. Characterizing these tyrosyl radicals will allow for further exploitation of the DHP's versatility, whether as a bioremediation catalyst or as a model system for the study of peroxidase and other detoxifying functions.

Discovery and Characterization of DHP

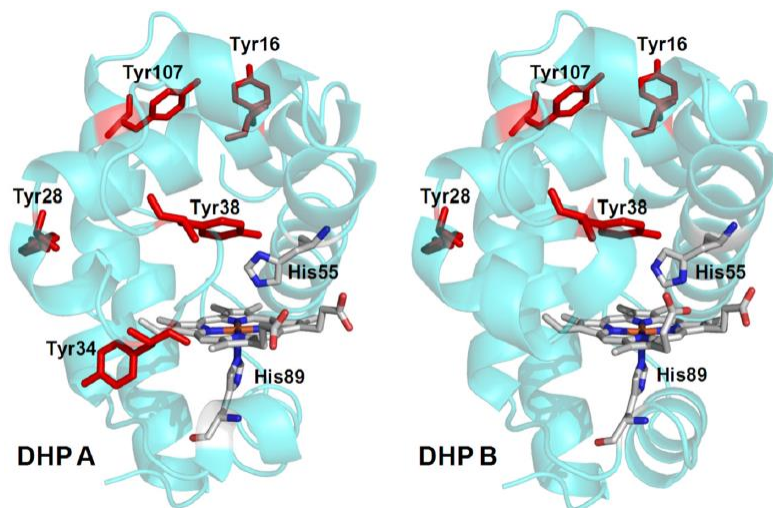*A. ornata* lives near several polychaetes and hemichordates that produce toxic halogenated compounds in order to repel predators. *A. ornata* deals with these compounds using the A and B isozymes of a dehalogenating enzyme dubbed Dehaloperoxidase (Figure 7.1).[2] This protein oxidizes toxic trihalophenols (TXP) to
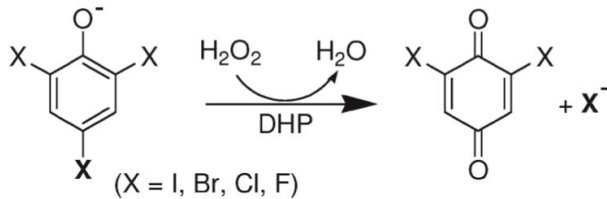

Figure 7.1: Isozymes DHP A and DHP B.[2]

Figure 7.2: Oxidation Reaction via DHP Peroxidase Mechanism.[7]

less toxic dihaloquinones (DXQ) using a peroxidase mechanism (Figure 7.2).[7] DHP makes up 3 % of the soluble protein in *A. ornata*, indicating its importance as both a globin and detoxifying compound.[8] DHP was discovered in 1996 when the Woodin lab examined *A. ornata's* proteome to find the protein responsible for the worm's immunity to bromophenols.[9] DHP exhibited a high activity in this study.[9] Dawson et. al. found the genes responsible for producing the two isozymes in 2004.[9] DHP A contains five more amino acids than DHP B, including a Y34 residue.[2,9] The two isozymes of DHP employ tyrosyl radical formation as a part of the peroxidase mechanism. The radicals form first on either the Y38 or Y34. If the radical forms on Y38, it will eventually transfer to the Y28.[2] The two isozymes exhibit the same radical formation at Y28 and Y38, but only the slightly larger isozyme DHP A possesses a Y34 residue, and thus is the only one that can form the Y34 radical.[2]

The protein's overall structure is very similar to other globins. In fact, when a sperm whale myoglobin structure is superimposed upon it, the alpha carbons are only an average of 1.8 Å apart. This distance means that these two proteins are as similar as distantly related globins, indicating that DHP most likely evolved from a globin.[10] The protein crystallizes as a dimer but is present in solution as a monomer, which originally caused some confusion as to whether or not it was the coelomic hemoglobin of the worm.[8,9] Others have found that DHP has many

105

characteristics that are halfway between those typical of globins and peroxidases, such as Fe-His55 distance, imidizolate character, and catalytic speed for the peroxidase mechanism (for which it is faster than Myoglobin (Mb), but slower than Horse Radish Peroxidase (HRP)).[3,8,11] Woodin also found that DHP has the ability to dehalogenate a wide range of compounds, which is supported by recent studies on the peroxygenase and oxidase functions of the enzyme.[9,12]

Mechanisms Involved in DHP Function

DHP is capable of carrying out four functions: globin, peroxidase, oxidase and peroxygenase. This unusual versatility raises several mechanistic questions, such as how the protein switches between the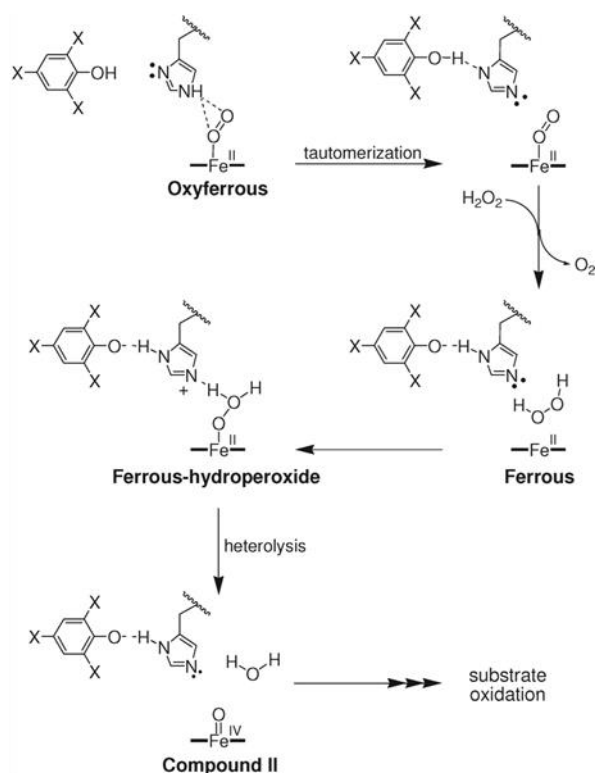 globin and peroxidase functions. It is known that both the substrate, trihalophenol (TXP), and H_2O_2 must be present to facilitate the functional switching between the two functions, but there is disagreement about the exact mechanism. The Ghiladi Lab hypothesizes that the binding of TXP in the presence of H_2O_2 is responsible for the initiation of the catalytic cycle (Figure 7.3). They base this theory on precedent in Kat G, HRP and other peroxidases, as well as a lack of any
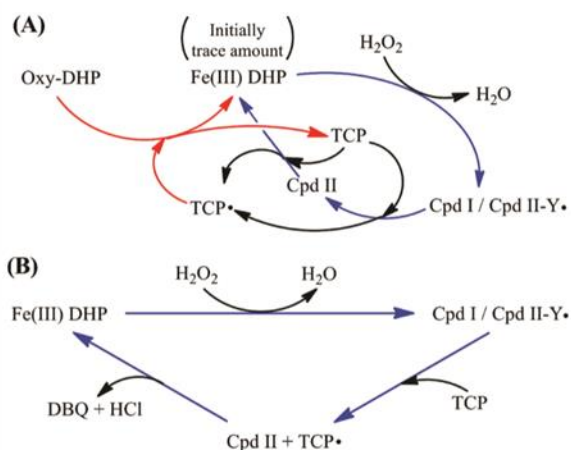

Figure 7.3: Substrate Binding Mechanism of Activation.[7]

106

Figure 7.4: Substrate Radical Activation.[3]

TXP radical species in multiple spectroscopic studies.[7] Meanwhile, the Dawson Lab proposes that trace amounts of ferric protein oxidize TXP to the TXP radical. This TXP· then converts the majority of the DHP from oxyferrous (Fe^{2+}) to the ferric (Fe^{3+}) form (Figure 7.4). Only once all of the protein has been converted to the ferric form does the catalytic cycle start, which accounts for the observed lag time between TXP addition and the initiation of the catalytic cycle.[3]

Another question raised by this unusual protein concerns the protein's ability to circumvent the classical Poulos and Kraut peroxidase mechanism in order to begin and end in the oxyferrous state.[1] The Poulos and Kraut pathway is modeled on Cytochrome C Peroxidase. First the RO_2H binds to the resting state (7.5a).[13] Then an intermediate is formed (7.5b) and next an activated transition state. This transition state contains a negatively charged oxygen 1, which is stabilized by the positive Arg48. The transition state then undergoes a heterolytic cleavage to form the leaving group ROH. This leaves behind a single oxygen attached to the Fe^{3+} (7.5d), which is in resonance with the Fe^{4+} protein (7.5e). These two states are referred to as Compound I. Next, the negatively charged oxygen atom abstracts a hydrogen from the nearby Trp51, forming a protein radical
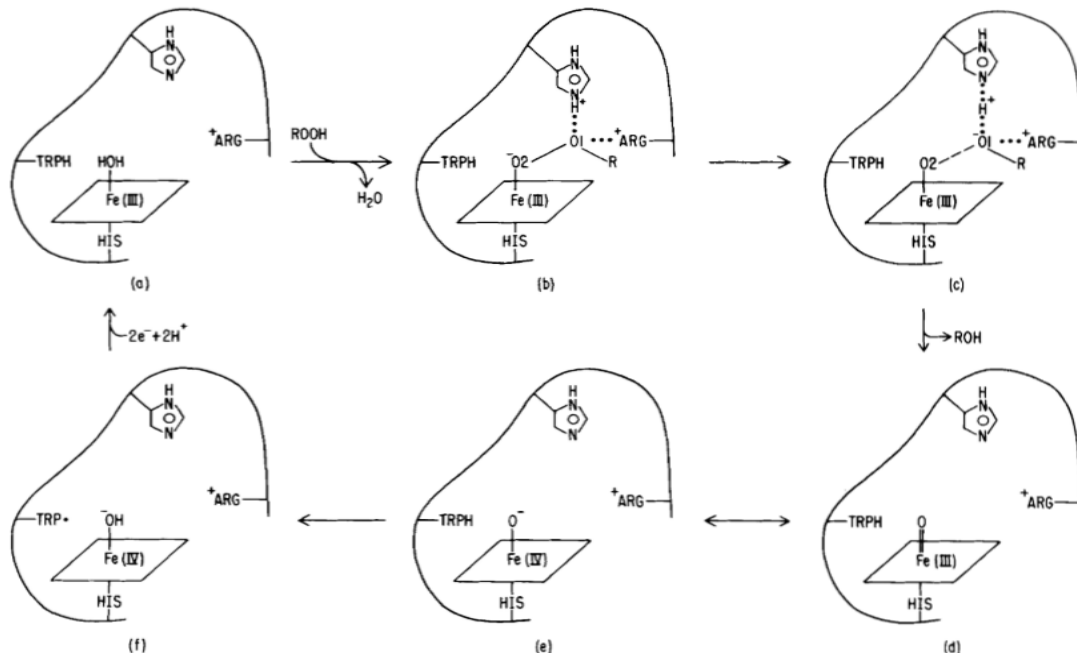
107

Figure 7.5: Poulos and Kraut Peroxidase Reaction Pathway.[13]

(7.5). Finally, two one-electron oxidations return the protein to its original Fe^{3+} state.[13] In most peroxidases, the ferric (Fe^{3+}) and ferryl (Fe^{4+}) states are essential, but the oxyferrous (Fe^{2+}) state is catalytically inactive. This presents a problem for DHP, whose primary function is to shuttle oxygen in the oxyferrous state.[1,14]

To solve this problem, DHP has evolved the ability to oxidize substrates starting from the oxyferrous state, as well as participate in a reaction with DXQ to return to the oxyferrous state upon completion of the catalytic cycle (Figure 7.6). After initiation of the catalytic cycle, TXP and H_2O_2 bind to DHP. H_2O_2 binds to the heme in the distal pocket causing the flexible, distal His55 to swing inward and form a hydrogen bond with the H_2O_2. In high concentrations, TXP will be present in the internal binding site above the α-edge of the heme, and at lower concentrations, it will bind to the suggested external site above the β- or δ-edge of
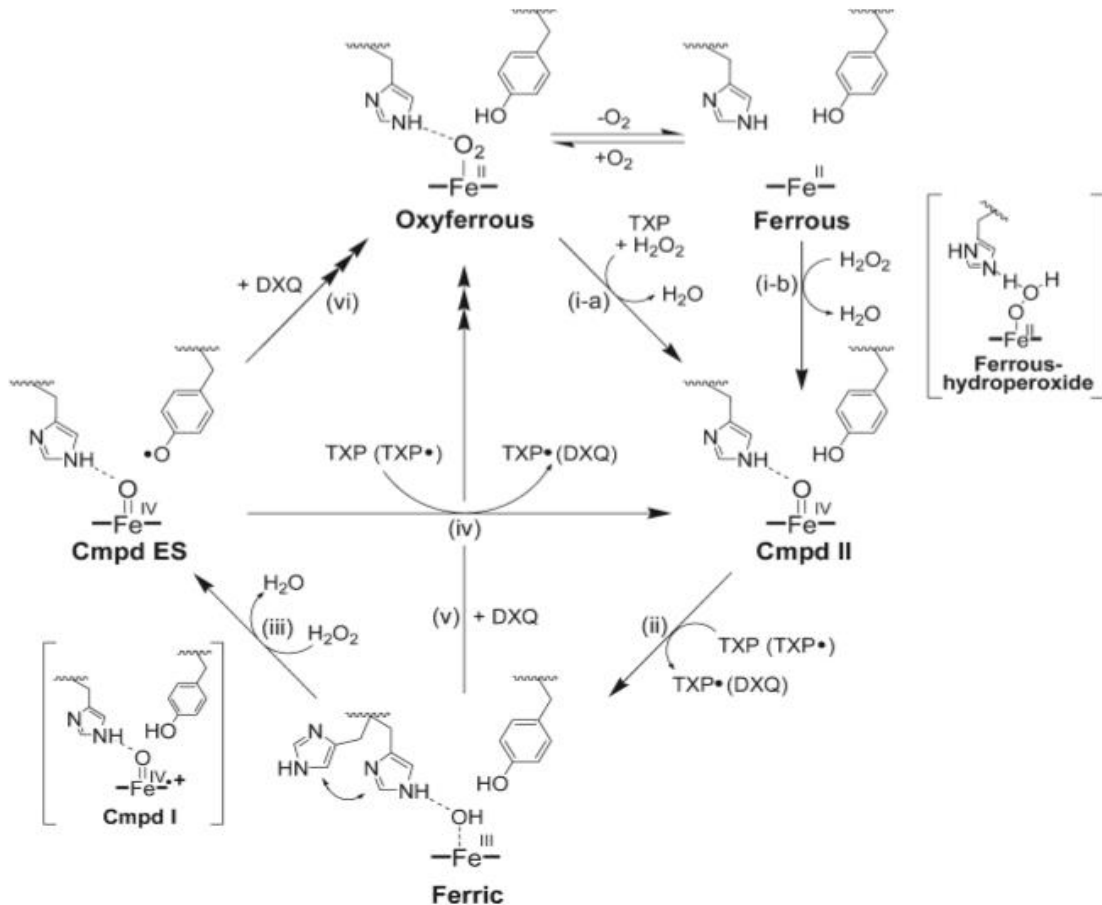
Figure 7.6: Total Peroxidase Cycle for DHP.[7]

the heme.[8] The histidine then acts as a donor base to allow for the cleavage of H_2O_2. This results in the formation of water and leaves an oxygen molecule bound to the Fe^{4+}, exactly like the Poulos and Kraut mechanism. In DHP this state is known as Compound II and is similar to Compound I in the Poulos and Kraut mechanism.[7,13] This form of the protein can then oxidize the TXP to TXP· and then to a dihaloquinone (DXQ), resulting in the ferric form of the enzyme. The ferric DHP can then be either converted either back to the oxyferrous form via DXQ or to the intermediate Compound I and then to Compound ES via excess H_2O_2. Compound ES contains a tyrosyl radical, similar to the tryptophenyl radical seen

in Figure 6f of the Poulos and Kraut mechanism.[7,13] The intermediate Compound I contains a porphyrin radical which is then converted to a series of tyrosyl radicals in the protein. Compound ES can then revert to the oxyferrous state via a reaction with DXQ, or can convert TXP radicals to DXQ to return to Compound II.[7]

Further separating DHP from typical peroxidases is its lack of the catalytic Arg-Fe-His triad which aids in the heterolytic cleavage of peroxides. It can be made up for by the proximal Methionine 86 or the Leu83-His89-Fe triad present in DHP. Met86 is near the proximal Histidine 89. Both functional groups polarize the Fe-O bond somewhat, although not as strongly as the Asp-His-Fe triad.[8] This mechanism can be inhibited by monohalogenated phenols, which bind in the distal pocket and force the H55 into the solvent exposed configuration, preventing the cleavage of H_2O_2 and the subsequent initiation of the peroxidase mechanism.[11] This dependence indicates the importance of the flexible H55.

Additional Capabilities of DHP

In addition to the globin and peroxidase functions, several new catalytic abilities and substrates have been identified for DHP, expanding its potential for bioremediation. It can oxidize hydroquinone (HQ) to 1,4-benzoquinone (1,4-BQ).[14] HQ acts as a competitive inhibitor, as it is preferentially oxidized over the native substrate, TXP. This could also account for the observed lag phase.[14] DHP can also oxidize toxic hydrogen sulfide (HS) using the highly flexible H55, indicating that DHP could be capable of many more detoxifying functions.[15]

110

Peroxygenase and concurrent oxidase functions have also been observed in DHP (Figure 7.7). The difference between peroxidase and peroxygenase chemistry is the origin of the oxygen used for oxidation. In the peroxidase function,
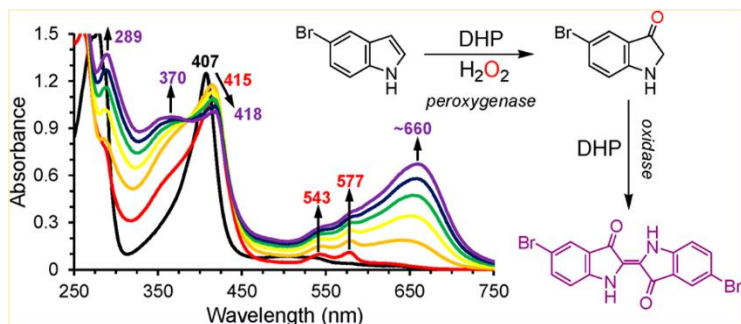

Figure 7.7: UV-Vis and Reaction of peroxygenase and oxidase functions.[12]

the oxygen comes from solvent water, and in the peroxygenase function the oxygen comes from the H_2O_2.[12,16] In the peroxygenase reaction, monohalindoles are oxidized to oxindolenine. The subsequent oxidase function then utilizes O_2, to form indigos. The presence of this function indicates that DHP can utilize other pathways besides the TXP· or TXP binding to switch from the globin to a catalytic function, and other pathways besides reactions with DXQ to return to the oxyferrous state.[12] These abilities provide the potential for other native detoxifying functions in DHP.[12]

Identification of Protein Radicals

The tyrosyl radicals at Y28, Y34 and Y38, were identified by Dumarieh et al. using a combination of EPR spectroscopy, site specific natural amino acid mutagenesis and computer simulations.[2] Initially, the Tyrosyl Radical Spectra Simulation Algorithm (TRSSA) was used to identify the Y34 and Y38 radicals present in WT DHP A Compound ES. The EPR spectra computed for each radical from the algorithm are shown in Figure 7.8. TRSSA was later used to determine
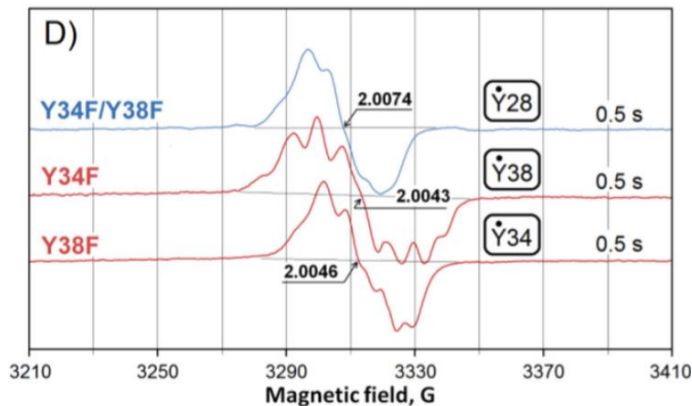
Figure 7.8: ERP Spectra for DHP tyrosyl radicals.[2]

that Y28 was the site of the third observed, previously unidentified radical. Experimental support for these computational results was garnered by systematically mutating the tyrosine(s) in question to phenylalanines for both isozymes of DHP. The existence of the Y28· was supported by the identical EPR spectra of DHP B Y38F and DHP A Y34F/Y38F, where residue 28 was the only potential radical position that remained in the active site to confer this signal. This EPR spectra also helped to confirm time-dependent radical migration from the Y34 to Y28, as it is identical to the spectrum observed after some time for DHP B. The Y38· was confirmed using the DHP B Y28F mutant, whose EPR spectra was similar to the one observed in WT DHP A at pH 5, which had been previously assigned the Y38 radical. The double mutant of DHP B did not exhibit any peroxidase activity, further confirming the importance of the tyrosyl radicals to the function of DHP.[2]

These tyrosyl radical assignments can be further examined through the use of an unnatural amino acid (UAA) probe, 3-fluorotyrosine (3-FY), which exhibits an EPR signal unique from that of tyrosine due to the $+\frac{3}{2}$ spin of the fluorine.[17] This UAA can be synthesized or purchased, then incorporated site specifically into the protein in the place of a tyrosine. Any radicals that form on 3-FY can be

112

distinguished from the natural tyrosyl radicals. This strategy would minimize any structure perturbations observed in phenylalanine mutants, and will allow for the normal radical pathway of the protein and the observation of the evolved tyrosyl radicals. This approach demonstrates the use of UAAs a powerful tool to confirm this radical pathway and to study the effect of experimental conditions on the radical pathway.

This peroxidase mechanism involves the formation of several tyrosyl radicals, which were determined using simulations and EPR spectroscopy of proteins mutated with non-radical forming phenylalanine. Further study of this unusual enzyme can be done through the use of UAAs, in particular 3-FY, which can be incorporated site specifically into DHP to probe the characteristics of the tyrosyl radicals. 3-FY· will yield an EPR signal unique from that of the natural tyrosyl radical, making it a powerful tool to use to study the peroxidase mechanism of DHP in real time.

Results and Discussion

The DHP mutant DNA was generated by QuikChange PCR site directed mutagenesis. The amber stop codon was inserted at positions 34 and 38, tyrosine residues that are involved in the radical pathway in only DHP A or in both isoforms, respectively. Initially, we selected the DHP A Y34TAG mutant for UAA incorporation. DHP containing the 3FY UAA was expressed at 10.5 L using pEVOL-3FY E3 synthetase. This synthetase was used instead of the pEVOL-3FY E11 synthetase because it is more efficient at charging the orthogonal tRNA with

Enzyme	Soret	Q-bands
DHP A Y34Y[3F]	408	505, 637
DHP A WT	407	504, 635

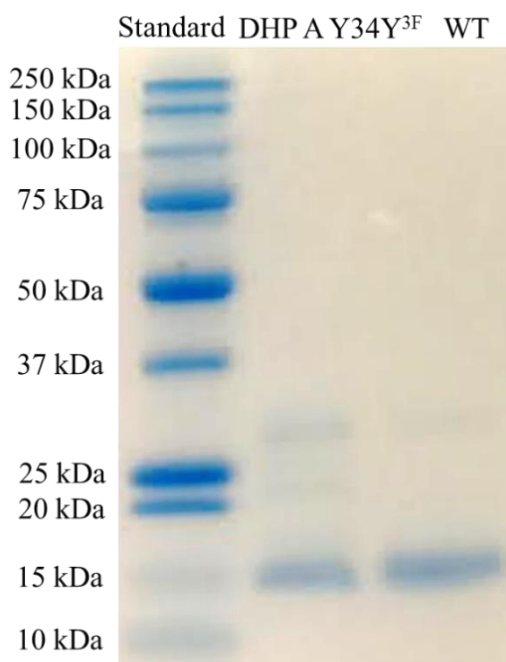Table 7.1: UV-vis characteristics of DHP A Y34Y[3F] and DHP A WT.[2]

Figure 7.9: DHP A Y34Y[3F] and DHP A WT.

3FY.[18] The purity of the expressed mutant was analyzed by UV-Vis and compared to the wild type (Table 7.1). The mutant was run on SDS-PAGE alongside previously expressed WT DHP A (Figure 7.9) to ensure that the mutant protein was full length. A mass spectrum of the mutant protein was performed to confirm that 3FY had been incorporated in the place of tyrosine. The mutant was 18.5 m/z heavier than the wild type at 16,556.0 m/z, the weight of adding a fluorine and subtracting a hydrogen from the wild type at 16,537.5 m/z. The mass spectra and SDS-PAGE both confirmed that the unnatural amino acid had been incorporated at position 34 into a full-length DHP A protein.

The mutant DHP was then assayed for peroxidase activity using established protocols and compared to wild type DHP A, DHP B and DHP A containing a phenylalanine substituted at the 34 position (Table 7.2).[2] The mutant DHP exhibited a slightly lower binding affinity (K_m) than the phenylalanine mutant and the wild type proteins, although the difference is only approximately two-fold. The

Enzyme	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1}µM^{-1})
DHP A Y34Y^{3F}	69.0 ± 15	1.01 ± 0.22	0.015
DHP A Y34F	27 ± 5	0.77 ± 0.02	0.029
DHP A WT	23 ± 1	0.61 ± 0.01	0.027
DHP B WT	22 ± 2	1.53 ± 0.03	0.070

Table 7.2: Kinetic Values for DHP A Y34Y^{3F}, DHP A Y34F, DHP A, and DHP B.[2]

same is true to a lesser extent for the rate (k_{cat}), which is slightly higher in the 3FY mutant than the other DHP proteins. Due to these slightly higher values in the mutant, the 3FY mutant catalytic efficiency (k_{cat}/K_m) value was lower than the other proteins. While these differences are outside of the range of error and statistically significant, they are not large enough to indicate that the 3FY mutant is performing the peroxidase function using a different mechanism than the other DHP proteins. This means that we can continue to stopped-flow studies on the peroxidase mechanism and eventually onto EPR studies to study the iron intermediates and radicals that arise during the peroxidase reaction in our mutant.

The DHP A Y34TAG single mutant was also induced at 10 mL 2,3-FY, 3,5-FY, 2,3,5-FY and 2,3,6-FY, then the protein was subjected to SDS-PAGE (Figure 7.10). Incorporation of these other fluorotyrosines would allow us to probe the radical mechanism more thoroughly, as each would give a unique EPR spectra. Two different fluorotyrosines could be incorporated into DHP A using another synthetase system, allowing for the simultaneous monitoring of all three tyrosine
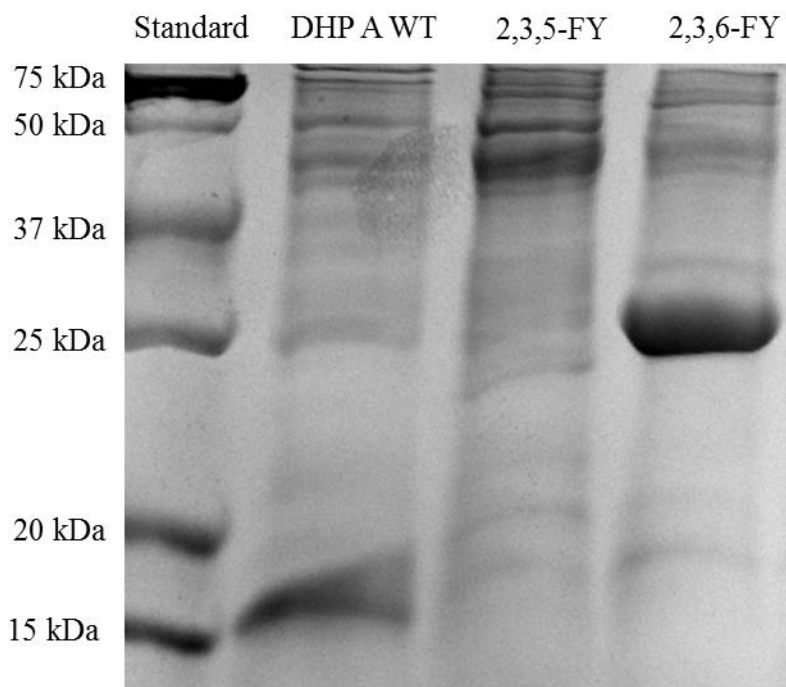
Figure 7.10: DHP A 34 expressed with 2,3,5-FY and 2,3,6-FY compared to DHP A WT.

residues involved in the radical pathway.[19] Only the 2,3,6-FY yielded any appreciable amount of protein, which is surprising considering that the pEVOL-3FY E3 is better at charging the other fluorotyrosines to the orthogonal tRNA.[18] In the future, this DHP A Y34Y^{2,3,6FY} protein will be re-expressed on a larger scale. This mutant will also be analyzed by UV-vis spectroscopy to analyze its purity and subjected to the peroxidase assay to determine what affect the triply substituted tyrosine has on the radical mechanism. Given the low yield of the previous experiment the terminal amber stop codon was replaced with the opal stop codon, TGA, to yield the DHP A Y34TAG double mutant. This was done to prevent insertion of the unnatural at two positions and to make the synthesis of DHP more compatible with *E. coli*, as TGA is used more frequently in *E. coli* than TAG.[20]

Transition of the experiments to the Young lab encountered some difficulty, as even expression of the WT protein became problematic. Thus, the expression of DHP A WT was used to troubleshoot the small-scale expression. Expressions

done by concentrating the induction mixture, inducing the cultures at the end of the log phase (OD_{600} 0.9), and expression of the DHP A 34 single mutant with another UAA, *p*-propargyloxyphenylalanine (*p*PrF), were attempted without any appreciable protein yield. Scaling up the expressions to 200 and 250 mL yielded a red solution with a soret at 416 nm indicating the presence of DHP A WT in the oxyferrous form. However, no protein at the appropriate molecular weight was visible via SDS-PAGE. This, combined with the low intensity of the soret lead to the conclusion that DHP A WT had been expressed at a very low concentration. This protein will be assayed for peroxidase activity to confirm that functional DHP was expressed, then this method will be applied to express DHP A $Y34Y^{2,3,6FY}$ for further analysis.

DHP A Y34TAG has proven challenging to express in good yield, but nevertheless it is possible to express this protein with an unnatural amino acid, in particular 3-fluorotyrosine and 2,3,6-fluorotyrosine. In the future, this expression method will be optimized and extended to the DHP A Y38TAG mutant to further investigate the radical mechanism of DHP. The kinetic values and intermediates of the DHP A 34 and 38 mutant proteins will be analyzed using UV-vis, stopped-flow UV-vis and rapid freeze quench EPR spectroscopy. This investigation will result in an enhanced understanding of the mechanism of this unique multifunctional hemoglobin and offer insights into the mechanisms of other hemoglobins.

Materials and Methods

General: All PCR was performed on a Biorad icycler, with the QuikChange II Kit. Mass spectra were acquired on a Thermo Fisher Scientific Exactive Plus MS orbitrap using a liquid chromatography-heated electrospray ionization source. The mass spectrum was acquired at the NCSU Mass Spectrometry Facility located in the Department of Chemistry. All reagents were obtained from Acros, USB, Biotium, Alfa Aesar or Amresco and used without further purification unless noted. All absorbance data was acquired on a Thermo Scientific Nanodrop 2000.

Quik-Change PCR Protocol: The plasmid harboring DHP A was diluted to 24.7 ng/µL, 12.3 ng/µL, and 6.2 ng/µL. The PCR reaction mixture contained 1 µL of one of the dilutions of the plasmids, 1 µL 10 mM DNTPs, 1 µL of Pfu Polymerase, 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 5 µL Pfu Buffer, and 40µL of ultrapure water. The reaction mixture was subjected to the following heating protocol: 95 °C (1 min); eighteen cycles of melting (95 °C, 30 s), annealing (55 °C, 30 s), and extension (68 °C, 6 min); a final extension period (68 °C, 6 min) and then held at 4 °C. Next the parent plasmid was digested by adding 2 µL of DPN1 [20,000 units/µL] and subjecting the mixtures to the following heating protocol: 37 °C 2 h and heat deactivation (80 °C, 15 min). The reaction was transformed (2 µL) into BL21 DE3 *E. coli* cells by heat shock. The transformed cells were plated (250 µL) onto agar containing ampicillin and incubated at 37 °C overnight. The resulting cells were inoculated and grown in liquid culture overnight, then miniprepped using IBI High Speed Plasmid Mini Kit. The resulting plasmids

were analyzed for successful insertion of the TAG mutation by sequencing at Genewiz. The primers were obtained from IDT DNA Technologies Inc. Y38TAG 5'-GTC AGA TTT GCC GAC CTA GTT TTT GAA GTA GCG-3' (forward), 5'-CGC TAC TTC AAA AAC TAG GTC GGC AAA TCT GAC- 3' (reverse), Y34TAG 5'-CGA CAT AGT TTT TGA ACT AGC GGC GCT CGT-3' (forward), 5'-GAC GAG CGC CGC TAG TTC AAA AAC TAT GTC G-3' (reverse).

Expression and Purification of Protein: The plasmid harboring a variable DHP TAG mutant (0.5 µL) was co-transformed with a pEVOL-3FY E11 plasmid (0.5 µL) into BL21 (DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (0.034 mg/mL), ampicillin (0.050 mg/mL) at 37 °C overnight. Two colonies were then used to inoculate two separate aliquots of LB media (5 mL) containing ampicillin and chloramphenicol. The cultures were incubated at 37 ° C overnight and used to inoculate an expression culture (1.5 L LB media, 1.5 mL 50 mg/mL Amp, 1.5 mL 34 mg/mL Chlor, 10 mL 0.01 mg/mL Hemin) at an OD_{600} 0.1. The cultures were incubated at 37 ° C to an OD_{600} between 0.6 and 0.8 at 600 nm, and protein expression was induced by addition of 3-Fluorotyrosine (15 mL, 100 mM) and 20 % arabinose (1.5 mL) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 1.5 mL). The cultures were allowed to shake at 32 °C for 16-20 h then centrifuged at 4,150 rpm for 20 minutes and stored at -80 °C for 16 h. The cell pellet was re-suspended in a solution of 100 mM sodium phosphate, 20 mM imidazole. The cells were lysed by addition of 1 mg/mL Lysozyme, 10 mg/mL Soybean Trypsin

119

Inhibitor, phenylmethylsulfonyl fluoride (75 mg/mL), Tosyl phenylalanyl chloromethyl ketone (10 mg/mL), Triton x-100 (1 mL) and sonication. The mixture was centrifuged for 1 h at 12,000 rpm. The supernatant was then centrifuged for 30 min at 12,000 rpm. The supernatant was then loaded onto an equilibrated nickel column for affinity chromatography, where the protein was washed with a pH 8 solution of 0.1 M sodium phosphate, 20 mM imidazole, 0.5 M NaCl until the flow through was clear. The protein was then eluted with a pH 8 solution of 0.1 M sodium phosphate and 200 mM imidazole, 0.5 NaCl. Excess potassium ferrocyanide was added to the elution to convert all the protein to the ferric form. The solution was then concentrated to one-fifth of the original solution and buffer exchanged via PD10 column. The resulting protein was then loaded onto an equilibrated CM- 52 column. Minor fractions were collected by washing the column with a solution of 10 mM potassium phosphate at pH 7. Major fractions were then collected by washing the columns with 38 mM potassium phosphate at pH 7. Protein concentrations were determined by UV-Vis measurements at 407 nm. The purity and size of the protein was analyzed by UV-Vis, SDS PAGE (12%) and LCMS.
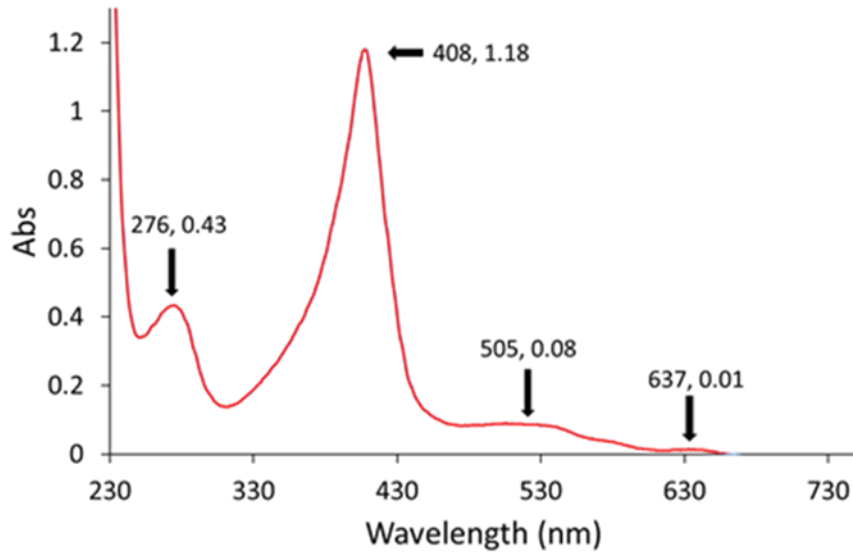
Figure 7.11: UV-vis spectrum of ferric DHP A 34+3FY.

Peroxidase Assay: The purified protein was assayed for peroxidase functionality

to compare kinetic values to the wild type. The protein (0.5 µM) was mixed with
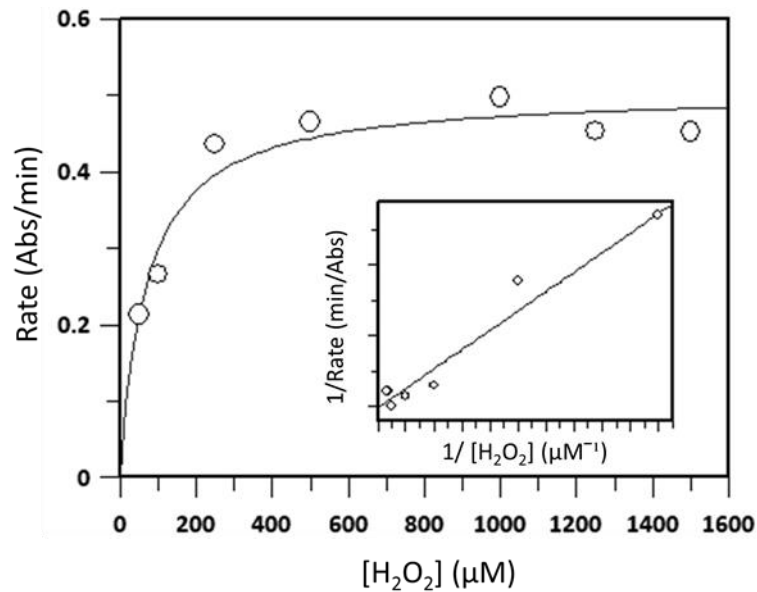

Figure 7.12: Michaelis-Menten plot of DHP A 34+3FY performing the peroxidase reaction.

trichlorophenol (150 µM) and varying concentrations of hydrogen peroxide (50,

100, 250, 500, 1000, 1250 µM) to a total volume of 250 µL in 100 mM Potassium

121

Phosphate Buffer (pH 7). Absorbance readings were taken every 0.6 s at 276 nm and 312 nm.

References

1. Franzen, S.; Thompson, M. K.; Ghiladi, R. A. The dehaloperoxidase paradox. *BBA-Proteins Proteomics* **2012***, 1824*, 578-588.

2. Dumarieh, R.; D'Antonio, J.; Deliz-Liang, A.; Smirnova, T.; Svistunenko, D. A.; Ghiladi, R. A. Tyrosyl Radicals in Dehaloperoxidase: how nature deals with evolving an oxygen-binding globin to a biologically relevant peroxidase. *J. Biol. Chem.* **2013***, 288*, 33470-33482.

3. Du, J.; Sono, M.; Dawson, J. H. Functional Switching of Amphitrite ornata Dehaloperoxidase from O-2-Binding Globin to Peroxidase Enzyme Facilitated by Halophenol Substrate and H2O2. *Biochemistry (N. Y.)* **2010***, 49*, 6064-6069.

4. Wang, Q.; Parrish, A. R.; Wang, L. Expanding the Genetic Code for Biological Studies. *Chem. Biol.* **2009***, 16*, 323-336.

5. Pond, M. P.; Wenke, B. B.; Preimesberger, M. R.; Rice, S. L.; Lecomte, J. T. J. 3-Fluorotyrosine as a Complementary Probe of Hemoglobin Structure and Dynamics: A 19F-NMR Study of Synechococcus sp PCC 7002 GlbN. *Chem. Biodivers.* **2012***, 9*, 1703-1717.

6. Yu, Y.; Lv, X.; Li, J.; Zhou, Q.; Cui, C.; Hosseinzadeh, P.; Mukherjee, A.; Nilges, M. J.; Wang, J.; Lu, Y. Defining the Role of Tyrosine and Rational

Tuning of Oxidase Activity by Genetic Incorporation of Unnatural Tyrosine Analogs. *J. Am. Chem. Soc.* **2015***, 137*, 4594-4597.

7. D'Antonio, J.; Ghiladi, R. A. Reactivity of Deoxy- and Oxyferrous Dehaloperoxidase B from Amphitrite ornata: Identification of Compound II and Its Ferrous-Hydroperoxide Precursor. *Biochemistry* **2011***, 50*, 5999-6011.

8. de Serrano, V.; D'Antonio, J.; Franzen, S.; Ghiladi, R. A. Structure of dehaloperoxidase B at 1.58 angstrom resolution and structural characterization of the AB dimer from Amphitrite ornata. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2010***, 66*, 529-538.

9. Chen, Y.; Woodin, S.; Lincoln, D.; Lovell, C. An unusual dehalogenating peroxidase from the marine terebellid polychaete Amphitrite ornata. *J. Biol. Chem.* **1996***, 271*, 4609-4612.

10. LaCount, M.; Zhang, E.; Chen, Y.; Han, K.; Whitton, M.; Lincoln, D.; Woodin, S.; Lebioda, L. The crystal structure and amino acid sequence of dehaloperoxidase from Amphitrite ornata indicate common ancestry with globins. *J. Biol. Chem.* **2000***, 275*, 18712-18716.

11. D'Antonio, J.; D'Antonio, E. L.; Thompson, M. K.; Bowden, E. F.; Franzen, S.; Smirnova, T.; Ghiladi, R. A. Spectroscopic and Mechanistic Investigations of Dehaloperoxidase B from Amphitrite ornata. *Biochemistry* **2010***, 49*, 6600-6616.

12. Barrios, D. A.; D'Antonio, J.; McCombs, N. L.; Zhao, J.; Franzen, S.; Schmidt, A. C.; Sombers, L. A.; Ghiladi, R. A. Peroxygenase and Oxidase Activities of Dehaloperoxidase-Hemoglobin from Amphitrite ornata. *J. Am. Chem. Soc.* **2014***, 136*, 7914-7925.

13. Poulos, T. L.; Kraut, J. The Stereochemistry of Peroxidase Catalysis. *J. Biol. Chem.* **1980***, 255*, 8199-8205.

14. Zhao, J.; Zhao, J.; Franzen, S. The Regulatory Implications of Hydroquinone for the Multifunctional Enzyme Dehaloperoxidase-Hemoglobin from Amphitrite ornata. *Biochemistry* **2013***, 117*, 14615-14624.

15. Nicoletti, F. P.; Thompson, M. K.; Franzen, S.; Smulevich, G. Degradation of sulfide by dehaloperoxidase-hemoglobin from Amphitrite ornata. *J. Biol. Inorg. Chem.* **2011***, 16*, 611-619.

16. Baciocchi, E.; Gerini, M.; Lapi, A. O-18 incorporation in the oxidation of N-methylcarbazole by lignin peroxidase and a model compound: a mechanistic insight into the oxidative N-demethylation of aromatic tertiary amines. *Chem. Commun.* **2002**, 946-947.

17. Seyedsayamdost, M.; Reece, S.; Nocera, D.; Stubbe, J. Mono-, di-, tri-, and tetra-substituted fluorotyrosines: New probes for enzymes that use tyrosyl radicals in catalysis. *J. Am. Chem. Soc.* **2006***, 128*, 1569-1579.

18. Minnihan, E. C.; Young, D. D.; Schultz, P. G.; Stubbe, J. Incorporation of Fluorotyrosines into Ribonucleotide Reductase Using an Evolved,

Polyspecific Aminoacyl-tRNA Synthetase. *J. Am. Chem. Soc.* **2011***, 133*, 15942-15945.

19. Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli. *Biochemistry (N. Y.)* **2013***, 52*, 1828-1837.

20. Escherichia coli/Codon Usage. Open Wet Ware. http://www.openwetware.org/wiki/Escherichia_coli/Codon_usage (accessed April 20, 2017).