


5-2016

## Utilizing Synthetic Tools to Address Biological Issues

Marshall Scott Padilla  
*College of William and Mary*

Follow this and additional works at: <https://scholarworks.wm.edu/honorsthesis>

 Part of the [Biochemistry Commons](#), [Molecular Biology Commons](#), [Organic Chemistry Commons](#), and the [Other Chemistry Commons](#)

---

### Recommended Citation

Padilla, Marshall Scott, "Utilizing Synthetic Tools to Address Biological Issues" (2016). *Undergraduate Honors Theses*. Paper 918.

<https://scholarworks.wm.edu/honorsthesis/918>

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact [scholarworks@wm.edu](mailto:scholarworks@wm.edu).

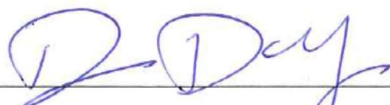
# Utilizing Synthetic Tools to Address Biological Issues

A thesis submitted in partial fulfillment of the requirement  
For the degree of Bachelor of Science in Chemistry from  
The College of William & Mary

by

Marshall Scott Padilla

Accepted for Honors



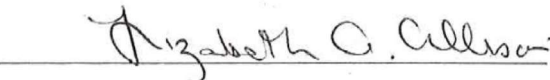
Douglas Young, Committee Chair



Lisa Landino



Robert Hinkle



Lizabeth Allison

Williamsburg, VA

May 3, 2016

# **Utilizing Synthetic Tools to Address Biological Issues**

Marshall S. Padilla  
Newport News, Virginia

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

Chemistry Department

The College of William and Mary

May 3<sup>rd</sup>, 2016

## Table of Contents

Table of Figures	iii
Table of Schemes	iv
Table of Tables	v
Abstract	vi
Chapter 1	1
Chapter 2	14
Chapter 3	31
Chapter 4	46
Chapter 5	60
Chapter 6	68
Chapter 7	78
Acknowledgments	86

## Table of Figures

<i>Figure 1.1</i>	Amino acid general structure	1
<i>Figure 1.2</i>	Translation mechanism	2
<i>Figure 1.3</i>	List of unnatural amino acids	3
<i>Figure 1.4</i>	pEVOL plasmid diagram	5
<i>Figure 1.5</i>	Synthetase selection procedure for an UAA	5
<i>Figure 1.6</i>	Green Fluorescent Protein	7
<i>Figure 1.7</i>	pET-GFP plasmid diagram	7
<i>Figure 1.8</i>	Photocaged serine and photocaged lysine	11
<i>Figure 1.9</i>	Temperature profile of a vessel in a microwave vs oil bath	13
<i>Figure 2.1</i>	Photoswitching of Dronpa and Padron	14
<i>Figure 2.2</i>	Azobenzene derivatives in ion channels	15
<i>Figure 2.3</i>	Flavin and LOV mechanism	16
<i>Figure 2.4</i>	Azobenzene unnatural amino acid	16
<i>Figure 2.5</i>	Emission scans for GFP mutants with <b>1</b> at different residues	18
<i>Figure 2.6</i>	Emission scans for GFP with <b>1</b> at residue 66	19
<i>Figure 2.7</i>	Spatial control experiment of GFP with <b>1</b>	19
<i>Figure 2.8</i>	Azobenzene alkyne unnatural amino acid	20
<i>Figure 2.9</i>	Synthetase assay for <b>5</b>	24
<i>Figure 3.1</i>	Proposed nitro-propargyl tyrosine UAA	31
<i>Figure 3.2</i>	Synthetase assay for <b>11</b>	37
<i>Figure 3.3</i>	SDS-PAGE analysis of GFP containing <b>11</b>	38
<i>Figure 3.4</i>	SDS-PAGE analysis of bioconjugation and decaging of GFP containing <b>11</b> with a fluorophore	38
<i>Figure 3.5</i>	Antibody diagram	39
<i>Figure 4.1</i>	A series of distinct azides	52
<i>Figure 5.1</i>	CoolMate vessel cooling diagram	60
<i>Figure 5.2</i>	Two UAAs incorporated into GFP and bioconjugated with a CoolMate	61
<i>Figure 5.3</i>	SDS-PAGE analysis of click bioconjugation with GFP mutants	61
<i>Figure 5.4</i>	SDS-PAGE analysis of Glaser-Hay bioconjugation with a GFP mutant	63
<i>Figure 6.1</i>	Eight asymmetric polyynes	70
<i>Figure 6.2</i>	Ethyne benzene alcohol asymmetric polyynes with a hexyne capping group	71
<i>Figure 6.3</i>	Emission spectra for anisole polyynes at 330 nm	72
<i>Figure 6.4</i>	Emission spectra for hexyne polyynes at 330 nm	72
<i>Figure 6.5</i>	Emission spectra of a series of asymmetric triynes	73
<i>Figure 7.1</i>	Proposed caged selenocysteine UAA	79
<i>Figure 7.2</i>	First synthetase assay for <b>41</b>	81

## Table of Schemes

<i>Scheme 1.1</i>	ONBY decaging	10
<i>Scheme 2.1</i>	Photoisomerization of azobenzene	17
<i>Scheme 2.2</i>	Mills reaction mechanism	22
<i>Scheme 2.3</i>	First synthesis of <b>5</b>	21
<i>Scheme 2.4</i>	Second synthesis of <b>5</b>	22
<i>Scheme 2.5</i>	Glaser-Hay bioconjugation with <b>5</b>	23
<i>Scheme 3.1</i>	Decaging mechanism of <b>11</b>	32
<i>Scheme 3.2</i>	Total synthesis of <b>11</b>	33
<i>Scheme 3.3</i>	Nucleophilic substitution mechanism to form <b>6</b>	33
<i>Scheme 3.4</i>	Nitration mechanism to form <b>7</b>	34
<i>Scheme 3.5</i>	Reduction mechanism to form <b>8</b>	34
<i>Scheme 3.6</i>	Boration mechanism to form <b>9</b>	35
<i>Scheme 3.7</i>	Microwave malonic ester synthesis mechanism to form <b>10</b>	36
<i>Scheme 3.8</i>	Microwave conversion and deprotection mechanism to form <b>11</b>	36
<i>Scheme 4.1</i>	1,3-dipolar cycloaddition reaction	46
<i>Scheme 4.2</i>	1,3-dipolar cycloaddition mechanism	46
<i>Scheme 4.3</i>	copper-free 1,3-dipolar cycloaddition reaction	47
<i>Scheme 4.4</i>	Trityl chloride protection of propargyl alcohol	51
<i>Scheme 4.5</i>	Immobilization, click reaction, and resin cleaving of propargyl alcohol	52
<i>Scheme 5.1</i>	CoolMate click bioconjugations with GFP containing azide and alkyne UAAs	62
<i>Scheme 5.2</i>	CoolMate Glaser-Hay bioconjugations with GFP containing an alkyne UAA	63
<i>Scheme 6.1</i>	Symmetric Glaser-Hay reaction	68
<i>Scheme 6.2</i>	Non-chemoselectivity of Glaser-Hay reaction	68
<i>Scheme 6.3</i>	Chemoselectivity of solid-support Glaser-Hay	69
<i>Scheme 6.4</i>	Alkyne loading onto resin	69
<i>Scheme 7.1</i>	Synthesis of <b>41</b>	80

## Table of Tables

*Table 1.1*

Several attempted copper-free click reactions

50

## **Abstract**

With the advent of chemoselective reactions and unnatural amino acids (UAAs), the fields of molecular biology and organic synthesis are merging. Researchers are developing synthetic tools and producing small molecules that are able to affect and investigate large biomolecules and complex living systems. This thesis undertakes a survey of synthetic techniques to develop novel tools that can be employed to address a variety of relevant biological questions. Specifically, we are investigating alternatives and improvements to caging groups, including photoreversible azobenzene UAAs and an UAA caging group possessing a bioorthogonal handle. Also, we are developing a novel system to undergo copper-free 1,3-dipolar cycloadditions, utilizing microwave irradiation. Adapting this methodology, we showcase the first transition and metal free bioconjugations utilizing microwave irradiation through CoolMate technology. Finally, we are examining a method to insert a highly reactive selenocysteine amino acid into proteins via the integration of a caging group and unnatural amino acid technologies.



# Chapter One: Introduction

Within the realm of cellular biology and biochemistry, proteins are perhaps the most dynamic of the major macromolecules. Proteins act as the functional units of cells, performing the majority of the cellular chores. Proteins include enzymes which catalyze reactions, antibodies which protect the body from detrimental foreign particles, and transporters which allow certain substances in the cells. Each of the over 100,000 proteins

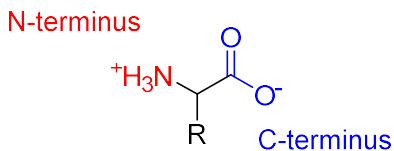


Figure 1.1: General structure of an amino acid. The N-terminus consists of an amino group, and the C-terminus consists of a carboxyl group.

are composed of twenty canonical amino acids in different numbers and combinations.<sup>1</sup> Amino acids have a basic structure of an amino N-terminus and a carboxyl C-terminus, connected by a singular carbon, which contains a functional R-group (Figure 1.1). The R-groups afford amino acid diversity, with the most simplistic being glycine, possessing a hydrogen atom as the R-group. Depending on the R-group, amino acids can have distinct chemical properties, such as polarity, hydrophobicity, and acidity.<sup>1,2</sup> Almost all naturally occurring amino acids are found in the L-configuration (which has the same nomenclature as the S configuration in stereochemistry).<sup>1</sup> Amino acids are connected through a peptide bond, which is the result of a dehydration reaction between the N-terminus of one amino acid with the C-terminus.<sup>1</sup> Conversely, water is used to hydrolyze a peptide bond to form individual amino acids.<sup>1</sup>

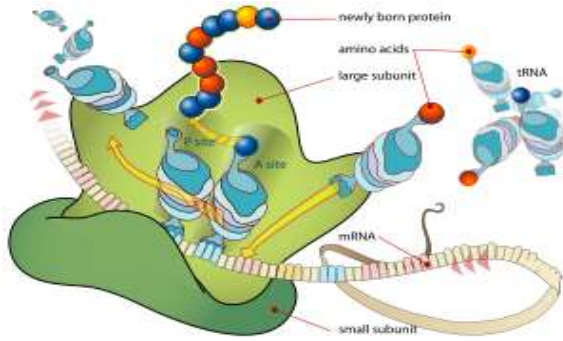


Figure 1.2: Translation of amino acids from mRNA to form polypeptides. Amino acids are carried by tRNAs which have anticodons that bind to codons on the mRNA transcript. This process happens on ribosomes, which are a combination of protein and ribosomal RNA (rRNA).<sup>2,3</sup>

In cells, amino acids are polymerized into polypeptides during translation, where messenger RNA (mRNA) sequences are converted into amino acids via ribosomes and tRNAs. The original mRNA is fabricated using DNA as a template during transcription, where segments of DNA are copied into RNA, spliced into exons and

introns, and then the exons combine to make messenger RNA.<sup>1</sup> Nucleic acids differ in their nitrogenous base groups, where DNA has adenine, guanine, thymine and cytosine, and RNA has adenine, guanine, uracil, and cytosine.<sup>2</sup> Triplet groups of nucleotides, known as codons, correspond to specific amino acids. Due to the number of possible nucleotide configurations and the relatively small number of amino acids, multiple nucleotide triplet combinations result in the same amino acid.<sup>2</sup> This creates degeneracy in the genetic code, which is useful evolutionarily as it minimizes harmful mutations.<sup>2</sup> Amino acids pair with codons through aminoacyl tRNA synthetases (aaRSs) and transfer RNA (tRNA).<sup>1,2</sup> These aaRSs bind tRNAs, which contain an anticodon matching the codon of an mRNA. The specific amino acid also binds to the aminoacyl tRNA synthetase and is covalently linked to the tRNA molecule.<sup>1,2</sup> Synthetases work in conjunction with ribosomes to couple amino acids together to form polypeptides and ultimately proteins (Figure 1.2).<sup>2,3</sup>

## Unnatural Amino Acids

Outside of the naturally occurring amino acids, researchers and genetic engineers have synthesized unnatural amino acids (Figure 1.3).<sup>4,5</sup> Unnatural amino acids (UAAs) allow for increased protein diversity by the addition of non-canonical moieties. For example, an amino acid can be constructed to undergo more complex organic reactions, such as Suzuki couplings and “click” reactions.<sup>6</sup> To date, about one hundred amino acids have been synthesized and successfully integrated into proteins.<sup>7-8</sup>

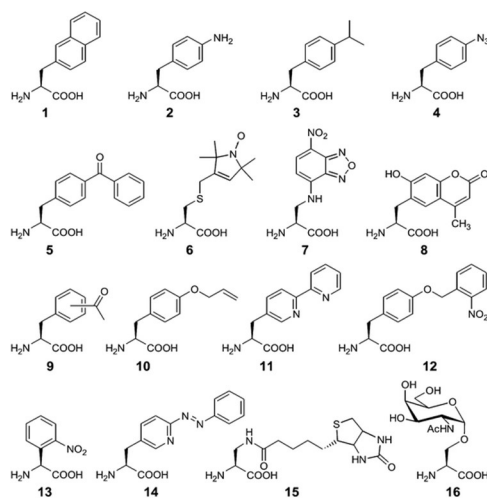


Figure 1.3: A list of unnatural amino acids successfully synthesized. Each UAA has a unique R group, such as iron containing groups, polyaromatics, and terminal alkynes.<sup>5</sup>

UAAs have been synthesized and incorporated into proteins for a variety of purposes. For example, a benzoyl phenylalanine derivative UAA has been incorporated into glutathione S-transferase.<sup>9</sup> Upon irradiation, the benzophenone moiety is able to cross link protein subunits, acting as a photophysical probe.<sup>9</sup> A boronate phenylalanine UAA was co-transformed into an antibody and was found to outcompete wildtype antibodies in binding to acyclic glucamine resins.<sup>10</sup> Unnatural fluorinated amino acids, such as m-fluorophenylalanine, 4-fluoroproline, and trifluoroleucine, have been utilized to direct specific protein-protein interactions, especially with regards to protein folding.<sup>11</sup> These fluorinated UAAs are also useful as probes, as they allow monitoring via <sup>19</sup>F NMR, an analytical technique with high sensitivity relative to cellular environments.<sup>12</sup> Moreover, they alter the

chemical properties of the residue, providing new probes for electron transfer events, and residues with tunable pKa values.<sup>11</sup>

### **Methods for Translation of Unnatural Amino Acids**

The utility of UAAs, while immense, is also challenging. It is rather inefficient to chemically synthesize proteins with the desired amino acid or unnatural amino acid; therefore, the cell's own machinery is often exploited. Bacteria are able to produce proteins in bulk if induced with specific small molecules.<sup>2</sup> While the induction of protein synthesis is rather simple, the actual mechanism for inserting an unnatural amino acid into a protein is very complicated. This is because synthetase/tRNA pairs are generally coupled with a specific amino acid, and therefore will not accept a non-canonical amino acid.<sup>2</sup> One solution to this involves introducing novel exogenous aaRSs that will carry out translation using only the UAA. This method requires that the tRNA and the tRNA synthetase be orthogonal to the endogenous translational machinery, so that no other synthetase will recognize the tRNA for the UAA.<sup>12</sup> Since the genetic code is inherently degenerate, there are multiple sequences for stop codons (UAA, UGA, and UAG), which means that these codons can be exploited for the incorporation of an unnatural amino acid.<sup>13</sup> One advantage in using these stop codons is that they do not encode an amino acid, but rather they signal the termination of translation. Also, only one stop codon is necessary to terminate translation, and if the suppressor tRNA synthetase/tRNA pair is not present, translation simply halts.<sup>1,2</sup>

One important methodology that involves tRNA synthetases is the engineering of optimized pEVOL plasmids to introduce the exogenous translational machinery (Figure 1.4).<sup>14</sup> Bacteria have circular double-stranded DNA sequences called plasmids that are

separate from the bacterial chromosomal DNA.<sup>2</sup> Plasmids usually contain some type of genetic advantage, such as antibiotic resistance, and these plasmids can be transferred amongst bacteria in several different fashions.<sup>2</sup> Researchers have been able to manipulate

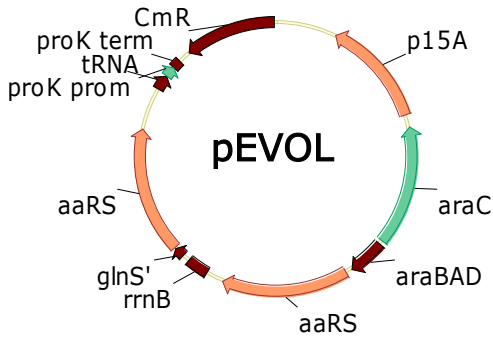


Figure 1.4: A pEVOL plasmid that codes for the production of aaRSs that allow for the amino acid to be translated into a protein.<sup>14</sup> There is also an araBAD that enhances protein expression.<sup>14</sup>

these plasmids via molecular biology techniques to provide a platform for the expression of exogenous genes. In this method, unnatural amino acids are introduced into proteins in *Escherichia coli* using aaRSs and suppressor tRNA pairs evolved from *Methanocaldococcus jannaschii*.<sup>14</sup>

Suppressor tRNA refers to utilizing the amber stop codon (UAG for RNA and TAG for DNA) as the unique codon that attaches to the anticodon (CUA) on the orthogonal tRNA. In this method, these plasmid vectors are linked to antibiotic resistance and protein synthesis induction.<sup>14</sup>

Selection of aaRSs is accomplished via a series of positive and negative selections from a library ( $\sim 10^8$ ) of aaRS active site mutants (Figure 1.5).<sup>4</sup> The positive selection is performed by transforming the cells with the library, as well as a mutated gene encoding chloramphenicol acetyl-transferase (CAT) with a TAG codon.<sup>4</sup> CAT inactivates that antibiotic chloramphenicol, through a catalyzed acetylation. Therefore, when cells are grown in the

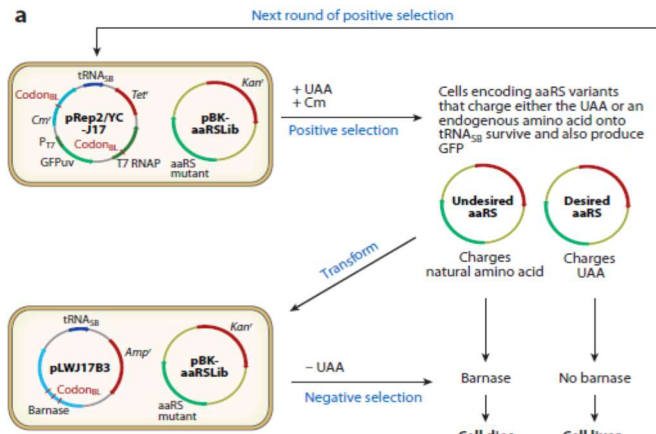


Figure 1.5: Selection mechanism for aaRSs. Positive selection transforms plasmids that charge an UAA or a canonical amino acid onto a tRNA.<sup>4</sup> Negative selection selects against aaRSs that encode canonical amino acids.<sup>4</sup>

presence of chloramphenicol, the cells that possess an aaRS that can bind to the tRNA will produce CAT and survive.<sup>4</sup> This process occurs because the plasmid was mutated to introduce a codon that is complementary to the orthogonal UAA tRNA. Also, the aaRS has been mutated at the tRNA binding site to recognize the AUC anticodon, as well as at the amino acid binding side, to provide UAA recognition. Negative selection proceeds by transforming the same aaRS with a plasmid encoding a mutated barnase gene, a toxic ribonuclease.<sup>4</sup> Cells are grown in the absence of UAAs to select against aaRSs that aminoacylate natural acids, by expressing barnase if the aaRS can read through the TAG codon.<sup>4</sup> This prevents the selected aaRS from incorporating a canonical amino acid.<sup>4</sup> Positive and negative selections are repeated multiple times to ensure an aaRS is evolved to specifically incorporate the desired UAA.

With the production of a variety of complex UAAs, creating and tuning individual novel aaRSs for individual unnatural amino acids can prove to be cumbersome. One alternative method is creating a synthetase that is able to encode a variety of amino acids.<sup>15</sup> These permissive synthetases allow for multiple UAAs to be integrated into proteins, rather than requiring a unique aaRS for each UAA.<sup>15</sup> One example of this is the *p*-cyanophenylalanine aminoacyl-tRNA synthetase (*p*CNF-RS), which is able to incorporate at least eighteen different UAAs in response to a TAG codon.<sup>15</sup> These UAAs include trifluoroketone, alkynyl, and hydrazine UAAs.<sup>15</sup> Another important facet for these unnatural aaRS's is that the twenty canonical amino acids cannot be utilized by this synthetase, providing excellent bioorthogonality.<sup>15</sup>

## Green Fluorescent Protein and Applications

One model system for unnatural amino acids is Green Fluorescent Protein (GFP) expression, a fluorescent protein found in the jellyfish *Aequoria victoria*. GFP is composed of a beta barrel with a tyrosine chromophore at residue 66 (Figure 1.6).<sup>16</sup> Residues 65-67 (Ser-Tyr-Gly)

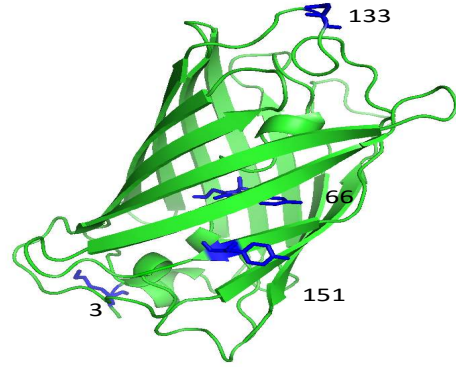


Figure 1.6: Green Fluorescent Protein, with various residues shown within and around the beta barrel.<sup>16</sup>

are theorized to undergo a condensation reaction resulting in a chromophore.<sup>16</sup> An imidazolinone is formed by nucleophilic attack by the glycine amide on the carbonyl residue of serine, and this is followed by a dehydration reaction.<sup>16</sup> It is then hypothesized that molecular oxygen dehydrogenates the  $\alpha$ - $\beta$  bond of tyrosine, allowing conjugation of the phenol ring to the imidazolinone.<sup>16</sup>

GFP provides a powerful tool for analysis of UAAs, especially with regards to monitoring the ability of synthetases to incorporate unnatural amino acids.<sup>14</sup> Competent bacteria are transformed with a pEVOL plasmid, with the respective aaRS/tRNA pair, as

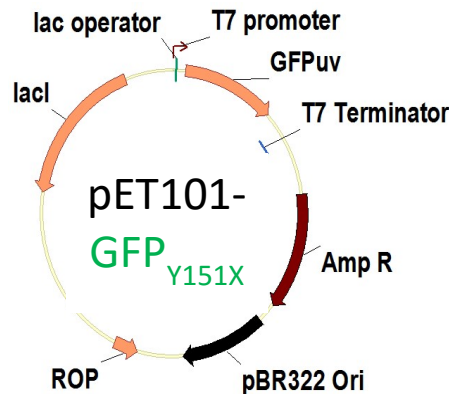


Figure 1.7: A TAG plasmid that codes for production of GFP.<sup>14</sup> Also, the plasmid incorporates a lac section that enhances protein expression, as well as ampicillin resistance.<sup>14</sup>

well as a plasmid with a GFP TAG mutation (Figure 1.7).<sup>14</sup> These mutations allow bacterial cells to transcribe and translate Green Fluorescent Protein. The TAG mutation is associated with a specific residue e.g. GFP 151 TAG and GFP 66 TAG, meaning that an unnatural amino acid can be placed specifically at that residue. With this procedure, bacterial cells are able to fluoresce only if the UAA is

incorporated, which indicates the successful suppression of the stop codon with an UAA.<sup>14</sup> Moreover, a comparison can be made between wild type GFP and GFP with an unnatural amino acid to elucidate the effect of the UAA on the fluorescence and ultimately protein function.<sup>14</sup>

## **Synthetic Plasmids**

As previously mentioned, certain segments of engineered plasmids are utilized for regulation of protein expression. In DNA, parts of the genetic material are devoted to modulating transcription.<sup>2</sup> Portions such as promoters help initiate transcription of a gene, which is recognized by RNA polymerase, the protein that converts DNA to RNA.<sup>2</sup> Operators are segments of DNA where proteins can bind, and typically act as repressors, which limits or stops expression.<sup>1,2</sup> In the pET-GFP plasmid, there exists part of the *lac* operon, which normally regulates lactose metabolism, but here acts to control expression of GFP.<sup>13</sup> One method of inducing GFP expression is with isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), a reagent that binds to the lac repressor protein, resulting in its removal from the operator of the *lac* operon, allowing transcription to occur. RNA polymerase then can generate mRNA, which can be translated to afford GFP (Figure 1.7).<sup>5,13,17</sup> The pEVOL plasmid also contains an *araBAD* promoter that regulates production of the aminoacyl-tRNA synthetase.<sup>5,13,18</sup> In *E. coli*, the *araBAD* operon is used to catabolize arabinose, which is a five carbon monosaccharide used in formation of pectin.<sup>18</sup> P<sub>BAD</sub> is a promoter on the *araBAD* operon that is activated by arabinose itself, and therefore, when arabinose is introduced into a cellular culture, the P<sub>BAD</sub> promoter facilitates production of the aaRS in the pEVOL plasmid (Figure 1.4).<sup>5,13</sup>



One factor that must be addressed when discussing transformation of bacterial cells with aminoacyl-tRNA synthetases/tRNA pairs is ensuring that bacterial cultures have the specific synthetase/tRNA pair. One method to guarantee proper incorporation is associating the engineered plasmid with antibiotic resistance. Generally, the GFP-TAG plasmid contains an ampicillin resistance sequence (Figure 1.7).<sup>5,13,14,15</sup> With this sequence, the bacteria are able to synthesize a  $\beta$ -lactamase enzyme that hydrolyzes  $\beta$ -lactam groups, which is the core structure of ampicillin, and thus deactivating the molecule.<sup>19</sup> When the bacteria are transformed, they are grown on agar plates with ampicillin, and therefore, cells that do not retain the engineered plasmid do not inherit antibiotic resistance and will die. This same methodology can be applied with chloramphenicol resistance, to act as a secondary safeguard to ensure proper transformation.<sup>5,13,14,15</sup>

### **Photochemistry and Caging Groups**

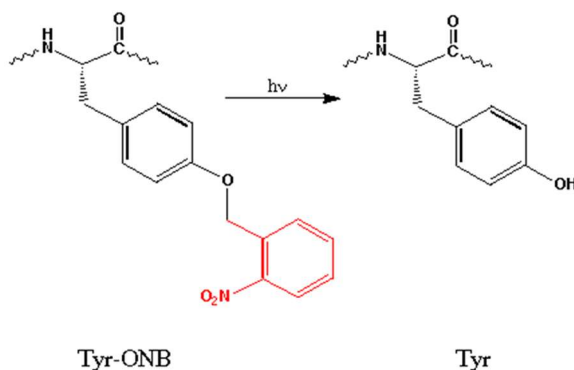
Manipulation of proteins provides a powerful tool as a therapeutic. Since amino acids are the molecules which determine shape and function of proteins, the editing of a single amino acid can have drastic effects on protein functionality. Unnatural amino acids have been incorporated into both prokaryotic and eukaryotic cells, which allows experimentation with bacteria, yeast, and even mammalian cells.<sup>10,13-15</sup> Many diseases are caused by malfunctioning proteins, especially misfolded protein; therefore, options to internally correct these harmful proteins are necessary.<sup>9</sup>

One of the more interesting and important areas in biochemistry, especially with regards to amino acids, is the manipulation of proteins via photochemistry. Photochemistry provides great utility to controlling biological macromolecules, as light can be controlled with extreme precision, specifically with respect to location and time.<sup>20</sup> Fluorescent and

UV/Vis equipment are also found ubiquitously in laboratories, making these technologies readily applicable. One advantage of using light is that certain reactions can be triggered using a specific wavelength light, without damaging the cell.<sup>20</sup> Also, this non-phototoxic irradiation does not affect the cell normally; therefore, if certain photosensitive molecules are placed inside the cell, they can be safely reacted without fear of simultaneous reactions.

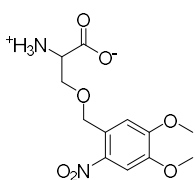
Recently, biochemists have taken interest in the utilization of caging groups. These caging groups are protecting groups that inactivate the molecule, but can easily be removed via irradiation of light.<sup>20</sup> The most common caging groups are based on ortho-nitrobenzyl moieties; however, more advanced caging groups have been developed, such as a nitrodibenzofuran chromophore and 8-bromo-7-hydroxyquioline.<sup>20-22</sup> One notable example of caging groups involves caged IPTG to control expression of genes under control of the *lac* operator.<sup>23</sup> Caged IPTG was introduced into bacterial cells, and due to the caging group, IPTG was unable to bind to the lac repressor, but when irradiated, the caging group disassociated and IPTG removed the lac repressor allowing gene expression.<sup>23</sup> This is significant, as IPTG decaging mimics the effect of adding decaged IPTG allowing gene expression to be turned on in a spatial and temporal fashion.<sup>23</sup>

While proteins can be affected by the caging of small effector molecules, they can also be regulated by caging of amino acids.<sup>20</sup> This, however, is more difficult, as attempts to cage isolated amino acids on a protein surface are non-specific due to the fact that location of the caging group cannot always be ensured.<sup>20</sup> Utilizing orthogonal synthetase/tRNA pairs has proved to be

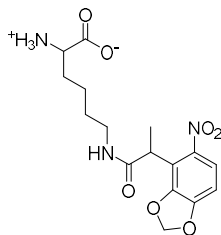


useful in caging amino acids at specific residues.<sup>20</sup> One example of this is light activated transcription via a photocaged tyrosine at residue 639 of a bacteriophage (Scheme 1.1).<sup>24</sup> The same methodology as described above was performed, where an amber stop codon was incorporated into a plasmid and transformed with the bacteriophage, such that a tyrosine with an ortho-nitrobenzyl caging group could be placed into an RNA polymerase enzyme.<sup>24</sup> The caging group blocked the active site from the nucleotide, preventing transcription; however, after a brief radiation with non-cytotoxic UV light, the caging group was removed from the tyrosine residue and RNA polymerase functionality was restored.<sup>24</sup>

Caged amino acids have also found utility in mammalian cells. For example a caged lysine has been incorporated to control nuclear import of the tumor suppressor p53 (Figure 1.8).<sup>25</sup> An ortho-nitrobenzyl cage was synthesized and attached to lysine, such that an irradiation of light would degrade the amino acid in a non-harmful manner.<sup>25</sup> Researchers



A



B

Figure 1.8: (A) Photocaged serine. (B) Photocaged lysine.

utilized an orthogonal *MbPyIRS*/tRNA pair that is congruent with the amber stop codon, in order to facilitate protein synthesis with the caged lysine.<sup>285</sup>

A single caged lysine was placed at position 305 of p53 utilizing the above method, and found that p53 was largely unable to enter the nucleus of mammalian cells, due to the caging group disrupting the nuclear localization signal binding.<sup>25</sup> Nuclear localization signals are used to import macromolecules into the nucleus, and without the tag, p53 was unable to enter. Subsequently, when the cell was photolysed with one second of 365 nm light, the lysine residue became decaged, and p53 was able to be imported from the cytosol to the nucleus.<sup>25</sup>

In another experiment, a photocaged serine (4,5-dimethoxy-2-nitrobenzylserine) was inserted into *Saccharomyces cerevisiae* utilizing the previously described amber suppression (Figure 1.8).<sup>26</sup> The caged serine unnatural amino acid was incorporated into the transcription factor Pho4 at serine residue 128.<sup>26</sup> Pho4 is needed for the expression of a phosphate-repressible acid phosphatase, which acts to remove phosphate groups, and is especially involved in controlling the cellular response to inorganic phosphate.<sup>26,27</sup> By caging the serine molecule, Pho4 was unable to be phosphorylated, which inhibited export.<sup>26</sup> After irradiation, the free serine residue was able to be phosphorylated, and Pho4 transport was restored.<sup>26</sup> Experimentation was performed by fusing Pho4 with GFP, and tracking the movement of the transcription factor in real time.<sup>26</sup>

## **Microwaves**

Within the field of synthetic chemistry, microwaves have gained prominence in the last thirty years due to the instrument's reduced reaction time, high yields, and nuanced reaction effects.<sup>28</sup> Microwaves function by exposing a substance to microwave radiation (wavelengths ranging from 1 m - 1 mm, and frequencies ranging from 300 MHz - 300 GHz).<sup>28</sup> When exposed to a microwave electromagnetic field, the dipoles of molecules will align with alternative motion of the electromagnetic waves. In a process known as dielectric heating, the alternating molecules collide and distribute energy with each other via their vibrating motions.<sup>28</sup> The rise in average kinetic energy is proportional to a rise in temperature, which will uniformly heat the substance undergoing irradiation. Due to the necessity of a dipole, substances without a dipole moment typically do not generate heat under microwave conditions.

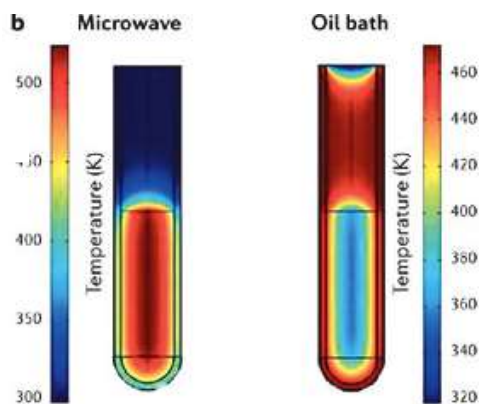


Figure 1.9: Temperature readings of two heating methods. Microwaves uniformly heat the reaction mixture, compared to an oil bath, which heats the reaction mixture non-uniformly.<sup>30</sup>

Since first used in the late 1970s in synthetic chemistry, microwave technology offers a powerful technique in the quick and efficient reaction of organic and inorganic compounds.<sup>29</sup> The main utility of microwaves is their ability to uniformly heat in a rapid fashion, compared to more standard methods of heating such as oil baths and heating jackets, which heat reaction vessels,

but in a non-uniform fashion (Figure 1.9).<sup>30</sup> Currently, microwave technology is used for a variety of synthetic reactions, such as Suzuki Coupling<sup>31</sup>, Krapcho decarboxylation<sup>32</sup>, and more broadly, in the synthesis of analgesic drugs<sup>33</sup>. One example of promising microwave technology is in 1,2,3 triazole cycloadditions (“click” chemistry).<sup>34</sup>

## Chapter 2: Photoisomerable Azobenzene UAAs

The discovery of GFP prompted a surge in the engineering of switchable GFP hybrids. While GFP is able to fluoresce, a useful feature, the ability for a fluorescent protein to switch its state of fluorescence provides a superior resource for researchers. For example, photoswitching proteins can assist in live-cell imaging, information storage, protein

control, and protein tracking.<sup>35</sup> The first few switchable fluorescent proteins that were generated were unable to provide a high contrast in their switching ability, limiting their utility.<sup>35</sup>

These proteins included Yellow Fluorescent Protein (YFP), Cyan Fluorescent Protein (CFP), and Citrine.<sup>35</sup> The first widely used switchable

protein was Kindling Fluorescent Protein (KFP), which was able to track protein movements *in vivo*.<sup>35</sup> Shortly following the development of KFP, Dronpa, a photoswitchable fluorescent protein derived from stony coral, was engineered.<sup>36</sup> Dronpa has the ability for its baseline “on” stage to be turned “off” via irradiation with light (Figure 2.1).<sup>35,36</sup> Several mutants of the protein were created that alter Dronpa’s photoswitching kinetics, such as Dronpa-2 and Dronpa-3.<sup>36</sup> Another Dronpa mutant, Padron, has a reverse characteristic from Dronpa, where its baseline state is “off”, and irradiation of light switches the photoswitchable fluorescent protein “on”.<sup>36</sup>

Green Fluorescent Protein has a particular excitation spectrum, where there are normally two excitation peaks at 395 nm and 475 nm. When UV light is irradiated, or the pH is changed, the large peak at 395 nm diminishes while the smaller peak at 475 nm

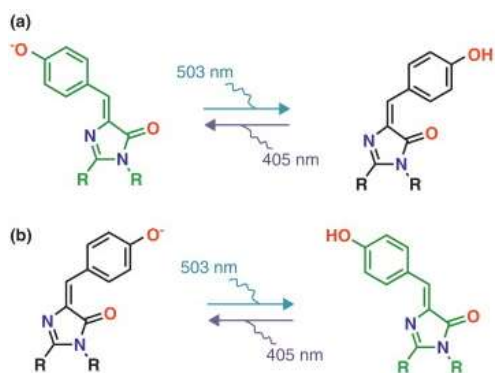


Figure 2.1: (a) Photoswitching involves cis–trans isomerization in Dronpa (b) trans–cis isomerization in Padron.<sup>35</sup>

risers.<sup>16</sup> This change has to do with altering protonation states of the tyrosine amino acid at residue 66.<sup>16,36</sup> GFP has been coupled with UAA research, for example, using orthogonal aminoacyl-tRNA synthetases, unnatural amino acids were able to replace the fluorophore tyrosine.<sup>16</sup> Unnatural tyrosines with the phenol replaced with boronate, azido, nitro, and keto groups were put into residue 66 and analyzed for alterations in fluorescent properties.<sup>16</sup> The scans showed reduced or no fluorescence intensities, while also showing shifted emission maxima, for the UAA-protein complexes with reduced intensities.<sup>16,36</sup>

### Photo-Reversible Amino Acids and Domains

One major problem with caging groups is that they are not reversible, meaning after irradiation and decaging, the effector molecule or amino acid cannot be recaged. Therefore, when researchers need photoreversibility, one

method they employ is photoisomerable molecules, such as molecules with azobenzene moieties.<sup>37</sup> Azobenzene naturally exists in the more stable trans formation; however, when irradiated with light, azobenzene molecules shift to a cis formation (Scheme 2.1).<sup>37</sup> This has been utilized in controlling ion channels, where an azobenzene molecule can become

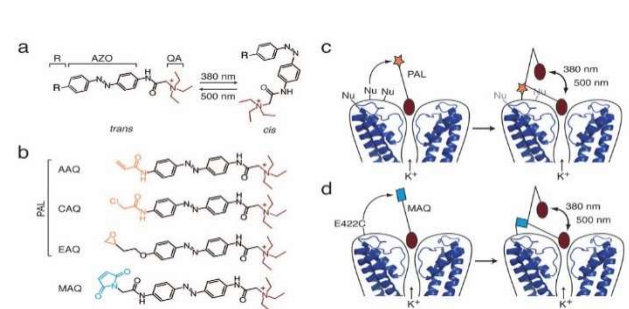
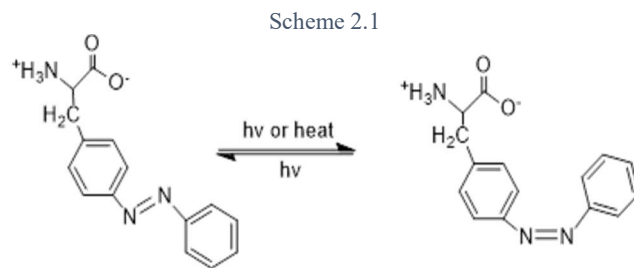


Figure 2.2: Examples of azobenzene derivatives, and their utility as ion channel regulators.<sup>37</sup>

attached to the protein via reacting with a nucleophile via ligand binding, or subsequently become incorporated into the protein's primary structure.<sup>37</sup> In one experiment, the azobenzene molecule was reacted with a charged

ammonium substrate on the terminal benzene group, such that when the molecule was irradiated, it would shift to the cis position, blocking ions from entering the ion channel (Figure 2.2).<sup>37</sup> Gratifyingly, the azobenzene molecule can shift back either thermally or with irradiation at a separate wavelength, reverting the ion channel to normal functionality.<sup>37</sup>

One other reversible system employed is the fusing of Light-Oxygen-sensing domains (LOVs) to proteins.<sup>38</sup> LOV domains are bound to a flavin mononucleotide (FMN)

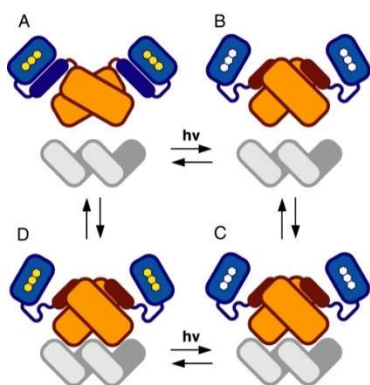


Figure 2.3: An example of a conformational change of a flavin mononucleotide, unwinding the LOV domain, and signaling an effector protein.<sup>38</sup>

cofactor, which undergoes a conformational change when irradiated with light at 450-470 nm.<sup>38</sup> The change in protein structure is a result of the unwinding of a helical domain near the LOV domain, which acts as a signaling message to attached effector protein (Figure 2.3).<sup>38</sup> In one experiment,

a LOV domain replaced the chemosensor domain of a histidine kinase.<sup>39</sup> The kinase is traditionally light inert;

however, after replacement with the LOV domain, kinase

activity was repressed when illuminated with light.<sup>39</sup> Kinase activity was restored after thermal recovery, but only recovered after an intermediary lag phase.<sup>39</sup> With the

importance of reversible moieties apparent, we propose an azobenzene unnatural amino acid to be synthesized and inserted into GFP (Figure 2.4). The fluorescence of GFP will be measured while the azobenzene UAA is in the trans and cis positions.

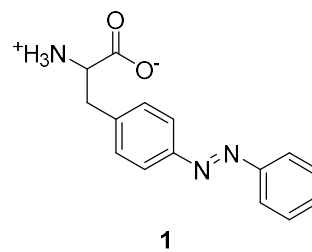


Figure 2.4: A proposed azobenzene unnatural amino acid.

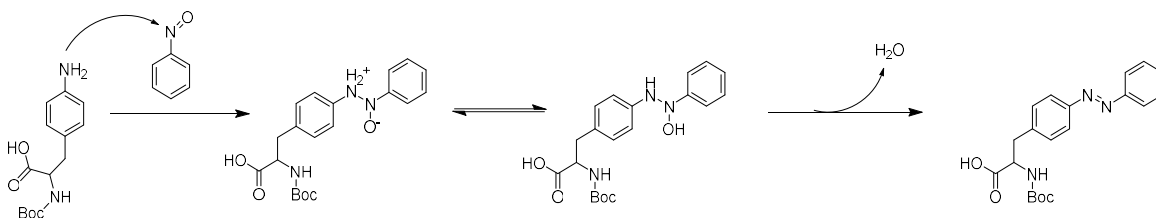


## Results and Discussion

### Azobenzene UAA:

The synthesis of **1** was performed under relatively mild conditions, utilizing a Mills reaction.<sup>40,41</sup> The Mills reaction proceeds via attack of the nitroso by the aniline, followed by removal of water (Scheme 2.2). Hydrochloric acid was then used to remove the Boc group. Due to the native charged state of amino acids, normal purification methods were not possible. Therefore, the reaction was simply concentrated with a rotary evaporator, for a modest 61% yield. When making a stock solution for protein expression, **1** was not able to be dissolved in water. However, the insolubility was addressed via sonication and addition of NaOH, which was needed to neutralize **1** from the deprotection step. To enable incorporation of **1** into GFP, competent cells were co-transformed with an Azobenzene pEVOL-tRNA synthetase plasmid and four separate pET-GFP-TAG plasmids containing TAG mutations at residues 3, 66, 133, and 151. Mutant GFPs were obtained utilizing the general GFP expression as will be described. Cell cultures were induced with **1**, to afford GFP with an incorporated azobenzene unnatural amino acid.

Scheme 2.2



Various buffers were analyzed for the fluorescence measurements, including 5X Tris Buffer, 1X Tris Buffer, and PBS. It was originally assumed that due to the low salt

concentrations, 1X buffer and PBS would be the most consistent, but it was demonstrated empirically that 5X Tris buffer yielded the most consistent spectra. Each of the four GFP mutants with **1** were examined before, immediately after, and at several hours after irradiation at 365 nm. Different irradiation time lengths were employed, and 10 minutes of irradiation was found to be optimal for full photoisomerization of the azobenzene UAA. Full thermal photoisomerization to the trans position was observed after approximately 16 hrs. From the fluorescence experiments, addition of **1** to each of the GFP four residues was demonstrated to shift the emission spectrum of the protein (Figure 2.5). Irradiation resulted in drastic emission shifts to higher wavelengths, while thermal recovery restored emission to pre-irradiation levels (Figure 2.6). This demonstrates the photoisomerability of the azobenzene UAA, as it is able to alter protein function when irradiated, and also restore native function.

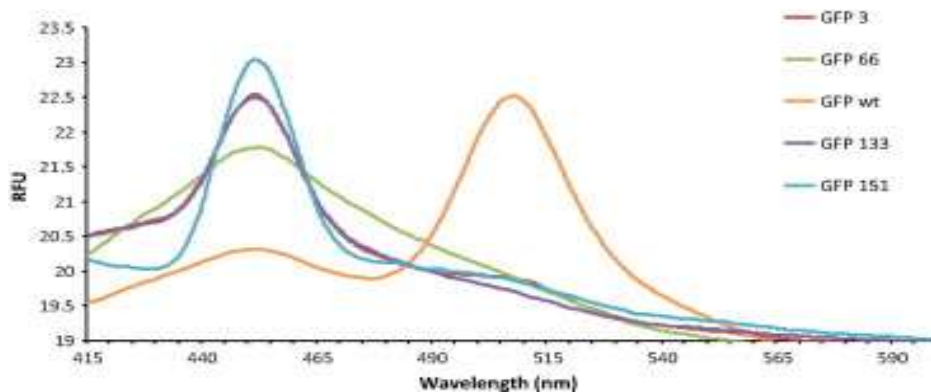


Figure 2.5: Emission scans for GFP mutants containing **1** at various residues, compared to GFP wt (orange).<sup>42</sup> Interestingly, for each GFP mutant, the emission peak rose significantly for the 475 nm peak, while falling greatly at the 510 nm peak.

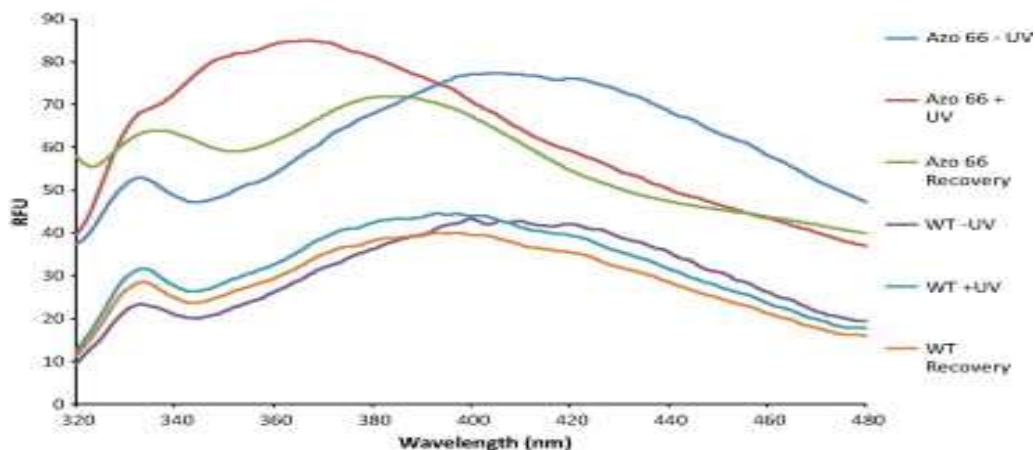


Figure 2.6: Emission scans for GFP mutants containing **1** at residue 66. UV irradiation prompted great shifts to lower wavelengths of light, compared to GFP wt.<sup>42</sup> Also, recovery of the GFP mutant is shown via the green line, demonstrating the photoisomerizability of the UAA.

Given the photoswitching observed by fluorescence spectroscopy, spatial control over GFP fluorescence was attempted. LB Agar plates containing the inductants and the unnatural amino acid **1** were prepared and *E. coli* harbouring the pET-GFP-Y66TAG plasmid and the pEVOL-Azobenzene plasmid were plated and allowed to grow overnight at 37° C. After a bacterial lawn was

formed, an aluminum foil mask was placed over half of the plate to block UV irradiation. When irradiated for 15 minutes or longer, photobleaching occurred, preventing the detection of

photoisomerization. The optimal

irradiation conditions were ultimately found to be 10 min at 365 nm. Excitingly, a differential fluorescence was observed only on cells that were exposed to UV irradiation (Figure 2.7). After a 4 h recovery in the dark at room temperature the plate was imaged

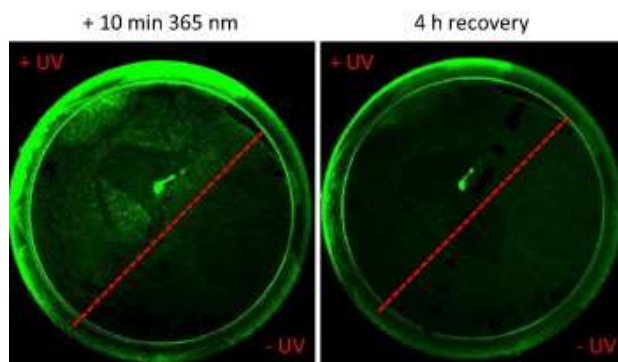


Figure 2.7: A spatial control experiment demonstrating both the ability of **1** to change the fluorescence of GFP after UV irradiation, and reversibly photoisomerize, restoring pre-irradiation fluorescence.<sup>42</sup>

again, and the differential fluorescence was absent, signifying a reversion to the trans form of the fluorophore. The result demonstrates the ability to ‘switch’ the GFP fluorescence in a spatial and temporal fashion. Moreover, only slight photobleaching was observed when the wild type GFP was expressed and irradiated on the plates, and no recovery was observed

### Azobenzene Alkyne UAA:

The ability to control the wavelength of light needed to photoisomerize an azobenzene unnatural amino would be useful if applied to living systems. Azobenzene isomerizes at 365 nm, a cytotoxic wavelength, which would not be applicable *in vivo*.

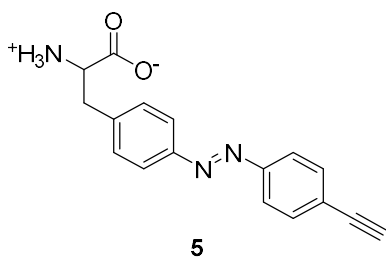
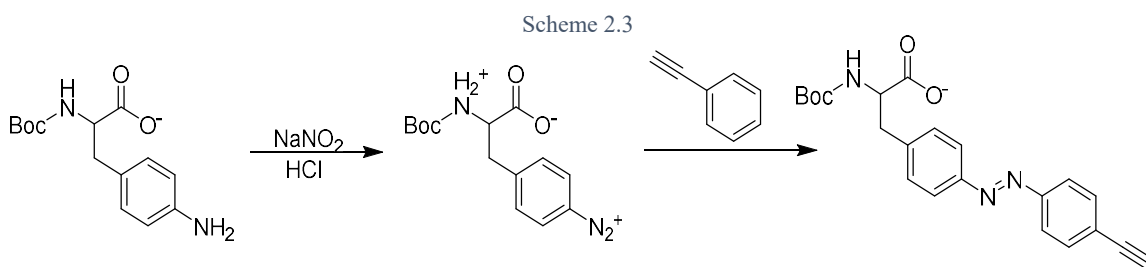


Figure 2.8: An azobenzene alkyne UAA

Derivatizing the *para* position of the terminal benzene is hypothesized to shift the isomerization wavelengths, such that photoisomerization of an azobenzene UAA would be accomplished at non-toxic wavelengths.<sup>43</sup> Adding an alkyne derivative would enhance the conjugation of the azobenzene moiety, which could shift the isomerization wavelengths, but also act as a biorthogonal handle (Figure 2.8). Terminal alkyne groups can undergo bioconjugation via click and Glaser-Hay reactions, as will be described in the following sections. This allows for dyes, proteins, and other macromolecules to be attached to a protein with **5**.

Synthesis of **5** was first attempted via a standard azo coupling, in which a para-amino group is converted into a diazonium group via sodium nitrite and a strong acid (Scheme 2.3).<sup>43</sup> The diazonium group acts as an electrophile and will couple to the para position of activated aromatics at a neutral pH, hence the addition of sodium hydroxide. Hydrochloric acid was first investigated as the acid source; however, it appeared that the

tert-butyloxycarbonyl (Boc) group of the 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine was also being deprotected. This resulted in diazonium formation of the N-terminus and the para-amino group of the 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine molecule. Glacial acetic acid was then tested as the acid source, but it was not strong enough to generate the diazonium group. NMR of the reaction indicated that there was only starting material.



Consequently, an alternative synthetic method was then employed, utilizing a similar mechanism as employed for the azobenzene UAA **1** (Scheme 2.4). This involved oxidizing the amino group of TMS ethynyl aniline molecule into a nitroso group. The nitroso group was then reacted with the 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine in a condensation reaction to form the azo group. The product was difficult to separate via flash chromatography, due to the similarities of the reactants and products. Despite a high 30:1 DCM:MeOH eluent system, poor separation still occurred, which would inflate the yields because of the remaining starting material. Based off NMR data, the yield was determined to be a poor 20 %.





and GFP was induced in a 96-well plate with appropriate expression controls, as described in the general aaRS assessment protocol. Fluorescence of the cell cultures were measured on a microplate reader after 16 hours of expression. The pAzF synthetase appeared to incorporate **5** at a relatively higher rate than the control (Figure 2.9). GFP containing **5** was then expressed and purified utilizing the pAzF pEVOL plasmid; however, no GFP was produced on this large scale. This could be due to the pAzF plasmid incorporating **5** poorly, but enough to indicate a higher result than the control.

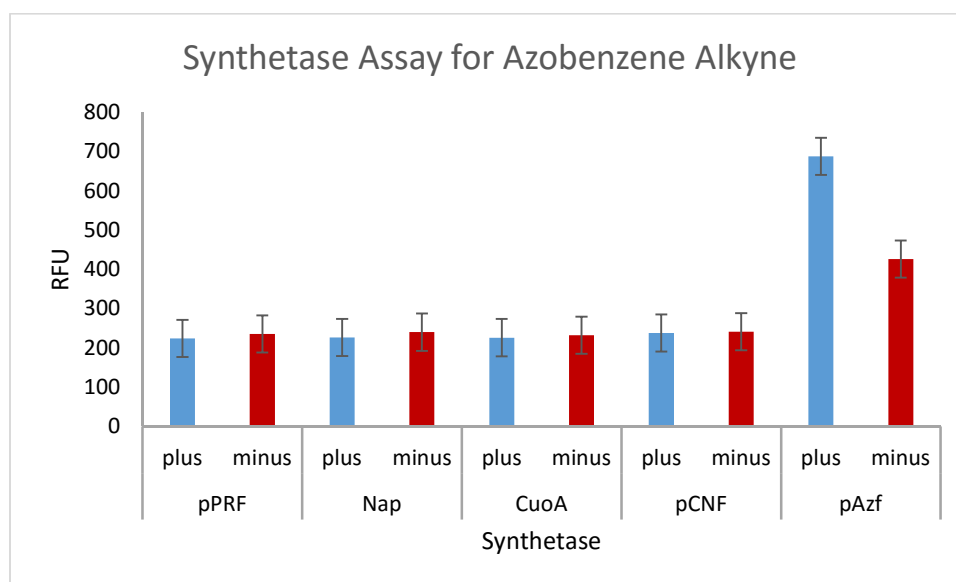


Figure 2.9: Synthetase assay for **5**. Five synthetases were tested for their ability to incorporate **5** into GFP at residue 151. The synthetase pAzf showed the most promise, as fluorescence measurements deviated the most from the control.

Despite current efforts, it appears that **5** has not been incorporated into GFP with high yields. One possibility is that the most recent synthesis is not **5**; however, an LCMS was run confirming that **5** has been properly synthesized. Another synthetase assay could be run, testing a variety of synthetases with large binding pockets. Alternatively, a selection could be performed that would evolve a synthetase to incorporate **5**.



## Experimental

**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. NMR data was acquired on a Varian Gemini 400 MHz. Fluorescence data were measured using a PerkinElmer LS 55 Luminescence Spectrometer and a Synergy Microplate reader. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit.

**General GFP Expression:** A pET-GFP-TAG plasmid (0.5  $\mu$ L) was co-transformed with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electorporator. The cells were then plated (100  $\mu$ L) 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 Chlr) at an OD<sub>600</sub> 0.1. The cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8, and protein expression was induced by addition of the respective unnatural amino acid (100  $\mu$ L, 100 mM) and 20 % arabinose (10  $\mu$ L) and 0.8 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG; 10  $\mu$ L). The cultures were incubated at 30° C for 16-20 h then pelleted at 5,000 rpm for 10 minutes and stored at -80° C for 3 hours. The cell pellet was re-suspended using 500  $\mu$ L 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 added to an equilibrated column of Ni-NTA resin (200  $\mu$ L) and the GFP was purified according to manufacturer's

protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h), and employed without further purification. Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+) after being stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

**General aaRS Assessment:** The aaRSs were assessed for unnatural amino acid incorporation by first co-transforming a pET-GFP-TAG variant plasmid (0.5  $\mu$ L) with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L), in a similar method as described previously. For each pEVOL plasmid, one colony was used to inoculate LB media (4 mL) in separate containers that contained ampicillin (4  $\mu$ L, 50 mg/mL) and chloramphenicol (4  $\mu$ L, 34 mg/mL). The cultures were incubated at 37° C overnight. Expression cultures were created containing LB media (4 mL), ampicillin (4  $\mu$ L, 50 mg/mL), and chloramphenicol (4  $\mu$ L, 34 mg/mL), and then were inoculated with the previous culture (250  $\mu$ L). The expression cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8. Protein expression was induced by addition of 20 % arabinose (4  $\mu$ L), and 0.8 mM IPTG (4  $\mu$ L). The cultures were then plated on a 96-well plate (200  $\mu$ L), with six wells for each plasmid. Three wells were designated as controls, and the other three were designated for the experiment. The designated UAA (100  $\mu$ L, 100 mM) was added to the experimental wells only. The plate was allowed to shake at 30° C for 16 h, after which it was measured for fluorescence (528 nm and 485 nm) utilizing a Synergy HT microplate reader.

## Azobenzene UAA

**(E)-2-ammonio-3-(4-(phenyldiazenyl)phenyl)propanoate, 1:** The unnatural amino acid was prepared according to literature conditions.<sup>40</sup> To flame dried round bottom was added 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine (1.007 g, 3.59 mmol, 1 equiv.) and glacial acetic acid (200 mL). Nitrosobenzene (579 mg, 540  $\mu$ mol, 1.5 equiv.) was then added and stirred for 16 hours. The reaction was then quenched with dI water (100 mL). The product was extracted with ethyl acetate (3 x 20 mL), and the organic layer was then back-extracted with brine (2x30 mL). The organic layer was isolated, dried with MgSO<sub>4</sub>, filtered, and solvent was removed by rotatory evaporation. The orange solid was purified via flash chromatography with a 9:1 DCM/Methanol eluent. The product was concentrated and dissolved in Dioxane (4 mL), and cooled in an ice bath. To the solution, 6N HCl (2 mL) was added and stirred for 3 hours. The solvent was then removed *in vacuo* to afford the deprotected azobenzene amino acid as an orange solid in a 61% yield (590 mg, 2.19 mmol).

**Fluorescence Measurements:** Each fluorescence spectrum was obtained in a glass cuvette with Tris Buffer (3 mL). To the cuvette, the desired GFP mutant protein was added (10  $\mu$ L; 0.3 mg/mL). Irradiation results were achieved via irradiating the cuvette with the protein at 365nm for 10 minutes, followed by immediate spectral measurements. The sample was excited at 300 nm and the emission spectral was recorded from 320 nm to 500 nm, 320 nm and the emission spectral was recorded from 350 nm to 550 nm, and 395 nm and the emission spectral was recorded from 415 nm to 600 nm. The excitation slit width was set to 10 nm, and emission slit width was set to 10 nm. Scan speed was set to 400 nm/min.

**Spatial Control of Azobenzene GFP Fluorescence:** Agar plates (10 mL LB agar/plate; 60 mm X 15 mm) containing Ampicillin (10  $\mu$ l), Chloramphenicol (10  $\mu$ l), 10% Arabinose (10  $\mu$ l), and 1 M Isopropyl  $\beta$ -D-1-thiogalactopyranoside (10  $\mu$ l) were prepared. Some plates were supplemented with a 100 mM L-phenylalanine-4'-azobenzene mixture to a final concentration of 1 mM, and other plates were prepared in the absence of the unnatural amino acid. *E.coli* BL21(DE3) cells (50  $\mu$ L) containing the pEVOL-Azobenene plasmid and pET-GFP-Y66TAG were added to each plate and incubated overnight at 37 °C to afford a bacterial lawn. A plate containing the UAA and all inductants was half-covered with aluminum foil and irradiated at 365 nm for 10 minutes following imaging on a BioRad gel imaging system with the appropriate UV filters. The plate was then allowed to recover for 4 hours at 37 °C followed by re-imaging.

## **Azobenzene Alkyne UAA**

### **First Method**

**(E)-2-((tert-butoxycarbonyl)amino)-3-(4-((4-ethynylphenyl)diazenyl)phenyl)-propanoic acid, 2:** 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine (100 mg, 0.357 mmol, 1 equiv.) was added to a flame dried round bottom, and dissolved with glacial acetic acid (1 ml). The solution was warmed until it dissolved, and then cooled to 0° C. In an Eppendorf tube, sodium nitrite (44 mg, 0.64 mol, 2 equiv.) was dissolved in deionized water (330  $\mu$ l). The contents of the Eppendorf were added to round bottom and allowed to stir for 45 mins at 0° C. To the solution was added ethynylbenzene (50  $\mu$ l, 0.455 mmol, 1.5 equiv.) and sodium hydroxide (3400  $\mu$ l, 1 M) to neutralize the solution. The solution was

allowed to stir for 2 hrs at 0° C. The reaction was quenched with sodium bicarbonate (100 µl). The product was extracted with ethyl acetate (3 x 20 mL), and the organic layer was then back-extracted with brine (3 x 30 mL). The organic layer was isolated, dried with MgSO<sub>4</sub>, filtered, and solvent was removed by rotatory evaporation. The subsequent solution was purified via flash chromatography with a 10:1 DCM/Methanol eluent.

## Second Method

**Trimethyl((4-nitrosophenyl)ethynyl)silane, 3:** To a flame dried vial was added Oxone (320 mg, 1.04 mmol, 2 equivalents) with deionized water (3 ml), and stirred for 10 mins. The contents of the vial were then added to a flame dried round bottom with 4-[(Trimethylsilyl)ethynyl]aniline (100 mg, 0.53 mmol, 1 equivalent) and dichloromethane (3 mL). The reaction was stirred overnight at room temperature. The product was then extracted with dichloromethane (3x20 mL), and the organic layer was then back-extracted with brine (3 x 30 mL). The organic layer was isolated, dried with MgSO<sub>4</sub>, filtered, and the solvent was removed by rotatory evaporation. The product was advanced to the next step without further purification.

**(E)-2-((tert-butoxycarbonyl)amino)-3-(4-(((trimethylsilyl)ethynyl)phenyl)diazonyl)phenyl)propanoic acid, 4:** In a flame dried vial, 4-Amino-N-(tert-butoxycarbonyl)-L-phenylalanine (150 mg, 0.54 mmol, 2 eq) was dissolved in glacial acetic acid (3 ml). The contents of the vial were transferred to a round bottom containing **3** (50 mg, 0.22 mmol, 1 eq) dissolved in acetic acid (3 ml), and left to stir overnight. The product was extracted with dichloromethane (3 x 20 mL), and the organic layer was then back-extracted with brine (3 x 30 mL). The organic layer was isolated, dried with MgSO<sub>4</sub>, filtered, and solvent was removed by rotatory evaporation. The subsequent solution was

purified via flash chromatography with a 20:1 DCM/Methanol eluent, resulting in a brown solid (50 mg, 0.11 mmol, 20 % yield). <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 8.258 (d, J=1.2 Hz, 2H), 8.151 (d, J=8.6 Hz, 2H), 7.859 (d, J=7.4 Hz, 2H), 7.630-7.552 (m, 2H), 4.609-4.607 (m, 1H), 3.268-3.208 (m, 2H), 1.255 (s, 6H), 0.007 (s, 6H).

**(E)-2-((tert-butoxycarbonyl)amino)-3-(4-((4-ethynylphenyl)diazenyl)phenyl)-**

**propanoic acid, 2:** In a flamed dried round bottom, **4** (50 mg, 0.11 mmol, 1 eq) was dissolved in dichloromethane (3 ml), and tetra-n-butylammonium fluoride (68 mg, 0.22 mmol, 2 eq). The reaction was stirred for 4 hrs at room temperature. The product was extracted with dichloromethane (3 x 20 mL), and the organic layer was then back-extracted with brine (3 x 30 mL). The organic layer was isolated, dried with MgSO<sub>4</sub>, filtered, and solvent was removed by rotatory evaporation. The subsequent solution was purified via flash chromatography with a 30:1 DCM/Methanol eluent, resulting in a brown solid (30 mg, 0.08 mmol, 70 % yield). <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 8.289 (d, J=9 Hz, 2H), 8.166 (d, J=8.6 Hz, 2H), 7.909 (m, 2H), 7.658-7.560 (m, 2H), 4.321 (t, 1H), 3.615-3.610 (m, 2H), 3.251 (t, 1H), 1.262 (s, 6H).

**(E)-1-carboxy-2-(4-((4-ethynylphenyl)diazenyl)phenyl)ethan-1-aminium, 5:** In a flamed dried round bottom, **2** (30 mg, 0.08 mmol, 1 eq) was dissolved in 1,4-Dioxane (4 mL), and cooled in an ice bath. To the solution, 6N HCl (2 mL) was added and stirred for 3 hours. The solvent was removed *in vacuo* to afford the deprotected azobenzene alkyne amino acid as a brown solid (23 mg, 0.08 mmol, 99% yield).

## Chapter 3: A Photosensitive Propargyl Tyrosine UAA

Proteins are diverse and complex macromolecules, and are often difficult to tag or modify, especially under physiological conditions. With the discovery of chemoselective reactions, macromolecules can be covalently linked, even in complex cellular environments.<sup>45</sup> This process is often referred to as a bioconjugation reaction, and provides a powerful tool to modify proteins through coupling with biophysical probes<sup>46,47</sup>, therapeutic agents<sup>48</sup>, or surfaces<sup>49</sup>. For example, the bioconjugation of naturally occurring thiol groups on cysteine residues with iodoacetamides has been utilized to determine the number of free cysteine residues in a protein.<sup>50</sup> Unnatural alkyne handles enable bioconjugation techniques such as click and Glaser-Hay reactions.<sup>44,51</sup> These reactions represent highly chemoselective covalent modifications that will be discussed in greater detail in the following sections.

As previously mentioned, caging groups offer the ability to exogenously control protein function in a spatial and temporal fashion. The caging groups inactivate the protein, and when cleaved with light, protein functionality is restored.<sup>23-25</sup> While the caging group can act as protein modifier, functional groups can be added to enhance its utility.

A proposed unnatural amino acid containing an ortho-nitrobenzene backbone with an alkyne handle would allow bioconjugation and photocleaving of the caging group (Figure

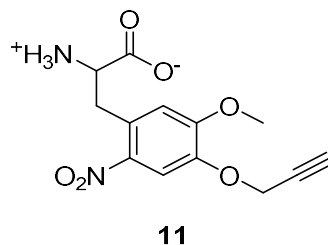
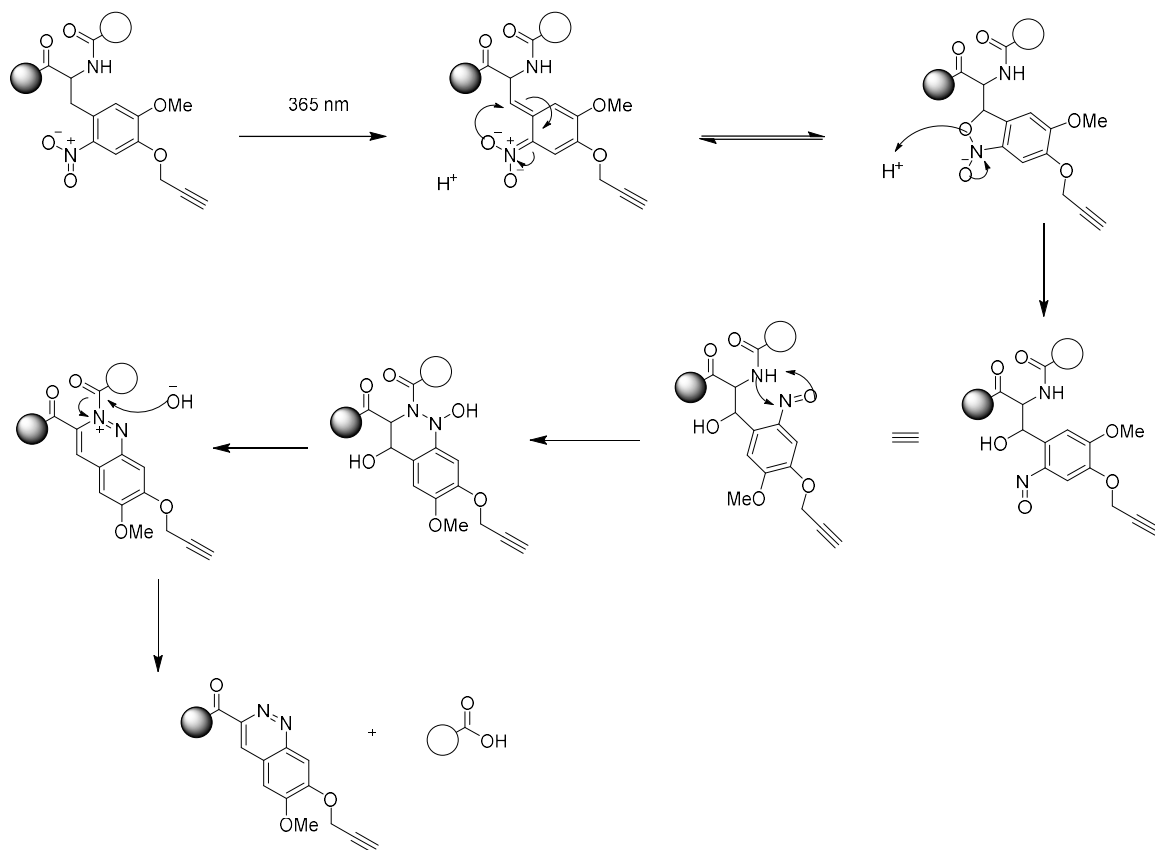


Figure 3.1: Proposed nitro-propargyl tyrosine unnatural amino acid. The UAA possesses both a caging group and an alkyne handle.

3.1). This UAA, once incorporated into a protein, could be conjugated to a drug, introduced into a cell, and then release the drug via photocleavage of the caging group (Scheme 3.1).<sup>52</sup>

This would allow both spatio-temporal control of drug release, and the removal of the drug from the sterics of the protein.

Scheme 3.1<sup>52</sup>



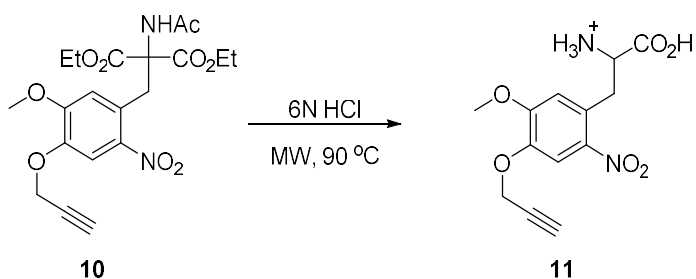
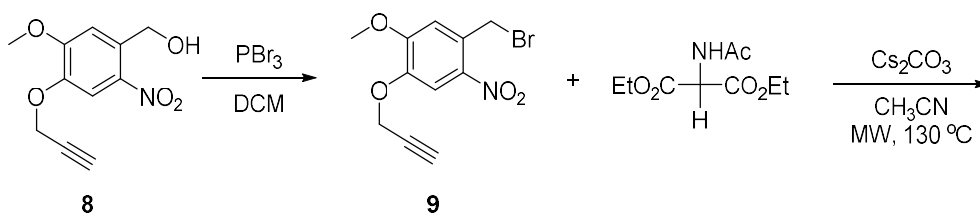
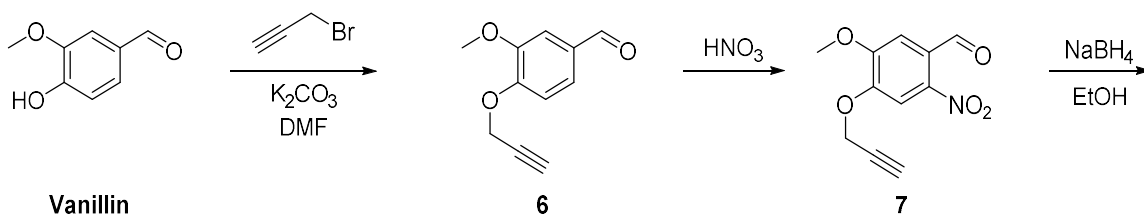
## Results and Discussion:

Synthesis of **11** proceeded via the proposed route (Scheme 3.2), with minor alterations.

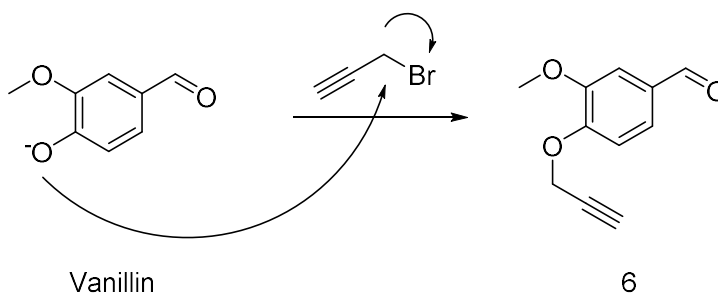
The propargylation of vanillin (**6**) was performed from a literature precedent and was achieved via nucleophilic attack of the phenoxide to the electrophilic propargyl bromide (Scheme 3.3).<sup>53</sup> It was found that using anhydrous DMF and increasing the equivalents of propargyl bromide from 3 equivalents to 5 equivalents, increased the yield from 70 % to 81 %.



Scheme 3.2



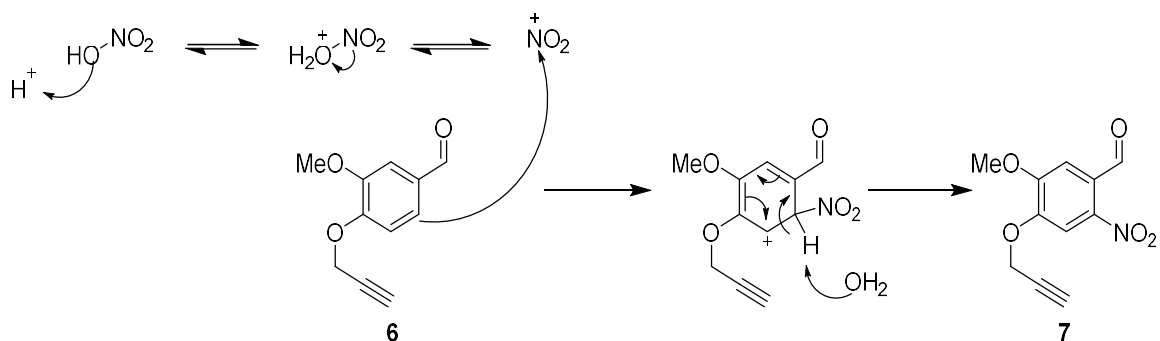
Scheme 3.3



Installation of a nitro group was accomplished by reacting **6** with nitric acid, via an electrophilic aromatic substitution (Scheme 3.4). The product was collected via Büchner filtration, with the solid being **7**. The  $\text{NO}_2^+$  cation is generated by loss of water, followed by attack by the nucleophilic aromatic group. Proton abstraction by water restores aromaticity. A decreased yield was observed on a large scale, but when the flow through from the

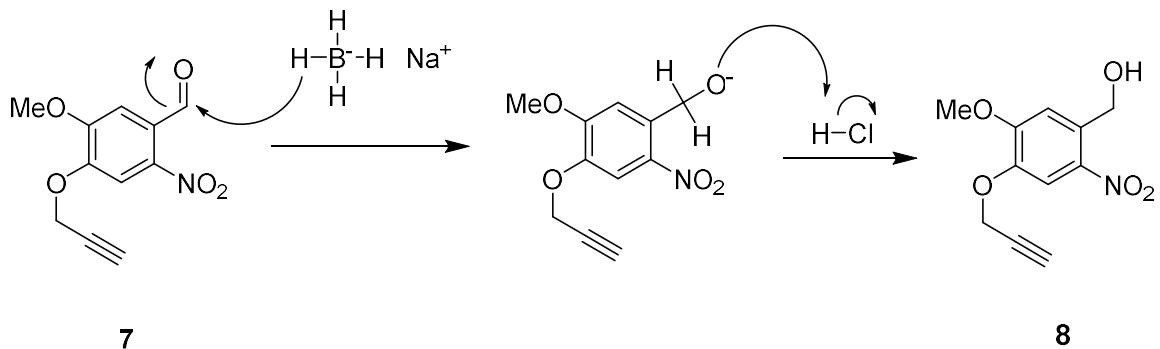
Büchner filtration was resubmitted to reaction conditions after cooling, the resulting yield increased to 81 %.

Scheme 3.4

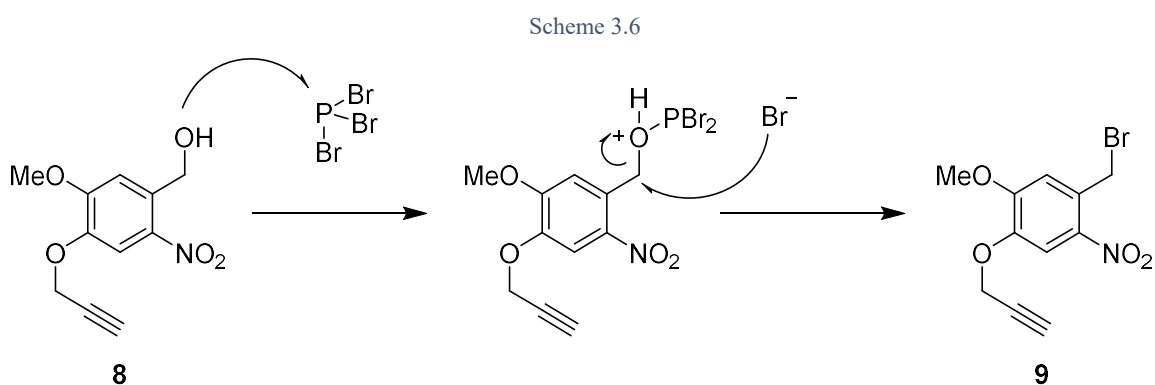


The reduction of **7** to **8** proceeded via attack of a hydride ion from sodium borohydride at the electrophilic aldehyde, generating a tetrahedral alkoxide intermediate (Scheme 3.5). In previous attempts, it was realized that the hydride ions were quenched by reaction with ethanolic protons, resulting in low yields. This was remedied by conducting the reaction in 1 M NaOH. The reaction was extracted and concentrated with a rotary evaporator, requiring no flash chromatography, resulting in a modest 76 % yield.

Scheme 3.5

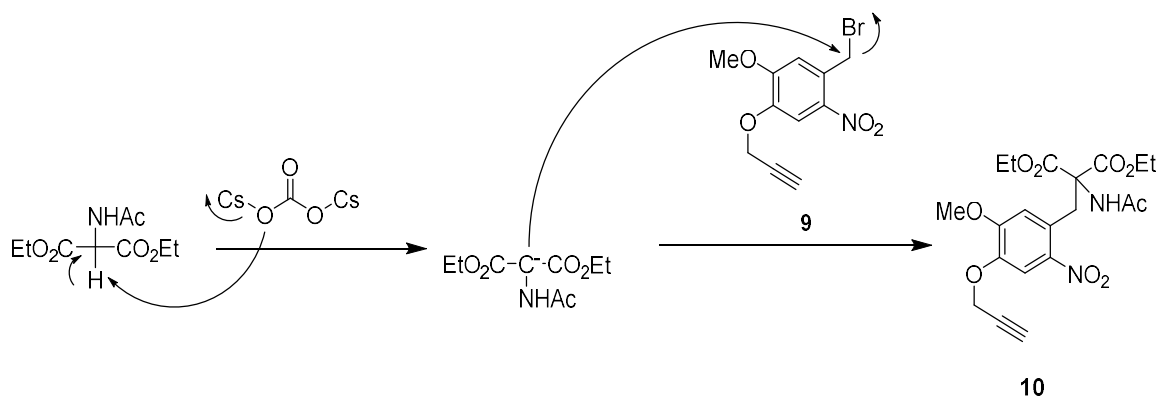


Bromination of **8** to **9** was then attempted using activated nucleophilic substitution. The nucleophilic alcohol of **8** reacts with  $\text{PBr}_3$ , displacing a bromine ion (Scheme 3.6). The resulting phosphorous ester is a good leaving group, and is displaced after nucleophilic attack by the bromine ion, forming the C-Br bond. The phosphonic dibromide is a better leaving group than bromine, thus equilibrium favors the brominated product. A 50 % yield was obtained after extraction and flash chromatography.



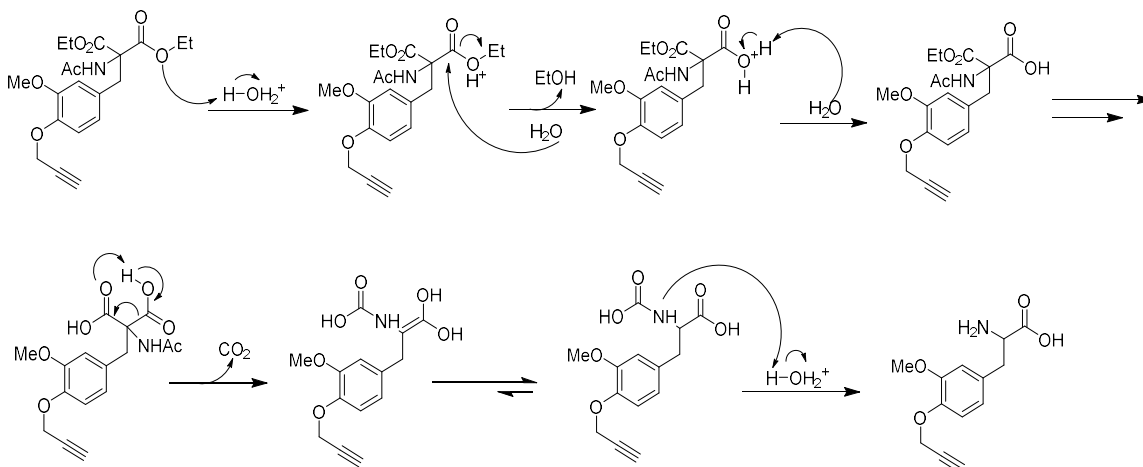
The formation of the amino acid precursor **10** was found to be optimal under microwave conditions. Previous attempts reacted NaH and diethyl acetamidomalonate in DMF; however, no product was observed. It was found that the reaction could proceed to completion under microwave irradiation, based off literature precedence (Scheme 3.7).<sup>54</sup> The microwave irradiation resulted in several changes to previously described procedures. Due to the high reactivity,  $\text{Cs}_2\text{CO}_3$  was employed in place of NaH. Moreover, the inability of DMF to tolerate large amounts of microwave energy prompted the use of acetonitrile. Due to the nature of the microwave vials, only 100 mg of **9** could be reacted at a time. After 10 minutes of irradiation at this temperature, a 79 % yield was achieved after extraction and flash chromatography.

Scheme 3.7



The conversion of **10** to **11** also proceeded via microwave irradiation using 6 N HCl (Scheme 3.8).<sup>54</sup> Poor solubility of compound **10** was observed; however, as the reaction proceeded, the more water soluble **11** was generated. This resulted in aggregation of **10**, and therefore the reaction had to be performed multiple times to achieve high yields. Due to the charged amino acid backbone, the reaction could not be extracted, nor purified by traditional silica gel column chromatography. Additionally, the deprotection step is not stereoselective, resulting in a racemic mixture of (D) and (L) diastereomers. Since only (L) amino acids can be translated by ribosomes, only half of **11** can be incorporated into proteins. However, based on NMR and LC/MS analysis, **11** was found to be present in

Scheme 3.8



relatively high purity and a yield of 95 %. With the UAA in hand, we were next ready to attempt incorporation into a protein.

Since no aaRS has previously been evolved to specifically incorporate **11**, an assay was performed to identify if any already existing aaRS was promiscuous for **11**. The pEVOL synthetases Nap, BipY, pCNF, and ONBY plasmids were chosen due to their large binding pockets, or similarities to their corresponding UAA, and were each transformed with a pET-GFP-TAG 151 plasmid. Cultures were grown, and GFP was induced in a 96-well plate with appropriate expression controls, as described in the general aaRS assessment protocol. Fluorescence of the cell cultures was measured on a microplate reader after 16 hours of expression (Figure 3.2). While most aaRSs were not able to accommodate the UAA, the results indicate that the pCNF-aaRS was able to incorporate **11** when compared to a negative control in the absence of the UAA.

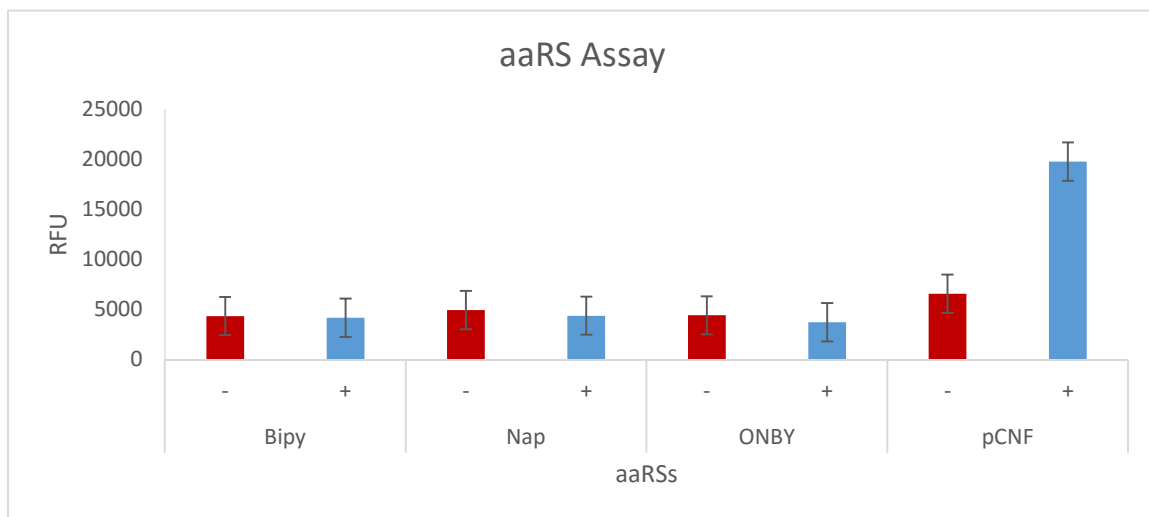


Figure 3.2: Synthetase assay for **11**. Four synthetases were tested for their ability to incorporate **11** into GFP at residue 151. The synthetase pCNF showed the most promise, as fluorescence (and therefore UAA incorporation) was highest relative to the control.

Incorporation of **11** into GFP at residue 151 via *p*CNF-aaRS was attempted on a 10 mL scale, and gratifyingly, protein expression was observed via SDS-PAGE (Figure 3.3). Next, the bioconjugation of the mutant GFP protein was examined. The GFP containing **11** was reacted with an AlexaFluor-488 alkyne fluorescent dye via a Glaser-Hay reaction, similar to the one utilized for **5** (Scheme 2.5). Initial attempts at

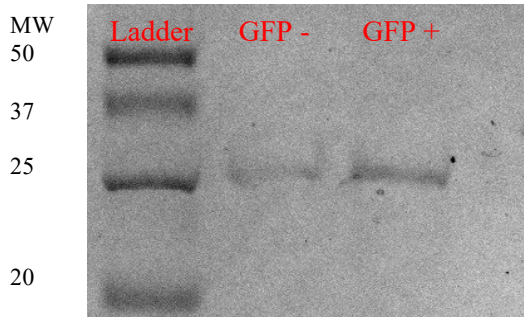


Figure 3.3: Coomassie-stained SDS-PAGE analysis of GFP containing **11** at residue 151. Some protein is seen in the GFP minus lane due to the promiscuous synthetase incorporating tyrosine. Molecular weight is in kDa.

bioconjugation failed to couple the two molecules, requiring further optimization of reaction conditions. The first experimental variable assessed was temperature, and the reaction was incubated at 22° C, a 37° C incubator, and at 4° C. Reaction time was also examined,

Lane	1	2	3	4	5	6
GFP		+1	+1	+1	WT	WT
Glaser-Hay		-	+	+	-	-
UV		-	-	+	-	+

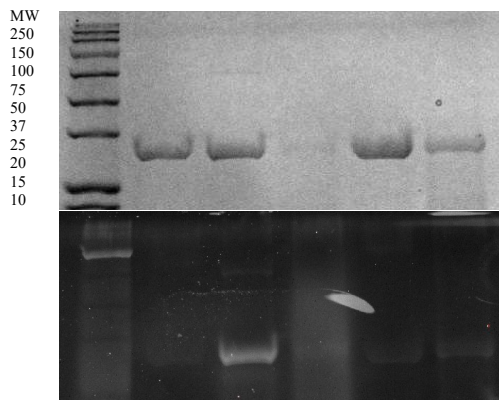


Figure 3.4: SDS-PAGE analysis of both the Glaser-Hay bioconjugation and the photoregulation of the conjugate. The top picture is a Coomassie-stained gel, while the bottom picture is a Spyro Ruby fluorescence reading of the gel. Lane 1: Protein Ladder; Lane 2: GFP containing **1** not subjected to Glaser-Hay conditions; Lane 3: GFP containing **1** coupled with AlexaFluor-488 Alkyne, but not irradiated with UV; Lane 4: GFP containing **1** coupled with AlexaFluor-488 Alkyne, and irradiated at 254 nm; Lane 5: Wild-type GFP not irradiated; Lane 6: Wild-type GFP irradiated at 254 nm. Molecular weight is in kDa.

molecule (Figure 3.4).

To assess decaying, the protein was subjected to UV irradiation with a transilluminator. The protein was irradiated for 10 minutes at a variety of wavelengths.

bioconjugation failed to couple the two molecules, requiring further optimization of reaction conditions. The first experimental variable assessed was temperature, and the reaction was incubated at 22° C, a 37° C incubator, and at 4° C. Reaction time was also examined, as the degradation of protein via copper has been previously observed with this particular reaction. It was determined by SDS-PAGE analysis that a reaction at 4° C for 4 hours yielded the bioconjugated

Partial decaging was observed after irradiation at 254 nm, but not at higher wavelengths. When irradiated at 254 nm for 30 minutes, a greater degree of decaging was observed by gel electrophoresis. It was originally hypothesized that decaging would sever the protein at the site of UAA incorporation, resulting in two protein fragments (Figure 3.2). However, a total protein degradation was observed when analyzed by SDS-PAGE (Figure 3.4). It is possible that the streak is due to radicals produced by decaging, which damaged the protein as well as proteases, which would cleave the protein. The degradation is not necessarily problematic, as if this UAA is coupled to a drug, the drug would have already been delivered into the desired location before severing, and would not be affected by the proteases.

Further work will be needed to optimize the decaging of GFP containing **11** in order to achieve a greater level of control. The mechanism could be further understood by incorporating **11** at different residues of GFP, conjugating to a dye, decaging, and observing the bands formed via SDS-PAGE. Also, the decaging conditions need to be

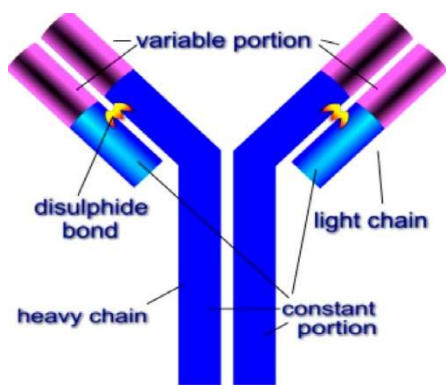


Figure 3.5: Structure of an antibody. The variable portions bind with high specificity to particular antigens.<sup>55</sup>

further optimized, to potentially minimize degradation from the high energy UV light. After optimization, **11** can be incorporated into an antibody and be conjugated by a small drug molecule. Antibodies are immune proteins that have two variable regions which are highly discriminatory, and bind to specific antigens (Figure 3.5).<sup>55</sup> The antibody

can be designed to reach the drug's target.<sup>55</sup> Once at the desired location, the antibody-drug complex can be photocleaved, allowing the drug to be released. This would allow less

dosage to be utilized, as the drug is carried directly to its target. Also, this methodology would be especially beneficial for cancer drugs, as they are expensive to make, and often kill non-cancer cells.<sup>56</sup> With this strategy, antibodies can be generated that target cells with cancer markers, bringing a drug directly to the cancer cells.<sup>57</sup> Also, the drug could be released at various intervals, depending on the strength of light. Alternatively, caging groups that decay at different wavelengths could be synthesized, providing another method for controlled drug release.

## Experimental

**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. Microwave reactions were conducted in a CEM Discovery microwave. NMR data was acquired on a Varian Gemini 400 MHz. Fluorescence data was measured using a PerkinElmer LS 55 Luminescence Spectrometer and a Synergy Microplate reader. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit. Glaser-Hay deprotection was performed with a VWR UV-AC hand lamp.

**General GFP Expression:** A pET-GFP-TAG plasmid (0.5  $\mu$ L) was co-transformed with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electroporator. The cells were then plated (100  $\mu$ L) 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 Chlr) at an OD<sub>600</sub> 0.1. The cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8, and protein expression was induced by addition of the respective unnatural amino acid (100 µL, 100 mM) and 20 % arabinose (10 µL) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 10 µL). The cultures were incubated at 30° C for 16-20 h then pelleted 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 added to an equilibrated column of Ni-NTA resin (200 µL) and the GFP was purified according to manufacturer's protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h), and employed without further purification.

**General aaRS Assessment:** The aaRSs were assessed for unnatural amino acid incorporation by first co-transforming a pET-GFP-TAG variant plasmid (0.5 µL) with a pEVOL-tRNA synthetase plasmid (0.5 µL), in a similar method as described previously. For each pEVOL plasmid, one colony was used to inoculate LB media (4 mL) in separate containers that contained ampicillin (4 µL, 50 mg/mL) and chloramphenicol (4 µL, 34 mg/mL). The cultures were incubated at 37° C overnight. Expression cultures were created containing LB media (4 mL), ampicillin (4 µL, 50 mg/mL), and chloramphenicol (4 µL, 34 mg/mL), and then were inoculated with the previous culture (250 µL). The expression cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8. Protein expression was induced by addition of 20 % arabinose (4 µL), and 0.8 mM IPTG (4 µL). The cultures were then plated on a 96-well plate (200 µL), with six wells for each plasmid. Three wells were

designated as controls, and the other three were designated for the experiment. The designated UAA (100  $\mu$ L, 100 mM) was added to the experimental wells only. The plate was allowed to shake at 30° C for 16 h, after which it was measured for fluorescence (528 nm and 485 nm) utilizing a Synergy HT microplate reader.

**(3-methoxy-4-(prop-2-yn-1-yloxy)benzaldehyde), 6:** The synthesis was conducted using literature conditions.<sup>53</sup> Vanillin (1.5 g, 9.86 mmol, 1 equivalent) was stirred with Cs<sub>2</sub>CO<sub>3</sub> (3.213 g, 9.86 mmol, 1 eq) for five minutes in anhydrous DMF (25 mL). To the mixture was added propargyl bromide (3.742 mL, 49.3 mmol, 5 equivalents). The reaction was stirred at room temperature for 24 hours. The reaction was then quenched with dI H<sub>2</sub>O (30 mL) and extracted with EtOAc (3x30 mL) and back extracted with brine (3x80 mL). The organic layer was dried with Mg<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated by rotatory evaporation. The resulting solution was purified via flash chromatography (3:1 Hexanes: EtOAc), resulting in a white solid (1.521 g, 8.00 mmol, 81% yield).

**(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)methanol, 7:** A flask was charged with **6** (1.5 g, 8.00 mmol, 1 eq), wrapped in aluminum foil and incubated on ice. Chilled HNO<sub>3</sub> (50 mL, excess) was added to the flask and allowed to stir for 25 minutes at 0° C. The reaction was then warmed to room temperature for 2 hours. The reaction was quenched with chilled dI H<sub>2</sub>O (100 mL) and the precipitate was collected by vacuum filtration and washed with ice water (3x30 mL). The resulting product was a yellow solid (1.534 g, 6.52 mmol, 81% yield) and required no further purification. <sup>1</sup>H NMR (400MHz; CDCl<sub>3</sub>):  $\delta$  10.45 (s, 1H), 7.79 (s, 1H), 7.43 (s, 1H), 4.90 (s, 2H), 4.02 (s, 3H), 2.62 (s, 1H); <sup>13</sup>C NMR (400MHz; CDCl<sub>3</sub>):  $\delta$  187.9, 153.9, 146.7, 140.2, 126.6, 110.5, 109.6, 78.0, 76.5, 57.4, 57.0.

**(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)methanol, 8:** Compound **7** (1.53 g, 6.52 mmol, 1 eq) was dissolved in EtOH (120 mL) and wrapped in aluminum foil. NaBH<sub>4</sub> (740 mg, 19.6 mmol, 3 eq) was dissolved in NaOH (50 mL, 1 M). The resulting solution was then added to **7** and allowed to stir for 2.5 hours. The reaction was neutralized by HCl (50 mL, 1M), and extracted with EtOAc (3x30 mL) and back extracted with brine (3x70 mL). The organic layer was dried with Mg<sub>2</sub>SO<sub>4</sub> and concentrated using a rotary evaporator. The resulting compound was a pale yellow solid (1.78 g, 6.53 mmol, 76% yield) and required no further purification. <sup>1</sup>H NMR (400MHz; CDCl<sub>3</sub>): δ 7.89 (s, 1H), 7.22 (s, 1H), 4.98 (s, 2H), 4.84 (s, 2H), 4.01 (s, 3H), 2.58 (s, 1H), 1.56 (s, 1H); <sup>13</sup>C NMR (400MHz; CDCl<sub>3</sub>): δ 153.9, 146.7, 140.22, 126.6, 111.1, 78.0, 76.5, 63.0, 57.3, 29.9.

**1-(bromomethyl)-5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)benzene, 9:** A round bottom flask was charged with **8** (1.18 g, 4.96 mmol, 1 equivalent), DCM (40 mL) and was wrapped in foil and incubated in an ice bath. PBr<sub>3</sub> (1.50 mL, 14.9 mmol, 3 equivalent) was added to the reaction vessel and stirred for 6 hours on ice. The reaction was quenched with 15 mL of dI H<sub>2</sub>O, extracted with EtOAc (3x30 mL) and then back extracted with brine (3x70 mL). The organic layer was dried with Mg<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated via rotatory evaporation. The resulting solution was purified by flash chromatography (1:1 Hexane: EtOAc), yielding a yellow solid (739 mg, 2.46 mmol, 50% yield). <sup>1</sup>H NMR (400MHz; CDCl<sub>3</sub>): δ 7.85 (s, 1H), 6.98 (s, 1H), 4.87 (s, 2H), 4.84 (s, 2H), 4.01 (s, 3H), 2.60 (s, 1H), <sup>13</sup>C NMR (400MHz; CDCl<sub>3</sub>): δ 153.9, 146.7, 140.2, 128.7, 114.3, 111.3, 57.2, 56.8, 29.9.

**Diethyl 2-acetamido-2-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)benzyl)malonate, (10):** Diethyl acetamidomalonate (106 mg, 0.76 mmol, 2 eq) and Cs<sub>2</sub>CO<sub>3</sub> (114 mg, 0.35

mmol, 1 eq) was dissolved in CH<sub>3</sub>CN (1 mL) for five minutes. Compound **9** (106 mg, 0.35 mmol, 1 eq) was dissolved in CH<sub>3</sub>CN (2 mL), and added to the diethyl acetamidomalonate solution in a microwave vial. The solution was subjected to microwave irradiation in standard mode (130° C, 300W, 10 minutes) and quenched with dI H<sub>2</sub>O (5 mL). The reaction was extracted with ethyl acetate (3x15 mL) and back extracted with brine (3x60 mL), dried with Mg<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated with rotatory evaporation. The resulting oil was purified with flash chromatography (1:1 hexanes: EtOAc), yielding a peach-colored solid (122 mg, 0.28 mmol, 79% yield). <sup>1</sup>H NMR (400MHz; CDCl<sub>3</sub>): δ 7.69 (s, 1H), 6.76 (s, 1H), 6.52 (s, 1H), 4.80 (s, 2H), 4.33-4.13 (m, 4H), 4.10 (s, 3H), 3.90 (s, 2H), 2.59 (s, 1H), 1.94 (s, 3H), 1.29 (t, J = 7.2, 6H); <sup>13</sup>C NMR (400MHz; CDCl<sub>3</sub>): δ 169.7, 167.8, 152.9, 145.7, 142.1, 126.1, 116.0, 115.9, 78.0, 76.6, 66.1, 63.0, 57.1, 56.5, 35.2, 14.0, 14.0.

**1-carboxy-2-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethan-1-aminium,**

**(11):** Compound **10** (96.3 mg, 0.22 mmol, 1 eq) was stirred with 6N HCl (3 mL) under microwave irradiation in standard mode (90° C, 300W, 10 mins). The reaction was concentrated by rotatory evaporation, yielding a white solid (62 mg, 0.21 mmol, 95% yield). <sup>1</sup>H NMR (400MHz; CDCl<sub>3</sub>): δ 7.88 (s, 1H), 7.11 (s, 1H), 4.86 (d, J=6 Hz, 2H), 4.38-4.31 (m, 1H), 3.78 (s, 3H), 3.76-3.70 (m, 1H), 3.43-3.33 (m, 1H), 3.05 (s, 1H); <sup>13</sup>C NMR (400 MHz; CD<sub>3</sub>OD): δ 169.8, 168.3, 152.4, 146.3, 141.5, 122.9, 114.9, 112.1, 76.4, 56.1, 53.1, 39.7, 34.1

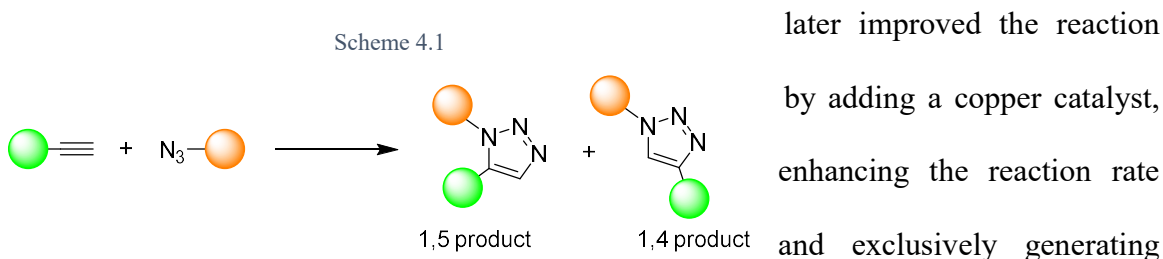
**Glaser-Hay Bioconjugation and Decaging:** GFP containing **11** at residue 151 (20 μl), AlexaFluor-488 Alkyne (15 μl, 1 mM), CuI (5 μl, 50mM), and tetramethylethylenediamine (5 μl, 50 mM) were added to an Eppendorf tube. A control reaction omitting the CuI and

TMEDA was conducted simultaneously. The tubes were incubated at 4° C for 4 hours. In order to remove excess copper and fluorophore, a buffer exchange was performed with Spin-X UF concentrator columns, and with PBS (8x200 µL) to a final volume of 50 µL. The GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h). Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), first via a Spyro Ruby irradiation to analyze fluorescence. Then, the gel was stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager. for protein via the Coomassie setting on the BioRad Molecular Imager.

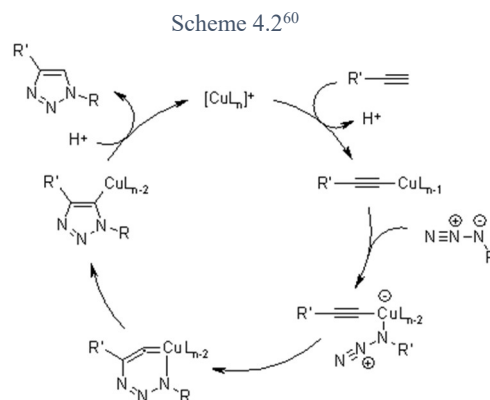
Decaging of GFP was achieved by transferring the conjugated GFP containing **11** Glaser-Hay (15 µL) in a clear eppendorf tube. A control was made by adding GFP wild type (15 µL) to a clear Eppendorf tube. Both tubes were irradiated at 254 nm for 30 mins using a VWR UV-AC hand lamp. The GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h). Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), first via a Spyro Ruby irradiation to analyze fluorescence. Then, the gel was stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

## Chapter 4: A Copper-Free Microwave 1,3-Dipolar Cycloaddition System

1,3-Dipolar Cycloadditions are highly selective reactions between alkyne and azide functional groups. First characterized by Rolf Huisgen in 1963, his conditions generated both the 1,4 and 1,5 adducts when reacted at 98° C for 18 hours (Scheme 4.1).<sup>58</sup> Sharpless



the 1,4 isomer (Scheme 4.2).<sup>59,60</sup> The utility of this reaction is highlighted by its use in complex biological systems, with relatively high specificity. Applications include its use to design HIV protease inhibitors, *in vivo* fluorescence labeling, creating new drugs, linking proteins together, and cell and protein immobilization.<sup>61</sup> Click chemistry has additionally been coupled with microwave technology to achieve great yields with regioselectivity, in a matter of minutes.<sup>62</sup>



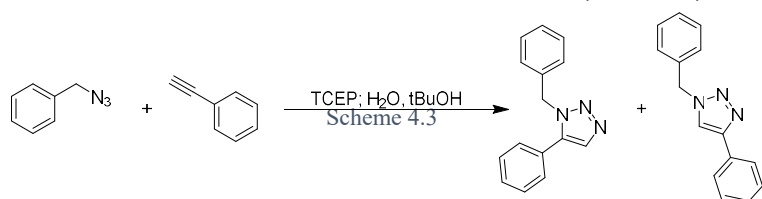
One problem associated with utilizing this reaction in biological environments is the toxicity associated with copper ions. Copper (II) ions react with hydrogen peroxide to

form a hydroxyl radical, which is a very powerful oxidizing agent. It has been shown to oxidize proteins and DNA, and disrupt the intermolecular interactions essential within large biomolecules.<sup>63</sup> Copper (II) protein degradation has also been associated with diabetes and aging via autoxidative glycosylation<sup>64</sup>, as well as degradation of albumin.<sup>65</sup>

Copper-free [3+2] cycloadditions between azides and acyclic alkynes have been demonstrated in cyclooctynes<sup>66</sup> and activated oxanorbornadienes<sup>67</sup>. Other researchers have utilized inverse-electron-demand Diels-Alder reactions to attach *trans*-cyclooctene compounds to a bipyridyltetrazine-tethered probe.<sup>68</sup> Also, photoinducible dipolar cycloadditions (photo-click reactions) between nitrile imines and alkenes have been examined.<sup>69</sup> These reactions however, take advantage of ring strain, and often require synthesizing complicated cyclic molecules, in the case of the [3+2] cycloadditions, or require photo-reactive molecules such as tetrazoles, in the case of the photo-click reactions.<sup>70</sup> We seek to develop a method to undergo copper-free 1,3-dipolar cycloadditions between alkynes and azides, regardless of structure, in high yield.

## Results and Discussion:

The optimized microwave-mediated 1,3 dipolar cycloaddition required many trials altering microwave conditions, and required substantial troubleshooting of separation and purification techniques. Proof-of-concept experiments were performed using benzyl azide (21  $\mu$ L, 0.17 mmol, 1 equiv.) and phenylacetylene (18.6  $\mu$ L, 0.17 mmol, 1 equivalent), as well as identical amounts of deionized water, tBuOH, and TCEP (Scheme 4.3) similar to



literature conditions. Note that copper was not used as a

reagent for these reactions. After microwave irradiation, the reaction was extracted in ethyl acetate and brine; however, it was theorized that the miscibility of tBuOH could leave the click products in the aqueous layer, reducing the experimental yield, and forcing an evaluation of organic extraction solvents. Extracting the reaction in dichloromethane produced similar yields, as ethyl acetate. Consequently, the solvent was then immediately removed after microwave irradiation by rotary evaporation, to remove both water and tBuOH, and then subjected to flash chromatography. This proved to afford a higher yield, and thus was adopted as a general protocol for all additional reactions.

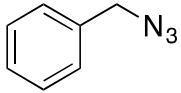
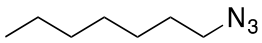
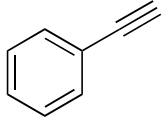
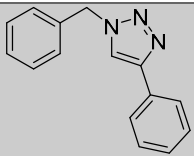
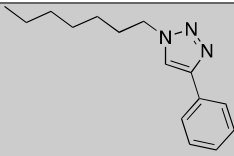
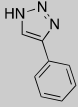
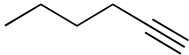
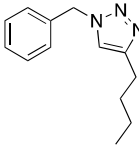
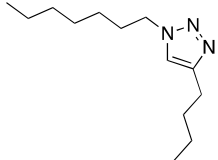
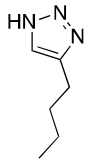
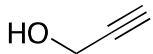
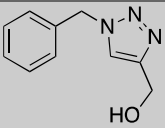
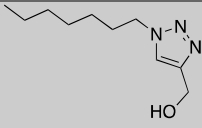
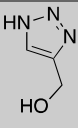
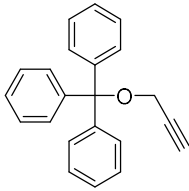
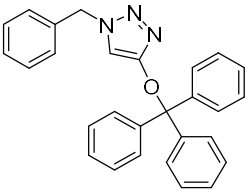
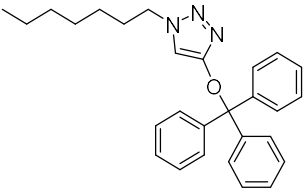
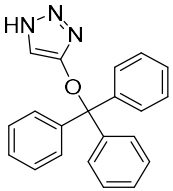
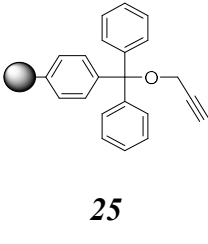
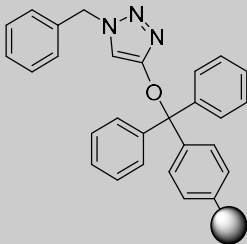
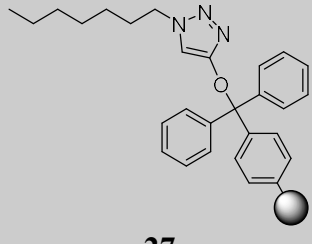
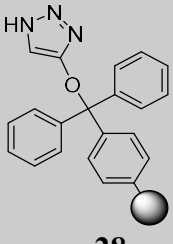
The microwave reaction was first attempted at 100 W, 100° C, for 10 mins on temperature mode, generating a poor yield of 21.9 %. However, a control reaction of phenylacetylene and benzyl azide in a 100° C oil bath resulted in an extremely low yield (less than 5 %), representing some advantage of the microwave. The temperature was raised to 125° C with the same conditions, resulting in a slightly higher yield of 26.0 %. Reaction time was then increased to 20 mins (at 100 W and 125° C), with yet another slight increase in yield of 31.4 %. Further attempts to increase temperature and time resulted in no significant improvement in yield. The reaction was then attempted in the power mode setting, which applies a constant power while allowing the temperature to fluctuate unless it achieves a certain limit at which the reaction is halted, which in this case was 168° C. Attempting this at 100 W and 200 W resulted in yields of 23.0 % and 26.0 % respectively. Interestingly, increasing the power to 300 W resulted in a significant increase in yield to 49.2 %, despite the reaction only proceeding for 1 minute. To probe deeper, the reaction was run twice at 300 W, resulting in a significantly lower yield of 21.9 %. This decreased yield could be due to extreme conditions that decomposed the reagents or the click product.



Another microwave procedure was investigated in order to generate higher yields. The peptide synthesis (SPS) procedure irradiates at a constant power until it reaches a specific temperature, where it then stops irradiation until the temperature drops to an acceptable range, and then it irradiates again. This cycle occurs for a set amount of time. For the click reactions, the maximum temperature was set to 168° C, the temperature at which irradiation again proceeds (denoted as  $dT$ ) was set for 15° C, and the total time was set for 20 minutes. This procedure has an advantage that it is able to irradiate at high power for a much longer length of time than power mode. Irradiation at 200 W on SPS mode resulted in a 48.1 % yield, while irradiation at 300 W on SPS mode resulted in a very high yield of 96.2 %. When performing a similar microwave reaction in the presence of a copper sulfate catalyst at 300 W on SPS mode, a similar yield was recorded.

Having obtained optimized reaction conditions, a series of azides and alkynes were examined utilizing the 300W SPS mode in order to elucidate the scope of the reaction (Table 4.1). It was determined that using a long aliphatic azide significantly reduced yield, regardless of the structure of the alkyne. Azidoheptane was reacted with phenylacetylene, 1-hexyne, and propargyl alcohol, producing yields of 19.3%, 15.8 %, and 14.5 % respectively. It was originally theorized that long chained aliphatic azides may not be well suited for the water and tBuOH solvents, or that they were microwave transparent as they possess very small dipoles. Therefore, to investigate the microwave suitability of nonpolar hydrocarbons, 1-hexyne was utilized as the alkyne reactant, undergoing reactions with benzyl azide and trimethylsilyl azide, and generating yields of 96.8 % and 37.6 % respectively. It should be noted that when using trimethylsilyl azide as a reactant, NMR of the click product indicated a TMS group; however, the acidity of the silica gel from flash

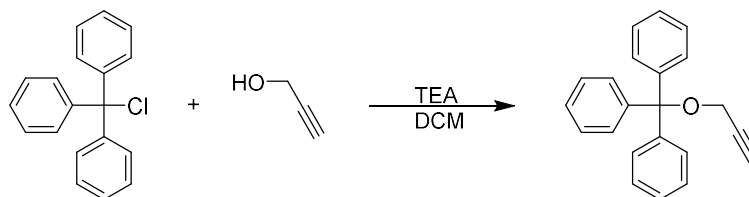
Table 4.1: A series of azides and alkynes were experimented using the SPS copper-free method. Note that the yields for compounds **26-28** represent the cleaved resin product (see experimental section).

			TMS-N <sub>3</sub>
	 <b>12</b> 96.2 %	 <b>13</b> 19.3 %	 <b>14</b> 64.8 %
	 <b>15</b> 96.8 %	 <b>16</b> 15.8 %	 <b>17</b> 37.6 %
	 <b>18</b> <5 %	 <b>19</b> <5 %	 <b>20</b> <5 %
	 <b>22</b> 14.7 %	 <b>23</b> 10.7 %	 <b>24</b> 12.1 %
	 <b>26</b> 75.5 %	 <b>27</b> 63.4 %	 <b>28</b> 81.3 %

chromatography cleaved the TMS group. The higher yields, especially with benzyl azide, create a more complex dynamic regarding the suitability of aliphatic hydrocarbons. It appears that aliphatic azides produce much lower yields, and the nature of the hydrocarbon chain of the alkyne molecule is somewhat independent of the yield. Moreover, it suggests that the aliphatic azide may simply be less stable under microwave irradiation and prone to degradation, resulting in decreased yields; however, this premise needs further testing.

Reactions with polar alkynes also failed to produce the desired products in appreciable yields. Propargyl alcohol was reacted with benzyl azide, azidoheptane, and trimethylsilyl azide, resulting in yields of less than 5 %. The most striking observation is that the reaction with benzyl azide resulted in a low yield, as the previous benzyl azide reactions with phenylacetylene and 1-hexyne have yields of 96.2 % and 96.8 % respectively. Propargyl alcohol was then protected with trityl chloride, to see if a protected alcohol would generate higher yields (Scheme 4.4). Fortunately, yields improved to 10-15 %, indicating that perhaps more non-polar molecules will react at higher yields.

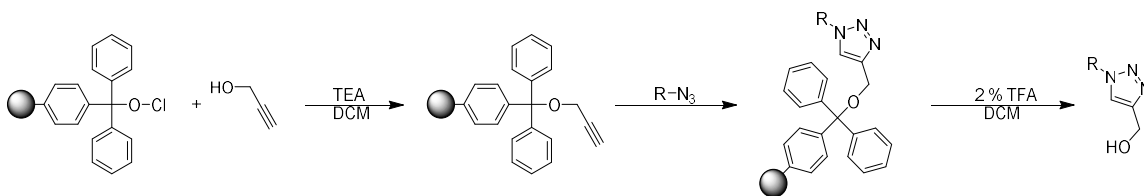
Scheme 4.4



Propargyl alcohol was then immobilized on a solid-supported resin to potentially confer higher microwave stability (Scheme 4.5). The protected alkyne was then reacted under similar microwave conditions with benzyl azide, azidoheptane, and trimethylsilyl azide, and then deprotected. Gratifyingly, the yields were much higher in generating the

propargyl alcohol click product, around 60 – 70 %. Unfortunately, due to the nature of the solid support, only 1 – 2 mgs, of product was obtained, making the determination of exact yields difficult to obtain. However, NMR spectra confirmed the presence of the product in high purity.

Scheme 4.5



To further probe the scope of this methodology, a series of molecularly distinct azides were examined with the same series of alkynes (Figure 4.1). The first azide investigated was 4-azidobenzoic acid (**29**), a highly polar azide to compare to the previously non-polar substrates. Utilizing the 300 W SPS conditions, very little product was obtained for any of the reactions, and the click products were unable to be separated via flash chromatography. One explanation for the poor yields could be that 4-azidobenzoic acid is an unstable molecule, and the microwave irradiation decomposed the reactant or click product. When reacted on power mode at various powers, similar low yields were observed. Two unnatural amino acid azide molecules were then tested under microwave

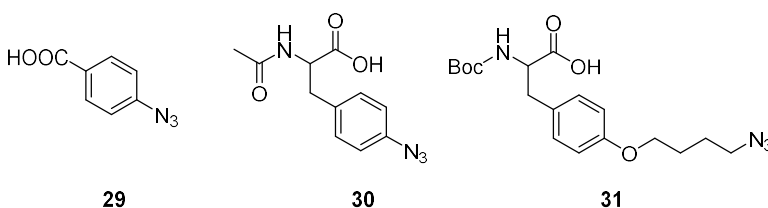


Figure 4.1: Three molecularly distinct azides tested for their propensity to undergo copper-free click reactions. **29** is commercially available azidobenzoic acid, **30** is protected para-azidophenylalanine, **31** is a protected tyrosine tethered azide.

conditions. The para-azidophenylalanine (**30**) unnatural amino was not able to successfully react with any of the previous

alkynes. This could be due to the nature of the aromatic azide providing unfavorable steric and electronic conditions for the reaction. To account for sterics, a protected C4 tethered azido tyrosine unnatural amino acid (**31**) was also utilized. However, no product was obtained when reacted with any alkyne. Although further tests need to be performed, polar molecules appear not to be able to undergo copper-free click reactions. This would explain the inability of the two unnatural amino acids to react, as the carboxyl C-terminus is extremely polar.

One of the more intriguing aspects of the microwave copper-free click chemistry was the ratio of 1,4 and 1,5 adducts. When copper sulfate was included in the microwave conditions, all reactions produced the 1,4 adduct. Without copper sulfate however, the ratio of adducts becomes complex. For example, when reacting benzyl azide, azidoheptane, and TMS-azide, with 1-hexyne, the respective ratio of 1,4 adducts to 1,5 adducts was 1:6, 1:1.7, and 3:1. Interestingly, when reacting the above azides with the resin protected propargyl alcohol, the ratios for the three click products were 1:3, in favor of the 1,5 adduct. For this particular example, it's possible that there is steric hindrance for the 1,4 position to form due to the triphenyl groups; however, the chemical properties of the azide did not affect the ratio.

In conclusion, it was demonstrated that microwave copper-free click chemistry is feasible, and can occur with high yields. Further investigation is needed to understand the molecular characteristics that control the reaction in the absence of copper catalysts. A range of polar azides and alkynes will need to be examined, in order to optimize polar microwave click reactions. These conditions include altering the cycle temperature range ( $dT$ ), reacting at different equivalents of alkynes and azides, and utilizing different solvents.

However, immobilizing polar compounds onto a solid-supported resin is an effective solution, as demonstrated from the high yields obtained from the immobilized propargyl alcohol. Also, similar experiments are needed to understand the conditions needed to obtain regioselectivity in the reactions.

Copper-free clicks are especially relevant *in vivo*, as copper has been demonstrated to aid in radical formation and protein degradation. Click chemistry is an important tool for bioconjugation, and the ability to click without cytotoxic copper would allow for safer click bioconjugation in living cells. However, one important factor is that proteins (and living cells) are not viable in microwaves, as protein degrades under the high temperature. Currently, we are working on adapting these conditions to proteins by conducting the reaction at much lower temperatures, as will be described in the next chapter. Moreover, microwave copper-free Cadiot-Chodkiewicz and Glaser-Hay couplings are also being explored.

## **Experimental:**

**General:** Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Reactions were conducted using a CEM Discovery Microwave. NMR data was acquired on a Varian Gemini 400 MHz. GC/MS analysis was conducted on an Agilent Technologies 6890N GC system interfaced with a 5973N mass selective detector. An Agilent J&W GC capillary column (30 m length, 0.32 mm diameter, 0.25  $\mu$ m film) was employed with a splitless injection (250° C inlet, 8.8 psi) with an initial 70 oC hold (2 min) and ramped for 15 min to 230° C.

**General 1,3-dipolar Cycloaddition Method:** To a microwave vial was added a small stir bar, an alkyne ( $1.7 \times 10^{-4}$  moles; 1 eq.), an azide ( $1.7 \times 10^{-4}$  moles; 1 eq.), *tris*(2-carboxyethyl)phosphine (TCEP; 50 mmol, 32  $\mu$ L), deionized water (970  $\mu$ L), and tert-Butyl alcohol (99 %, 1 mL). The solution was subjected to microwave irradiation (SPS mode, 20 minutes, 0-300 W, 168° C,  $\delta T = 15^\circ$  C). The resulting mixture was then quenched with chilled dI water (10 mL, 0° C). A rotary evaporator was used to remove water from the reaction, and the crude product was purified using flash chromatography (3:1 hexanes:EtOAc) to yield the product.

**1-benzyl-4-phenyl-1H-1,2,3-triazole, 12:** The solvent was removed in vacuo to give **12** as a solid (38 mg, 0.162 mmol, 96.2 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.798 (dd,  $J = 8.4$  Hz, 2H), 7.744 (s, 1H), 7.661 (s, 1H), 7.445-7.365 (m, 3H), 7.334-7.240 (m, 4H), 7.085-7.062 (m, 2H), 5.581 (s, 2H), 5.548 (s, 2H);  $^{13}\text{C}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  135.489, 133.289, 129.510, 129.160, 128.948, 128.895, 128.827, 128.796, 128.151, 127.150, 125.693, 119.470, 54.240, 51.804; GCMS ( $R_t = 12.54$ ) calculated for  $\text{C}_{15}\text{H}_{13}\text{N}_3$  235.29, found 235.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 1.1:1.

**1-heptyl-4-phenyl-1H-1,2,3-triazole, 13:** The solvent was removed in vacuo to give **13** as a solid (8 mg, 0.033 mmol, 19.3 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.834 (dd,  $J = 8.45$  Hz, 2H), 7.738 (s, 1H), 7.688 (s, 1H), 7.502-7.366 (m, 3H), 4.398 (t, 2H), 4.336 (t, 2H), 1.929 (quintet, 2H), 1.817 (quintet 2H), 1.373-1.165 (m; 8H), 0.896-0.824 (m, 3H);  $^{13}\text{C}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  129.358, 129.047, 128.751, 125.685, 48.291, 31.475, 30.071, 28.523, 26.352, 22.452, 13.961; GCMS ( $R_t = 13.77$  min) calculated for  $\text{C}_{15}\text{H}_{21}\text{N}_3$  243.35, found 243.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 2:1

**4-phenyl-1H-1,2,3-triazole, 14:** The solvent was removed in vacuo to give **14** as a solid (16 mg, 0.110 mmol, 64.8 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.976 (s, 1H), 7.826 (d, 2H), 7.479-7.367 (m, 3H);  $^{13}\text{C}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  128.971, 128.773, 126.110; GCMS ( $R_t = 8.71$  min) calculated for  $\text{C}_8\text{H}_7\text{N}_3$  145.16, found 145.1. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:1.2

**1-benzyl-4-butyl-1H-1,2,3-triazole, 15:** The solvent was removed in vacuo to give **15** as a solid (35 mg, 0.163 mmol, 96.8 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.939 (s, 1H), 7.405-7.388 (m, 2H), 7.378-7.263 (m, 3H), 5.560 (s, 2H), 3.090 (t, 2H), 1.767 (q, 2H), 0.995 (3, 2H), 0.914 (q, 3H);  $^{13}\text{C}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  129.335, 129.183, 128.349, 125.230, 54.475, 41.400, 17.413, 13.801; GCMS ( $R_t = 11.67$  min) calculated for  $\text{C}_{13}\text{H}_{17}\text{N}_3$  215.30, found 215.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:6.

**4-butyl-1-heptyl-1H-1,2,3-triazole, 16:** The solvent was removed in vacuo to give **16** as a solid (6 mg, 0.027 mmol, 15.8 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.729-7.698 (m, 1H), 7.542-7.520 (m, 1H), 4.297 (t, 2H), 4.235 (t, 2H), 1.721 (quintet, 2H), 1.413-1.259 (m, 14H), 0.988-0.849 (m, 6H);  $^{13}\text{C}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  130.883, 128.827, 65.865, 29.282, 28.917, 28.538, 25.950, 25.609, 23.742, 22.520, 14.021; GCMS ( $R_t = 10.49$  min) calculated for  $\text{C}_{13}\text{H}_{25}\text{N}_3$  223.36, found 223.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:1.7.

**4-butyl-1H-1,2,3-triazole, 17:** The solvent was removed in vacuo to give **17** as a solid (8 mg, 0.064 mmol, 37.6 %)  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  8.183 (s, 1H), 3.042 (t, 2H), 1.838-1.683 (m, 4H), 1.012 (t, 2H); GCMS ( $R_t = 8.78$  min) calculated for  $\text{C}_6\text{H}_{11}\text{N}_3$  125.18, found 125.1. Regioisomer ratio of 1,4:1,5 adducts was found to be 3:1.



**((prop-2-yn-1-yloxy)methanetriyl)tribenzene, 21:** To a flame dried vial was added (chloromethanetriyl)tribenzene (250 mgs, 0.90 mmol, 1.0 equiv.), propargyl alcohol (76  $\mu$ L, 1.31 mmol, 1.5 equiv.), and triethylamine (TEA; 624  $\mu$ L). The reaction was allowed to stir at room temperature for 2 days. The product was extracted with dichloromethane (3 x 20 mL), and the organic layer was then back-extracted with brine (3 x 30 mL). The organic layer was then isolated, dried with MgSO<sub>4</sub>, filtered, and the solvent was removed by rotatory evaporation. The product was purified via flash chromatography with a 3:1 hexanes/ethyl acetate eluent, and the solvent was removed *in vacuo* to give **21** as a solid (243 mgs,  $8.14 \times 10^{-4}$  mol, 90.7 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.536 (2, 4H), 7.371-7.261 (m, 20H), 3.811 (d, J=2.8 Hz, 2H), 2.866 (s, 1H).

**1-benzyl-4-(trityloxy)-1H-1,2,3-triazole, 22:** The solvent was removed in vacuo to give **22** as a solid (10 mg,  $2.33 \times 10^{-5}$  mol, 13.7 %). <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.564 (s, 1H), 7.478 (d, 4H), 7.472-7.177 (m, 16 H), 6.987-6.965 (m, 4H), 5.522 (s, 2H), 5.495 (s, 2H), 4.309 (s, 2H), 4.083 (s, 2H).

**1-heptyl-4-(trityloxy)-1H-1,2,3-triazole, 23:** The solvent was removed in vacuo to give **23** as a solid (8 mg,  $1.82 \times 10^{-5}$ , 10.7 %). <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.546 (s, 1H), 7.471-7.440 (m, 5H), 7.351-7.260 (m, 10H), 4.189 (s, 2H), 4.163 (t, 2H), 1.763 (quintet, 2H), 1.254-1.198 (m, 8H), 0.853 (t, 3H).

**4-(trityloxy)-1H-1,2,3-triazole, 24:** The solvent was removed in vacuo to give **24** as a solid (7 mg,  $2.06 \times 10^{-5}$  mol, 12.1 %). <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.605 (d, 2H), 7.585 (d, 2H), 7.547 (d, 4H), 7.342-7.230 (m, 10 H), 3.758 (s, 2H), 3.725 (s, 2H).

**Immobilization of Propargyl Alcohol onto Trityl Chloride Resin in Low Loading, 25:**

To flame dried vial was added trityl chloride resin (200 mg, 0.36 mmol, 1 equiv.) and dichloromethane (5 mL). The resin was swelled at room temperature with gentle stirring for 15 min. Propargyl alcohol (25.0  $\mu\text{L}$ , 0.433 mmol, 1.2 equiv.) was added to reaction, followed by triethylamine (10.0  $\mu\text{L}$ , 0.072 mmol, 0.2 equiv). 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 DCM and dried under vacuum for 45 min before further use.

**Microwave Click using Solid-Support Propargyl Alcohol:** To a microwave vial was added **25** (40 mgs,  $7.2 \times 10^{-6}$  mol, 1 equiv.), an azide ( $7.2 \times 10^{-5}$  mol, 10 equiv.), *tris*(2-carboxyethyl)phosphine (TCEP; 50 mmol, 32  $\mu\text{L}$ ), deionized water (970  $\mu\text{L}$ ), and tert-Butyl alcohol (99 %, 1 mL). The solution was subjected to microwave irradiation (SPS mode, 20 minutes, 0-300 W, 168° C,  $\delta T = 15^\circ$  C). The resulting mixture was quenched with chilled deionized water (5 mL, 0° C). The resin was transferred to a syringe filter and washed with DCM and MeOH (5 alternating rinses with 5 mL each).

**Compound 26:** Compound **26** was then cleaved from the resin by treatment with 2 % TFA (DCM, 1 h), and filtered into a vial. A short silica plug was then employed to remove unreacted starting material (1:1 EtOAc/Hex). The solvent was removed *in vacuo* to give **17** as a solid (1.0 mg,  $5.28 \times 10^{-6}$  mol, 73.5 %).  $^1\text{H}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  7.414-7.241 (m, 5H), 5.658 (s, 2H), 5.532 (s, 2H), 5.313 (s, 1H), 4.752 (s, 2H), 4.610 (s, 2H);  $^{13}\text{C}$  NMR (400 MHz;  $\text{CDCl}_3$ ):  $\delta$  147.266, 129.327, 129.214, 129.115, 128.736, 128.295, 127.635, 122.225, 55.348, 54.847, 53.291, 52.730, 29.699; GCMS ( $R_t = 10.84$  min) calculated for

$C_{10}H_{11}N_3O$  189.22, found at 189.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:3.

**Compound 27:** Compound **27** was then cleaved from the resin by treatment with 2 % TFA (DCM, 1 h), and filtered into a vial. A short silica plug was then employed to remove unreacted starting material (1:1 EtOAc/Hex). The solvent was removed *in vacuo* to give **18** as a solid (0.9 mg,  $4.56 \times 10^{-6}$  mol, 63.4 %).  $^1H$  NMR (400 MHz;  $CDCl_3$ ):  $\delta$  7.684 (s, 1H), 7.578 (s, 1H), 5.486 (s, 1H), 4.788 (d,  $J=4$  Hz, 2H), 4.418-4.352 (m, 2H), 1.930 (quintet, 2H), 1.346-1.250 (m, 8H), 0.895-0.860 (m, 3H);  $^{13}C$  NMR (400 MHz;  $CDCl_3$ ):  $\delta$  122.172, 55.302, 53.253, 51.106, 31.558, 31.497, 30.078, 28.674, 28.576, 26.512, 26.345, 22.490, 13.991; GCMS ( $R_t = 11.72$ ) calculated for  $C_{10}H_{19}N_3O$  197.28, found 196.2. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:3.

**Compound 28:** Compound **28** was then cleaved from the resin by treatment with 2 % TFA (DCM, 1 h), and filtered into a vial. A short silica plug was then employed to remove unreacted starting material (1:1 EtOAc/Hex). The solvent was removed *in vacuo* to give **19** as a solid (1.0 mg,  $5.84 \times 10^{-6}$  mol, 81.3 %).  $^1H$  NMR (400 MHz;  $CDCl_3$ ):  $\delta$  7.637 (s, 1H), 7.519 (s, 1H), 4.804 (s, 2H); GCMS ( $R_t = 10.84$  min) calculated for  $C_6H_{13}N_3OSi$  171.28, found 171.1. Regioisomer ratio of 1,4:1,5 adducts was found to be 1:3.

## Chapter 5: Developing Microwave Bioconjugations via CoolMate Technology

Microwave technology offers a fast and efficient method for chemical synthesis; however, it is not applicable for most biological applications, especially *in vivo*. This is because microwave irradiation results in extremely high localized temperatures, often exceeding 100° C. In humans, physiological temperature is at 37° C, and when this

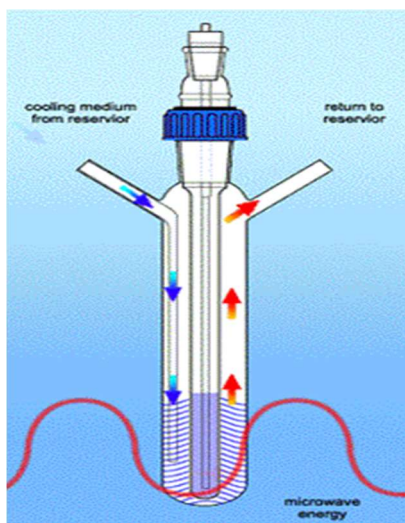


Figure 5.1: Diagram of a CoolMate vessel. A vial is kept cooled via a cooling medium, while simultaneously receiving microwaving irradiation.<sup>72</sup>

temperature is even moderately altered, macromolecules such as proteins and DNA begin to denature, and are no longer viable.<sup>1,2</sup> Despite this fact, some thermophilic enzymes, such as AdoHyc hydrolase and MTA phosphorylase from *S. solfataricus*, are able to withstand temperatures up to 90° C, and have been investigated in a microwave setting.<sup>71</sup> One alternative method to utilizing microwave irradiation for biological molecules

is via a CoolMate apparatus (Figure 5.1).<sup>72</sup> The Coolmate allows solutions to reach temperatures of -20° C to -30° C, affording microwave irradiation to be applied, and stopped once the temperature reaches physiological temperature.<sup>72</sup> CoolMate technology has been applied to the synthesis of organic molecules, such as diphenoxybutene and phenoxydiene.<sup>73</sup>

As previously mentioned, bioconjugation offers a unique method to append certain molecular moieties, such as dyes, and other substrates to proteins.<sup>45-49</sup> Microwave irradiation could offer a quick method to facilitate bioconjugation reactions with proteins,

as often these bioconjugation reactions can take many hours.<sup>44,45</sup> Also, bioconjugation reactions such as 1,3-dipolar cycloadditions utilize metals, which degrade the biomolecules the longer the reaction occurs. Through CoolMate technology, proteins can be efficiently conjugated using microwave irradiation, without thermal degradation. Moreover, due to the accelerated rate of the reaction, less protein would be degraded by catalytic metals. Two unnatural amino acids were examined for their abilities to undergo bioconjugations with dyes via both

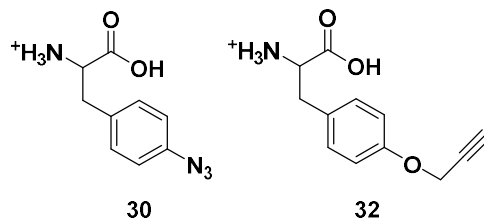


Figure 5.2: Two UAAs being tested for bioconjugation under CoolMate conditions. UAA **30** contains an azide moiety, while UAA **32** contains an alkyne moiety.

Glaser-Hay and 1,3-dipolar cycloaddition reactions under CoolMate conditions (Figure 5.2). UAA **30** and **32** both contain reactive moieties that allow them to react under physiological conditions, depending on the respective dye substrates.

## Results and Discussion:

Initial attempts at bioconjugation suffered from a lack of reproducibility. Both UAA **30** and **32** were investigated for their ability to undergo a click reaction with an alkyne or azide Alexafluor dye, respectively (Scheme 5.1). It was determined that only GFP containing UAA **32** was able to undergo click reactions. This is may be due to sterics, as

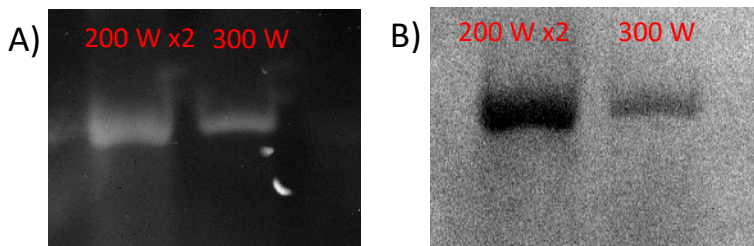
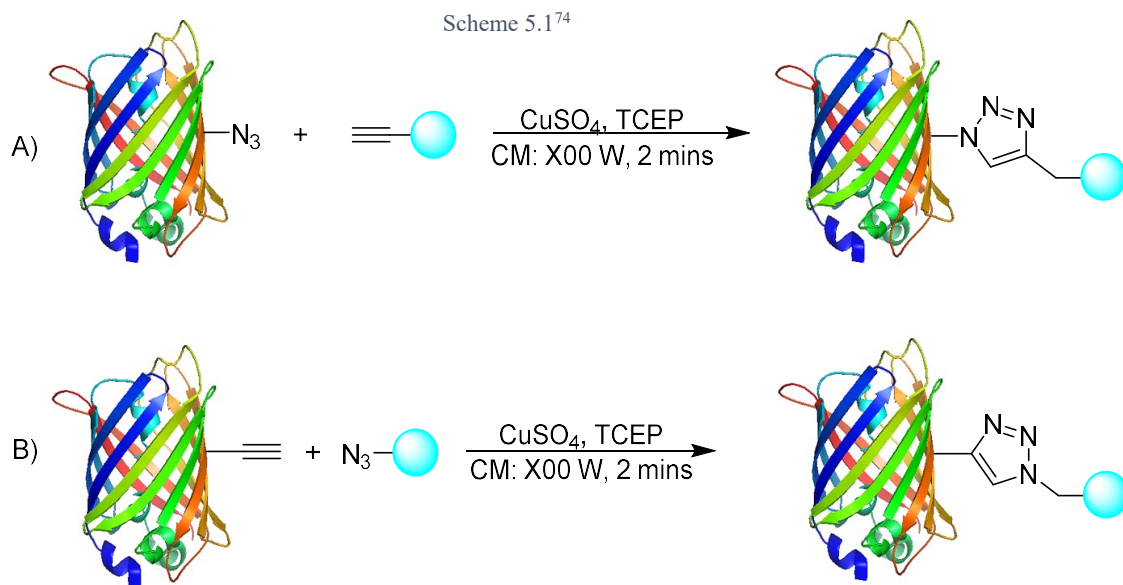


Figure 5.3: SDS-PAGE analysis of CoolMate clicks with GFP containing **32**. A) Fluorescence readings. Since there are fluorescent bands, it indicates that the protein has bioconjugated with the fluorescent dye. B) Coomassie stained gel, showing protein.

the azide group is in close proximity to the benzene ring on UAA **30**, and thus closer to the bulk of the protein. UAA **32** has a methylene extension from

the aromatic ring, which reduces some of the steric bulk. Preliminary reactions were performed using 1 mL of deionized water and 1 mL of tBuOH, but it was later discovered that reducing the concentrations to 500  $\mu$ L of deionized water and 500  $\mu$ L tBuOH resulted in higher yields of click product. This is due to an increased concentration increasing the propensity of the two reaction partners to react with appropriate orientations. The 1,3-dipolar cycloaddition reactions were examined at various power settings, and it was determined that reacting twice at 200 W and once at 300 W for 5 minutes (or when the reaction reached 40° C) yielded the best results (Figure 5.3). Unfortunately, these results have been very difficult to reproduce. This could be due to the fact that large amounts of protein are needed, and to repeat the process, additional GFP containing UAA **32** needs to be expressed and concentrated. The new expressions are often at differing concentrations, so is possible that a certain concentration of GFP is needed for the reaction to proceed.



GFP containing UAA **32** was also reacted with an alkyne dye, in order to test Glaser-Hay bioconjugation reactions under CoolMate conditions (Scheme 5.2). Similar to the click conditions, the volume of water was reduced from 2 mL to 1 mL, in order to

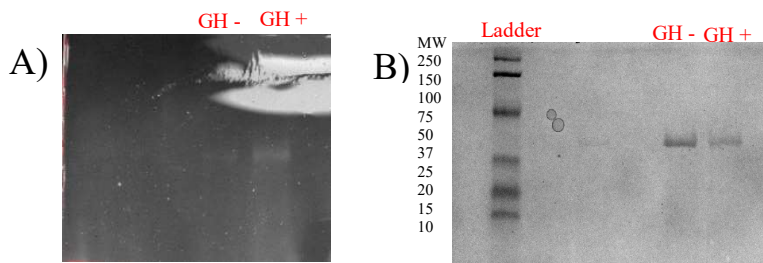
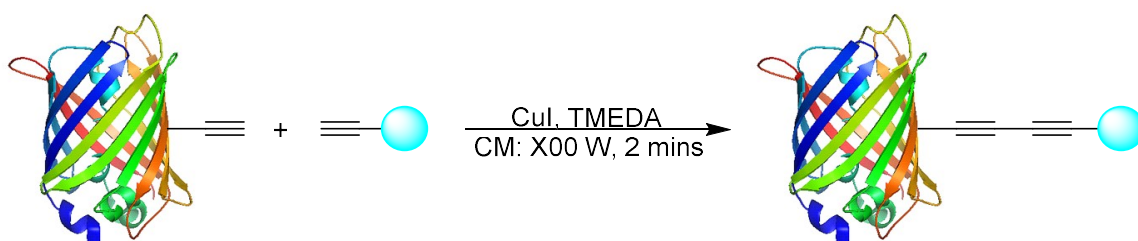


Figure 5.4: SDS-PAGE analysis of CoolMate Glaser-Hay with GFP containing **32**. A) Fluorescence readings. The GH + band indicates that the Glaser-Hay bioconjugation was successful as the fluorescent dye appears. Interestingly, the control (GH-) lane also shows a faint band, indicating that the Glaser-Hay reaction occurred without catalyst. B) Coomassie stained gel, showing protein.

increase the propensity of the two reaction partners to react with appropriate orientations. Various power settings were tested for the Glaser-Hay reaction,

and it was found that undergoing the reaction at 200 W twice for 5 minutes (or until the reaction reached 40° C), afforded bioconjugation product. Interestingly, it was also demonstrated that in the absence of copper iodide, Glaser-Hay product was still able to be formed (Figure 5.4). However, similar to the click bioconjugations, this procedure has poor reproducibility, and it is also theorized to be a result of inconsistency in protein expression.

Scheme 5.2<sup>74</sup>



Despite the difficulties with reproducibility, preliminary results indicate the success of both bioconjugation methods. This is the first known successful bioconjugations of

proteins utilizing microwave irradiation. Both bioconjugation reactions take less than five minutes to complete, which allows for a more efficient reaction. In the case of the click reactions, this means that there is less time for the protein to degrade from the copper (II) ions. In the case of the Glaser-Hay reactions, the success of the control indicates that the microwave irradiation provides enough energy to facilitate the reaction, even in the absence of copper. Both bioconjugation reactions are currently being optimized, in order to both improve reproducibility, and make the reaction more efficient. Also, the copper-free Glaser-Hay bioconjugation is also being optimized, in order to increase yields, and make the reaction viable for broader applications.

## **Experimental:**

**General:** Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Microwave reactions were conducted in a CEM Discovery microwave and CEM Discovery CoolMate microwave. NMR data was acquired on a Varian Gemini 400 MHz. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit.

**General GFP Expression:** A pET-GFP-TAG plasmid (0.5  $\mu$ L) was co-transformed with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electroporator. The cells were then plated (100  $\mu$ L) 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 Chlr) at an



OD<sub>600</sub> 0.1. The cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8, and protein expression was induced by addition of the respective unnatural amino acid (100 μL, 100 mM) and 20 % arabinose (10 μL) and 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; 10 μL). The cultures were incubated at 30° C for 16-20 h then pelleted 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 added to an equilibrated column of Ni-NTA resin (200 μL) and the GFP was purified according to manufacturer's protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h), and employed without further purification. Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), after being stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

**Coolmate Click Procedure:** GFP containing **30** or **32** at residue 151 (30 μl), AlexaFluor-488 dye (20 μl, 1 mM), deionized water (500 μL), *tris*(2-carboxyethyl)phosphine (TCEP; 32 μL, 50 mM), CuSO<sub>4</sub> (32 μl, 50mM), tert-Butyl alcohol (99%, 500 μL), and a stir bar were added to a Coolmate vial. Perfluoro polyether was added to a CEM Discover Coolmate. Liquid nitrogen or dry ice was added to the cooling chambers, to cool the Coolmate to -50° C. The flow rate was turned on to a moderate to high level, and the contents of the vial were cooled to -30° C and subjected to microwave irradiation (200 W or 300W) until the contents of the vial reached a temperature of 40° C. In order to remove

excess copper and fluorophore, a buffer exchange was performed with Spin-X UF concentrator columns, and with PBS (8x200  $\mu$ L) to a final volume of 50  $\mu$ L. A control reaction was performed utilizing the same procedure, in absence of copper sulfate. The GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h). Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), first via a Spyro Ruby irradiation to analyze fluorescence. Then, the gel was stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

**Coolmate Glaser-Hay Procedure:** GFP containing a UAA **32** at residue 151 (30  $\mu$ l), AlexaFluor-488 alkyne (20  $\mu$ l, 1 mM), CuI (5  $\mu$ l, 500 mM), tetramethylethylenediamine (TMEDA; 5  $\mu$ l, 500 mM), deionized water (1 mL), and a stir bar were added to a CoolMate vial. Perfluoro polyether was added to a CEM Discover Coolmate. Liquid nitrogen or dry ice was added to the cooling chambers, to cool the Coolmate to -50° C. The flow rate was turned on to a moderate to high level, and the contents of the vial cooled to -30° C, and were subjected to microwave irradiation (200 W or 300 W) until the contents of the vial reached a temperature of 40° C. In order to remove excess copper and fluorophore, a buffer exchange was performed with Spin-X UF concentrator columns, and with PBS (8x200  $\mu$ L) to a final volume of 50  $\mu$ L. A control reaction was performed utilizing the same procedure, in the absence of copper iodide. The GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h). Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), first via a Spyro Ruby irradiation to analyze fluorescence. Then, the gel was stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain

solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

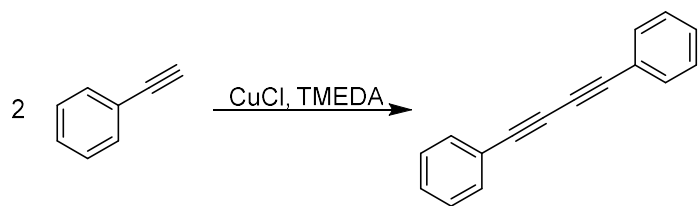
## Chapter 6: Developing a Solid-Support Resin Method for Asymmetric Polyynes Synthesis

Polyynes are molecules containing an acetylenic scaffold, and are core structures found in various natural products.<sup>75-77</sup> Over one thousand polyynes have been isolated from various organisms, and these compounds have been shown to exhibit therapeutic activity, such as antibacterial, anticancer, and antifungal properties.<sup>76-78</sup> Due to the highly conjugated  $\pi$ -system, the acetylenic scaffold can also act as a molecular wire, with optical

and conductive properties.<sup>79</sup> The polyynic moiety was first developed by Glaser, who utilized oxidative dimerization of phenylacetylene via a copper (I) catalyst.<sup>80</sup> Hay added catalytic amounts of copper (I) chloride with tetramethylenediamine (TMEDA) and oxygen.<sup>81</sup> This allowed the reaction to be run at more neutral conditions, decreasing the temperature needed to run the reaction and allowing a greater variety of solvents to be used.<sup>81</sup> The variant reaction, known as the Glaser-Hay reaction, forms symmetric polyynic molecules with relatively high yields (Scheme 6.1).<sup>81</sup>

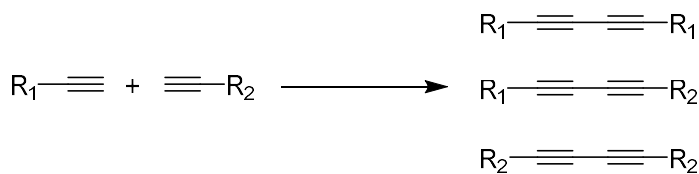
One caveat associated with the Glaser-Hay reaction is it does not exhibit chemoselectivity when two different alkynes are reacted, and instead will form three different diyne products (Scheme 6.2). The Young lab has previously employed polystyrene resin as a solid support to address

Scheme 6.4

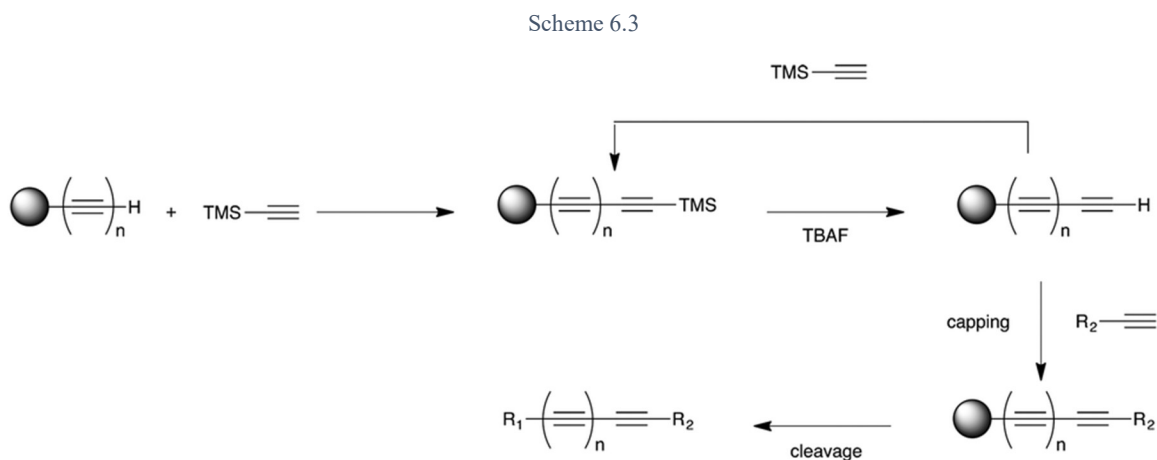


and conductive properties.<sup>79</sup> The polyynic moiety was first developed by Glaser, who utilized oxidative dimerization of

Scheme 6.2

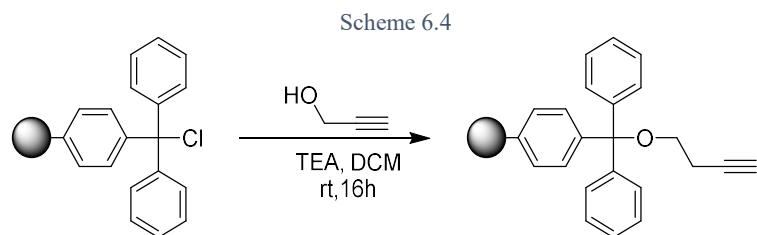


the chemoselectivity issues of the Glaser-Hay reaction, as well as eliminate much of the post-reaction workup.<sup>82</sup> Solid-support synthesis allows easy removal of excess reagents, as well as structural rigidity (Scheme 6.3).<sup>82</sup> Through this method, a library of asymmetric diynes can be synthesized. Also, different polyynes lengths can be achieved by coupling an immobilized alkyne with trimethylsilylacetylene (TMS acetylene), followed by cleavage of the protecting group. The synthesized diyne can undergo another round of polyyn extension, or capping with a different alkyne molecule. This allows the production of an asymmetric polyyn library.



## Results and Discussion:

My research focused on the synthesis of 8 polyynes (Figure 6.1), but many other polyynes were synthesized by other members of the lab. For each of the eight polyynes, propargyl alcohol was first immobilized. This was achieved by first swelling a 1% cross-



linked tritylchloride derivatized polystyrene resin in dichloromethane, followed by addition of propargyl

alcohol (Scheme 6.4). Based off previous results, it was found that alkyne loadings above 0.7 mmol g<sup>-1</sup> resulted in undesired dimerization of the immobilized alkynes.<sup>82</sup> Thus to ensure asymmetrical polyynes synthesis, propargyl alcohol was immobilized at loadings of 0.4-0.7 mmol g<sup>-1</sup>. For compounds, **33** and **37**, the next step involved a Glaser-Hay reaction with ethynylanisole and hexyne capping groups, respectively. The reaction was optimized with CuI and TMEDA in THF for 16 hours at 60° C, and then washed with cycles of DCM-MeOH (5 X 5 mL each). Alkyne extension followed a similar procedure, except that the Glaser-Hay reaction was done with TMS acetylene. The TMS group was then cleaved with TBAF in dichloromethane for 1 hour at room temperature. This step was repeated up to two more times, depending on the length of the polyynyl chain, where afterwards, the capping alkyne was then installed. Utilizing ethynylanisole and hexyne capping groups, eight asymmetric polyynes were synthesized (Figure 6.1).

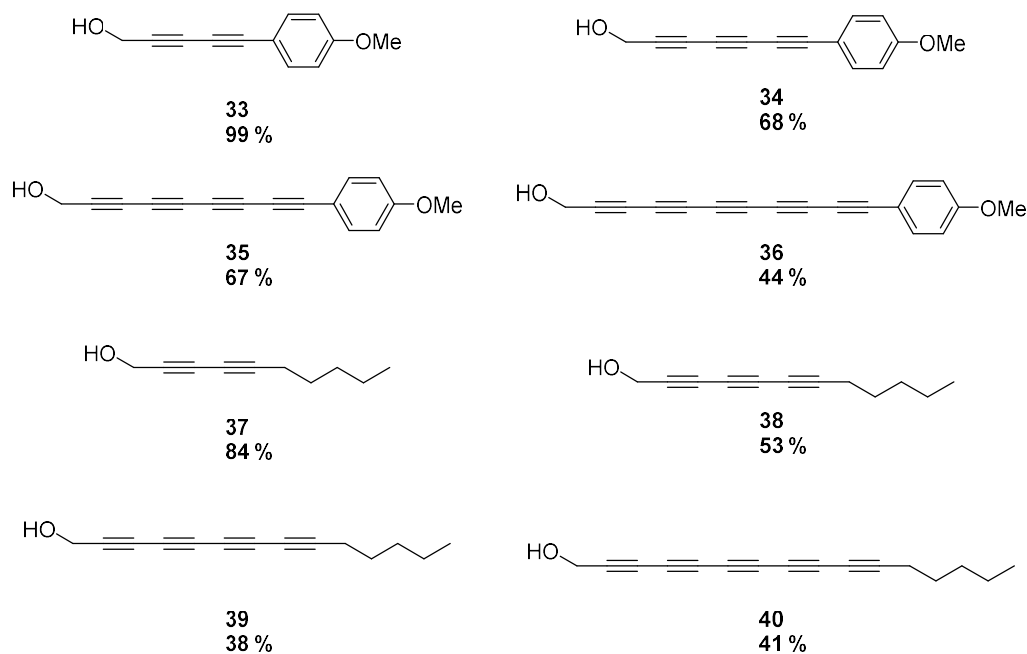


Figure 6.1: Eight asymmetric polyynes synthesized using ethynylanisole and hexyne capping groups

The yields of the asymmetric polyynes generally decreased with each alkyne extension, most likely due to the number of synthetic steps (up to 8 for the pentyne); however, yields were generally high (99-38 %). The hexyne capping group also resulted in lower yields than the ethynylanisole group. This could be a result of the increased conjugation of the anisole moiety creating a more

reactive alkyne. In addition to an immobilized propargyl alcohol, an ethynyl benzene group was

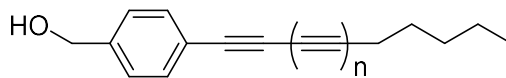


Figure 6.2: A proposed asymmetric polyyne with an ethynyl benzene alcohol and a hexyne capping group.

immobilized and capped with several alkynes. An ethynyl alcohol resin with a hexyne capping group was attempted (Figure 6.2); however, only the diyne was able to be synthesized. The hexyne capping group also showed lower than average yields for the immobilized propargyl alcohol.

Fluorescence measurements were obtained for various compounds, as these highly conjugated systems were hypothesized to exhibit unique fluorescent properties. For the ethynylanisole group, extended conjugation resulted in minor red and blue shifts, but the formation of the pentyne resulted in a large red shift (Figure 6.3). Other asymmetrical pentyne prepared by other lab members exhibited large blue shifts. This could be a result of the pentyne compounds transitioning into a polymeric state, exhibiting complex packing and interactions of the  $\pi$ -systems. The fluorescence measurements of the hexyne capping group differed greatly, as the triyne displayed a large red shift, while the other polyynes had smaller red and blue shifts (Figure 6.4). Despite the variability of the fluorescence, the results indicate that the optical properties of asymmetric polyynes can be tuned, depending on the capping group, and number of alkyne units. Fluorescence was also measured for a series of triynes with varying capping groups (Figure 6.5). Triynes with groups that

extended the conjugation, such as phenyl and anisole produced blue shifts, with a small blue shift observed for the propargyl alcohol capping group. Current research is being done to further investigate and characterize the optical and conducting properties of the prepared compounds. Also, the technology can be extended to the preparation of the more complex and biologically active natural product derivatives.

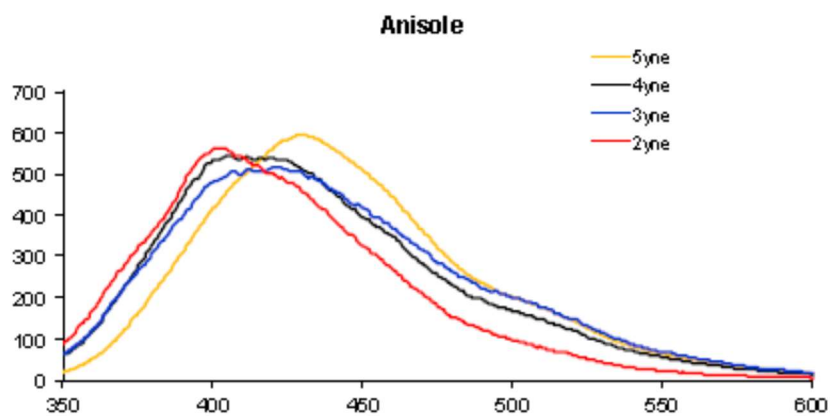


Figure 6.3: Emission spectra for the anisole polyynes after excitation at 330 nm.<sup>83</sup>

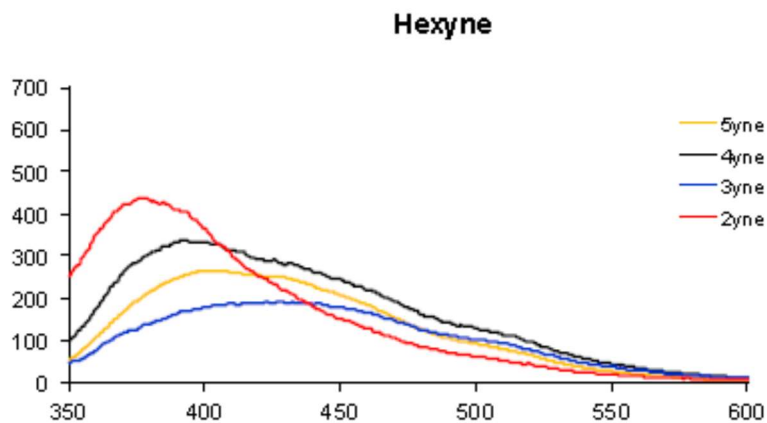


Figure 6.4: Emission spectra for the anisole polyynes after excitation at 330 nm.<sup>83</sup>



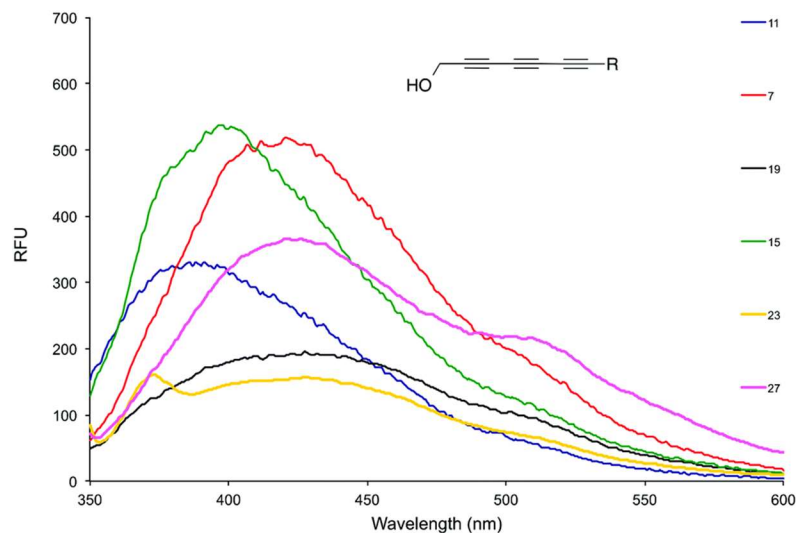


Figure 6.5: Emission spectra of propargyl alcohol loaded triynes, with different capping groups.<sup>83</sup> The capping groups are as follows: 7 – C<sub>2</sub>H<sub>2</sub>NH<sub>2</sub>, 11 – C<sub>6</sub>H<sub>5</sub>, 15 – CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>, 19 - CH<sub>2</sub>OH, 23 – C<sub>3</sub>H<sub>4</sub>N, 27 – C<sub>6</sub>H<sub>4</sub>OCH<sub>3</sub>. Each molecule was excited at 330 nm.<sup>83</sup>

## Experimental:

**General:** Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Tritylchloride resin, 100- 200 mesh, 1% DVB crosslinking, was purchased from Advanced Chemtech. Reactions were conducted under ambient atmosphere with non-distilled solvents. NMR data was acquired on a Varian Gemini 400 MHz. GC/MS analysis was conducted on an Agilent Technologies 6890N GC system interfaced with a 5973N mass selective detector. An Agilent J&W GC capillary column (30 m length, 0.32 mm diameter, 0.25 μm film) was employed with a splitless injection (250° C inlet, 8.8 psi) with an initial 70° C hold (2 min) and ramped for 15 min to 230° C. Fluorescence data was measured using a PerkinElmer LS 55 Luminescence Spectrometer.

**Immobilization of Alcohol onto Trityl Chloride Resin in Low Loading Conditions:** To flame dried vial was added trityl chloride resin (200 mg, 0.36 mmol, 1 equiv.) and

dichloromethane (5 mL). The resin was swelled at room temperature with gentle stirring for 15 min. Alcohol (25.0  $\mu$ L,  $\sim$ 1.2 equiv.) was added to reaction, followed by triethylamine (10.0  $\mu$ L, 0.072 mmol, 0.2 equiv). 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 DCM and dried under vacuum for 45 min before further use. Resin loading was assessed via cleavage of 15 mg resin in 2% TFA in DCM (200  $\mu$ L) at room temperature for 1 hour. After filtration of resin, loading was determined both by mass of cleaved product and by GC/MS analysis of integrated alcohol peaks from a standard curve and the cleaved resin.

**Polyne Extension Protocol:** Trimethylsilylacetylene (160  $\mu$ L, 1.05 mmol, 15 equiv.) was added to a flame dried vial containing the alcohol derivatized trityl resin (100 mg, 0.07 mmol, 1 equiv.), and tetrahydrofuran (2.0 mL). The copper catalyst (20 mg, 1.06 mmol.) and tetramethylethylenediamine (20  $\mu$ L, 0.132 mmol.) were added to a separate flame-dried vial then dissolved in tetrahydrofuran (2.0 mL). The catalyst mixture was then added to the resin in one portion and stirred at 60° C for 16 h. The resin was transferred to a syringe filter and washed with DCM and MeOH (5 alternating rinses with 5 mL each).

**Regeneration of Terminal Alkyne:** The TMS group was then cleaved by incubation in 1M tetra-n-butylammonium fluoride trihydrate in DCM (TBAF, 1 mL, 1 h). Then the reaction was again transferred to a syringe filter and washed with DCM and MeOH (5 alternating rinses with 5 mL each) and dried under vacuum for 45 minutes. Product was weighed and transferred to flame dried vial for future use.

**Polyyne Capping:** Soluble alkyne (15 equiv.) was added to a flame dried vial containing the desired alcohol derivatized trityl resin and tetrahydrofuran (2 mL). The copper catalyst (20 mg, 1.06 mmol.) and tetramethylethylenediamine (20  $\mu$ L, 0.132 mmol.) were added to a separate flame-dried vial then dissolved in tetrahydrofuran (2 mL). The catalyst mixture was then added to the resin in one portion and stirred at 60° C 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 product was then cleaved from the resin by treatment with 2% TFA (DCM, 1 h) and filtered into a vial. If necessary, a short silica plug was utilized to remove unreacted starting material (1:1 EtOAc/Hex).

**5-(4-methoxyphenyl)penta-2,4-diyn-1-ol, 33:** The solvent was removed in vacuo to give **33** as a solid (13 mg, 0.070 mmol, 99 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.44 (d, J = 7.6 Hz, 2H), 6.85 (d, J = 7.6 Hz, 2H), 4.14 (s, 2H), 3.82 (s, 3H); GCMS (Rt=9.93 min) calculated for C<sub>12</sub>H<sub>10</sub>O<sub>2</sub> 186.1, found 186.2.

**7-(4-methoxyphenyl)hepta-2,4,6-triyn-1-ol, 34:** The solvent was removed in vacuo to give **34** as a solid (10 mg, 0.049 mmol, 68 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.44 (d, J = 7.6 Hz, 2H), 6.86 (d, J = 7.6 Hz, 2H), 4.38 (s, 2H), 3.82 (s, 3H); GCMS (Rt t=11.02 min) calculated for C<sub>14</sub>H<sub>10</sub>O<sub>2</sub> 210.2, found 210.3.

**9-(4-methoxyphenyl)nona-2,4,6,8-tetrayn-1-ol, 35:** The solvent was removed in vacuo to give **35** as a solid (11 mg, 0.037 mmol, 67 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>):  $\delta$  7.46 (d, J = 7.6 Hz, 2H), 6.84 (d, J = 7.6 Hz, 2H), 4.38 (s, 2H), 3.82 (s, 3H); GCMS (Rt = 18.39 min) calculated for C<sub>16</sub>H<sub>10</sub>O<sub>2</sub> 234.3, found 234.2.

**11-(4-methoxyphenyl)undeca-2,4,6,8,10-pentayn-1-ol, 36:** The solvent was removed in vacuo to give **36** as a solid (8 mg, 0.031 mmol, 44 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 7.42 (d, J = 7.6 Hz, 2H), 6.85 (d, J = 7.6 Hz, 2H), 4.42 (s, 2H), 3.82 (s, 3H); GCMS (R<sub>t</sub> = 14.35 min) calculated for C<sub>18</sub>H<sub>10</sub>O<sub>2</sub> 258.3, found 258.2.

**deca-2,4-diyn-1-ol, 37:** The solvent was removed in vacuo to give **37** as a solid (8 mg, 0.058 mmol, 84 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 4.32 (s, 2H), 2.29 (t, J = 6.8 Hz, 3H), 1.52 (m, 2H), 1.42 (sextet, J = 7.2 Hz, 2H), 0.91 (t, J = 7.23 Hz, 3H); GCMS (R<sub>t</sub> = 7.29 min) calculated for C<sub>9</sub>H<sub>12</sub>O 136.1, found 136.0.

**dodeca-2,4,6-triyn-1-ol, 38:** The solvent was removed in vacuo to give **38** as a solid (6 mg, 0.037 mmol, 53 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 4.32 (s, 2H), 2.30 (t, J = 6.8 Hz, 3H), 1.55 (m, 2H), 1.42 (sextet, J = 7.2 Hz, 2H), 0.93 (t, J = 7.23 Hz, 3H); GCMS (R<sub>t</sub> = 10.61 min) calculated for C<sub>11</sub>H<sub>12</sub>O 160.2, found 160.2. 4

**tetradeca-2,4,6,8-tetrayn-1-ol, 39:** The solvent was removed in vacuo to give **39** as a solid (5 mg, 0.027 mmol, 38 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 4.36 (s, 2H), 2.30 (t, J = 6.8 Hz, 3H), 1.55 (m, 2H), 1.42 (sextet, J = 7.2 Hz, 2H), 1.00 (t, J = 7.23 Hz, 3H); GCMS (R<sub>t</sub> = 10.28 min) calculated for C<sub>13</sub>H<sub>12</sub>O 184.2, found 184.2.

**hexadeca-2,4,6,8,10-pentayn-1-ol, 40:** The solvent was removed in vacuo to give **40** as a solid (9 mg, 0.028 mmol, 41 %) <sup>1</sup>H NMR (400 MHz; CDCl<sub>3</sub>): δ 4.38 (s, 2H), 2.35 (t, J = 6.8 Hz, 3H), 1.60 (m, 2H), 1.42 (sextet, J = 7.2 Hz, 2H), 1.02 (t, J = 7.23 Hz, 3H); GCMS (R<sub>t</sub> = 11.71 min) calculated for C<sub>15</sub>H<sub>12</sub>O 208.3, found 208.3.

**Fluorescence Measurements:** Each fluorescence spectrum was obtained in a glass cuvette with DCM (3 mL). To the cuvette solvent was added the asymmetric polyynes (500 μL, 10

mmol). The sample was excited at 330 nm and the emission spectral was recorded from 350 nm to 600 nm; excitation slit width of 10 nm, an emission slit width of 10 nm, and at a scan speed of 400 nm/min.

## Chapter 7: A Photosensitive Selenocysteine UAA

There are certain amino acids that naturally exist outside of the canonical twenty that can also be utilized to make proteins. One such example is selenocysteine, which is a cysteine homolog with the sulfur atom replaced with a selenium atom. The other “unnatural” amino acid is pyrrolysine, which contains a pyrroline ring on the end of a lysine chain.<sup>84,85</sup> Amino acids resembling selenocysteine and pyrrolysine are known as proteinogenic amino acids, and they serve important purposes for cellular function.<sup>84</sup> Selenocysteine, for example, has a low pKa and low reduction potential due to the selenium atom, which makes it useful for antioxidant activity.<sup>84</sup> Pyrrolysine is used in methanogenic enzymes mainly found in archaea.<sup>5</sup> These proteinogenic amino acids are incorporated into proteins co-translationally, and are encoded in mRNA in the form of stop codons. Pyrrolysine typically is encoded by the UAG (amber) stop codon, while selenocysteine uses the UGA (umber) stop codon.<sup>84,85</sup>

Researchers have recently been interested in utilizing selenocysteine, especially due to its antioxidant activity.<sup>85</sup> Several selenoproteins have been discovered in humans, including glutathione peroxidase and thioredoxin reductase.<sup>85</sup> Naturally, selenocysteine is incorporated into proteins using the UGA stop codon, or a selenocysteine insertion sequence (SECIS), which is a unique untranslated portion of mRNA.<sup>85,86</sup> Traditionally, the high reactivity of selenium atoms has made genetically encoding selenocysteine into a protein difficult. This process is essential if there is a need to specify the site of the selenocysteine residue. To counteract this, a caging protecting group can be added to

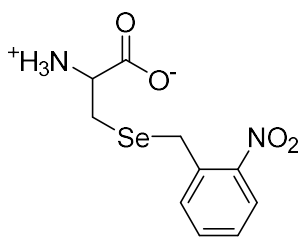


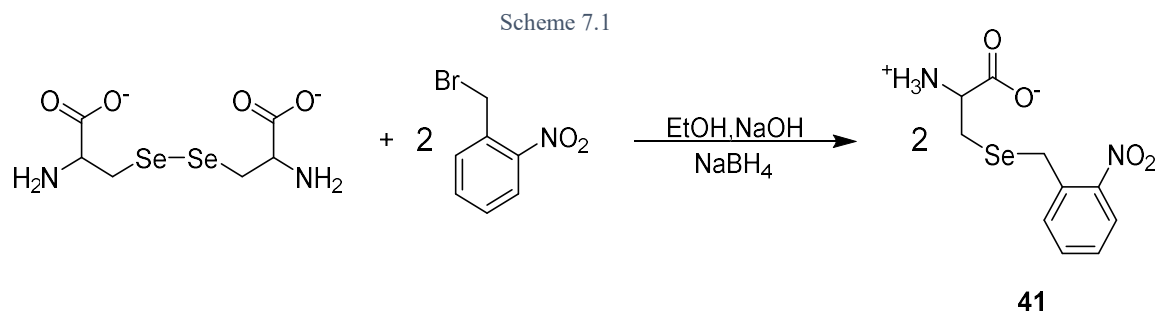
Figure 7.1: Proposed caged selenocysteine unnatural amino acid (**41**).

selenocysteine to make it less reactive, and able to be employed in aaRS selections. In one example, a 4,5-dimethoxy-2-nitrobenzyl (DMNB) caging group was attached to the selenium atom of selenocysteine, and incorporated with an enhanced green fluorescent protein (EGFP).<sup>87</sup> By using an orthogonal aaRS/tRNA pair, the caged selenocysteine was successfully introduced into the protein.<sup>87</sup> We propose an ortho-nitrobenzyl protecting group, to be incorporated into GFP (Figure 7.1).

The incorporation of a caged selenocysteine into proteins offers a probe to understand the importance of this proteogenic amino acid in biological redox reactions. Also, a link has been established between selenocysteine-containing proteins and neurodegenerative disorders such as Alzheimer's.<sup>88</sup> A caged selenocysteine unnatural amino acid can elucidate the role of selenocysteine in neurodegenerative diseases, as well as further understand the underlying mechanism.

## Results and Discussion:

The synthesis of **41** proceeded in a multi-pot reaction (Scheme 7.1). The addition of the caging group necessitated the reaction to be prevented from light exposure, and was thus wrapped in aluminum foil upon reaction with the ortho-nitrobenzyl bromide reagent. Due to the charged amino acid backbone of the product, the reaction could not be extracted, but flash chromatography was attempted as a purification method. Compound **41** dissolved



in a 10:1 DCM:methanol eluent mixture, and was transferred to the column. A gradient elution was then attempted, from 10:1 DCM:methanol to 1:1 DCM:Methanol, resulting in product recovery. Most likely, a high concentration of methanol is needed in the mobile phase in order to advance the product through the polar silica gel. This resulted in yields of over 100 %, which is most likely due to silica gel dissolution by methanol. The product was then analyzed via  $^1\text{H}$  NMR with  $\text{D}_2\text{O}$  as a solvent, and **41** was observed in a 55.5 % yield.

With **41** synthetically prepared, incorporation of the UAA into GFP was then attempted. Since no aaRS has been evolved to explicitly incorporate **41**, several aaRS assays were performed to examine aaRS promiscuity for **41**. The pEVOL synthetases pAcF 32/67, pCNF 32, ONBY, BipY, pCNF, and AzoF were investigated due to their large binding pockets, or structural similarities of their corresponding UAAs. Each aaRS plasmid was transformed with a pET-GFP-TAG-151 plasmid into *E. coli*. Cultures were grown, and GFP expression was induced in a 96-well plate with appropriate expression controls, as described in the general aaRS assessment protocol. Fluorescence of the cell cultures were measured on a microplate reader after 16 hours of expression (Figure 7.2). While most aaRSs were not able to accommodate the UAA, the results indicate that the pCNF-aaRS was able to incorporate **41**, when compared to a negative control grown in the absence of the UAA. The synthetase pCNF 32 was also able to incorporate an amino acid at a



relatively high rate. However, since the control incorporated an amino acid more than in the presence of **41**, it is more likely that tyrosine was being incorporated, and not **41**.

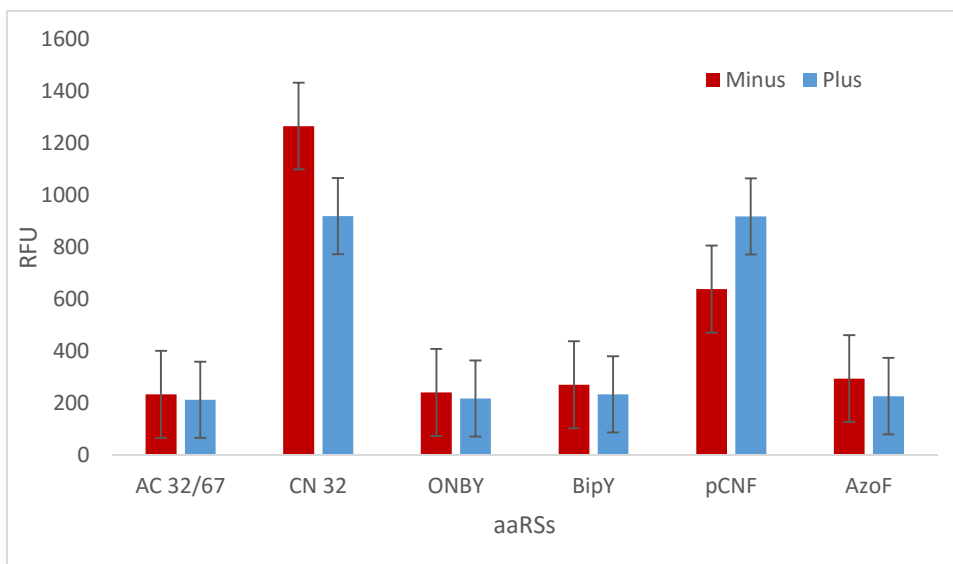


Figure 7.2: Synthetase assay for **41**. Six synthetases were tested for their ability to incorporate **41** into GFP at residue 151. The synthetase pCNF showed the most promise, as it incorporated an amino acid at a higher rate than the control.

In order to elucidate if these results were truly valid, large scale GFP expression (30 mL) was attempted utilizing the pCNF synthetase. However, little protein was expressed and isolated. This could be due to the pCNF plasmid incorporating **5** poorly, but enough to indicate a higher result than the control. Therefore, a second synthetase screen was performed (Figure 7.3). The synthetases ONiF, 3FY, CuoA, and Nap were selected based on their large binding pockets, or similarities of their respective unnatural amino acid, and were analyzed in the same manner as described above. Unfortunately, no aaRS incorporated **41** better than a control, suggesting the need for an aaRS evolution experiment.

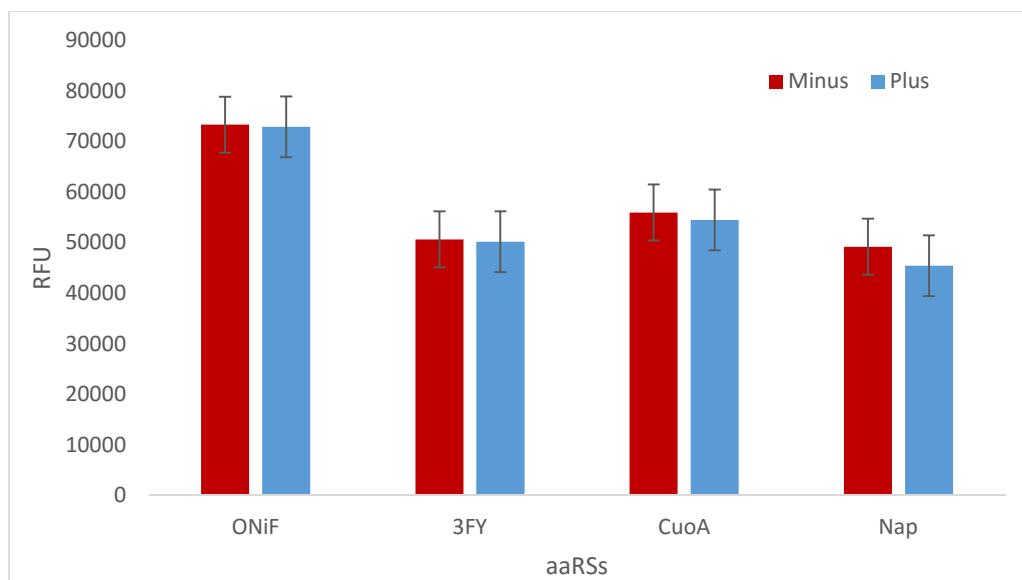


Figure 7.3: A second synthetase assay for **41**. Four synthetases were tested for their ability to incorporate **41** into GFP at residue 151. Unfortunately, no synthetase incorporated **41** in a greater amount than the control.

Several aaRSs have been employed, but none afforded a successful incorporation of **41**. Other synthetases are currently being investigated; however, **41** may be difficult to incorporate due to the benzene ring being farther away from the amino acid backbone. This is problematic due to the tested aaRSs all incorporating respective UAAs that have been derived from tyrosine, which has a singular methylene unit separating the benzene from the backbone. Thus, an extended aromatic moiety may not properly fit into the binding sites. Therefore, an aaRS double sieve selection is being performed from a large library of aaRSs ( $\sim 10^8$ ), in order to find a synthetase that exclusively incorporates **41**. Once a synthetase has been evolved or discovered that can incorporate **41**, we are going to collaborate with Dr. Landino, in order to utilize the UAA to probe biological redox reactions. Specifically, **41** will probe oxidative damage as it relates to Parkinson's and Alzheimer's.

## Experimental:

**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. NMR data was acquired on a Varian Gemini 400 MHz. Fluorescence data was measured using a Synergy Microplate reader. All GFP proteins were purified according to manufacturer's protocols using a Qiagen Ni-NTA Quik Spin Kit.

**General GFP Expression:** A pET-GFP-TAG plasmid (0.5  $\mu$ L) was co-transformed with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electorporator. The cells were then plated (100  $\mu$ L) 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 Chlr) at an OD<sub>600</sub> 0.1. The cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8, and protein expression was induced by addition of the respective unnatural amino acid (100  $\mu$ L, 100 mM) and 20 % arabinose (10  $\mu$ L) and 0.8 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG; 10  $\mu$ L). The cultures were incubated at 30° C for 16-20 h then pelleted at 5,000 rpm for 10 minutes and stored at -80° C for 3 hours. The cell pellet was re-suspended using 500  $\mu$ L 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 added to an equilibrated column of Ni-NTA resin (200  $\mu$ L) and the GFP was purified according to manufacturer's

protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10 % precast gels, 150 V, 1.5 h), and employed without further purification. Gels were imaged on a BioRad Molecular Imager (Gel Doc XR+), after being stained with Coomassie Brilliant Blue for an hour. The gel was then destained utilizing destain solution (60 % deionized water, 30 % methanol, 10 % acetic acid) until the gel was clear, after which the gel was analyzed via the Coomassie protocol on the BioRad Molecular Imager.

**General aaRS Assessment:** The aaRSs were assessed for unnatural amino acid incorporation by first co-transforming a pET-GFP-TAG variant plasmid (0.5  $\mu$ L) with a pEVOL-tRNA synthetase plasmid (0.5  $\mu$ L), in a similar method as described previously. For each pEVOL plasmid, one colony was used to inoculate LB media (4 mL) in separate containers that contained ampicillin (4  $\mu$ L, 50 mg/mL) and chloramphenicol (4  $\mu$ L, 34 mg/mL). The cultures were incubated at 37° C overnight. Expression cultures were created containing LB media (4 mL), ampicillin (4  $\mu$ L, 50 mg/mL), and chloramphenicol (4  $\mu$ L, 34 mg/mL), and then were inoculated with the previous culture (250  $\mu$ L). The expression cultures were incubated at 37° C to an OD<sub>600</sub> between 0.6 and 0.8. Protein expression was induced by addition of 20 % arabinose (4  $\mu$ L), and 0.8 mM IPTG (4  $\mu$ L). The cultures were then plated on a 96-well plate (200  $\mu$ L), with six wells for each plasmid. Three wells were designated as controls, and the other three were designated for the experiment. The designated UAA (100  $\mu$ L, 100 mM) was added to the experimental wells only. The plate was allowed to shake at 30° C for 16 h, after which it was measured for fluorescence (528 nm and 485 nm) utilizing a Synergy HT microplate reader.

**2-ammonio-3-((2-nitrobenzyl)selenanyl)propanoate, 41:** To a flamed dried vial was added seleno-L-cystine (22 mg,  $5.98 \times 10^{-5}$  mol, 1 equivalent), sodium hydroxide (320  $\mu$ L, 0.5 M),

and ethanol (80  $\mu\text{L}$ ). The reaction was allowed to stir at  $0^\circ\text{C}$ . Once dissolved, sodium borohydride (16 mg,  $4.23 \times 10^{-4}$  mol, 7 equivalents) was added, and the reaction was warmed to room temperature, which was maintained for 5 minutes. The reaction was then cooled to  $0^\circ\text{C}$ , and sodium hydroxide (160  $\mu\text{L}$ , 2M) was added, and allowed to stir for 5 minutes, while the reaction warmed to room temperature. Ortho-nitrobenzyl bromide (52 mg,  $2.39 \times 10^{-4}$  mol, 4 equivalents) was then added, and the reaction was allowed to stir at room temperature for an additional 1.5 hours. Hydrochloric acid (500  $\mu\text{L}$ , 1M) was added to neutralize the solution. The reaction was then concentrated via a rotary evaporator, and was purified via flash chromatography with gradient elution from 10:1 DCM/Methanol to 1:1 DCM/Methanol, resulting in a yellow solid (20 mg,  $6.60 \times 10^{-5}$  mol, 55.5 %).

## **Acknowledgments**

I would first like to thank Dr. Douglas Young for his insightful guidance and mentorship. I would also like to thank the Young Lab for their assistance and support, they have truly made lab a wonderful experience. In particular, Chris Farley '14 who taught me the basics of lab work, and Lindsay Chatkewitz, who has personally assisted me in many of my projects, and has been instrumental in achieving much of the results.

Also, the Roy R. Charles Center has supported me financially during the summers of 2014 and 2015 through the Charles Center Summer Research Stipend and the Llanso-Sherman Scholarship. I am also grateful to the various donors who contributed to my Honor's Fellowship fund.

## References

1. LODISH. *Molecular cell biology*. 7th ed. Mac Higher; 2013.
2. McKee T, McKee JR. *Biochemistry: The molecular basis of life*. Fourth Edition ed. New York: Oxford University Press; 2009.
3. Translation (biology) [Internet]: New World Encyclopedia; c2009 [cited 2009 March 19]. Available from:  
[http://www.newworldencyclopedia.org/entry/Translation\\_\(biology\)](http://www.newworldencyclopedia.org/entry/Translation_(biology)).
4. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry* **79**, 413–444 (2010).
5. Wang, L. & Schultz, P. G. Expanding the Genetic Code. *Angewandte Chemie International Edition* **44**, 34–66 (2005).
6. Johnson, J. A., Lu, Y. Y., Van Deventer, J. A. & Tirrell, D. A. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Current Opinion in Chemical Biology* **14**, 774–780 (2010).
7. Link, A. J., Mock, M. L. & Tirrell, D. A. Non-canonical amino acids in protein engineering. *Current Opinion in Biotechnology* **14**, 603–609 (2003).
8. Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* **4**, 645–652 (2000).
9. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *PNAS* **99**, 11020–11024 (2002).
10. Liu, C. C. *et al.* Protein evolution with an expanded genetic code. *PNAS* **105**, 17688–17693 (2008).

11. Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* **4**, 645–652 (2000).
12. Yoder, N. C. & Kumar, K. Fluorinated amino acids in protein design and engineering. *Chem. Soc. Rev.* **31**, 335–341 (2002).
13. Liu, D. R., Magliery, T. J., Pastrnak, M. & Schultz, P. G. Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10092–10097 (1997).
14. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* **395**, 361–374 (2010).
15. Young, D. D. *et al.* An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity. *Biochemistry* **50**, 1894–1900 (2011).
16. Tsien, R. Y. The Green Fluorescent Protein. *Annual Review of Biochemistry* **67**, 509–544 (1998).
17. IPTG: *BL21(DE3) competent cells, BL21 (DE3)pLysS competent cells, and BL21 competent cells instruction manual* (revision #066004e ed.). La Jolla, CA: Stratagene. (2006)
18. Watson, James (2003). *molecular biology of the gene*. p. 503.
19. Neu, H. C. Effect of beta-lactamase location in *Escherichia coli* on penicillin synergy. *Appl Microbiol* **17**, 783–786 (1969).
20. Riggsbee, C. W. & Deiters, A. Recent advances in the photochemical control of protein function. *Trends Biotechnol* **28**, 468–475 (2010).



21. Momotake, A., Lindegger, N., Niggli, E., Barsotti, R. J. & Ellis-Davies, G. C. R. The nitrodibenzofuran chromophore: a new caging group for ultra-efficient photolysis in living cells. *Nat. Methods* **3**, 35–40 (2006).
22. Zhu, Y., Pavlos, C. M., Toscano, J. P. & Dore, T. M. 8-Bromo-7-hydroxyquinoline as a Photoremovable Protecting Group for Physiological Use: Mechanism and Scope. *J. Am. Chem. Soc.* **128**, 4267–4276 (2006).
23. Young, D. D. & Deiters, A. Photochemical Activation of Protein Expression in Bacterial Cells. *Angewandte Chemie International Edition* **46**, 4290–4292 (2007).
24. Chou, C., Young, D. D. & Deiters, A. Photocaged t7 RNA polymerase for the light activation of transcription and gene function in pro- and eukaryotic cells. *Chembiochem* **11**, 972–977 (2010).
25. Gautier, A. *et al.* Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. *J. Am. Chem. Soc.* **132**, 4086–4088 (2010).
26. Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M. & Schultz, P. G. Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat Chem Biol* **3**, 769–772 (2007).
27. Shao, D., Creasy, C. L. & Bergman, L. W. A cysteine residue in helixII of the bHLH domain is essential for homodimerization of the yeast transcription factor Pho4p. *Nucl. Acids Res.* **26**, 710–714 (1998).
28. Kappe, C. O. Microwave dielectric heating in synthetic organic chemistry. *Chem. Soc. Rev.* **37**, 1127–1139 (2008).
29. Lidström, P., Tierney, J., Wathey, B. & Westman, J. Microwave assisted organic synthesis—a review. *Tetrahedron* **57**, 9225–9283 (2001).

30. Saxena, V. K. & Chandr, U. in *Microwave Heating* (ed. Chandra, U.) (InTech, 2011).
31. Sharma, A. K., Gowdahalli, K., Krzeminski, J. & Amin, S. Microwave-Assisted Suzuki Cross-Coupling Reaction, a Key Step in the Synthesis of Polycyclic Aromatic Hydrocarbons and Their Metabolites. *J. Org. Chem.* **72**, 8987–8989 (2007).
32. Mason, J. D. & Murphree, S. S. Microwave-Assisted Aqueous Krapcho Decarboxylation. *Synlett* **24**, 1391–1394 (2013).
33. Mirafzal, G. A. & Summer, J. M. Microwave Irradiation Reactions: Synthesis of Analgesic Drugs. *J. Chem. Educ.* **77**, 356 (2000).
34. Rijkers, D. T. S. *et al.* Efficient microwave-assisted synthesis of multivalent dendrimeric peptides using cycloaddition reaction (click) chemistry. *Chem. Commun.* 4581–4583 (2005).
35. Zhou, X. X. & Lin, M. Z. Photoswitchable fluorescent proteins: ten years of colorful chemistry and exciting applications. *Current Opinion in Chemical Biology* **17**, 682–690 (2013).
36. Taghizadeh, R. R. & Sherley, J. L. CFP and YFP, but Not GFP, Provide Stable Fluorescent Marking of Rat Hepatic Adult Stem Cells. *J Biomed Biotechnol* 2008, 453590 (2008).
37. Fortin, D. L. *et al.* Photochemical control of endogenous ion channels and cellular excitability. *Nat Meth* **5**, 331–338 (2008).
38. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *PNAS* **105**, 10709–10714 (2008).

39. Möglich, A., Ayers, R. A. & Moffat, K. Design and Signaling Mechanism of Light-Regulated Histidine Kinases. *J Mol Biol* **385**, 1433–1444 (2009).
40. Bose, M., Groff, D., Xie, J., Brustad, E. & Schultz, P. G. The Incorporation of a Photoisomerizable Amino Acid into Proteins in *E. coli*. *J. Am. Chem. Soc.* **128**, 388–389 (2006).
41. Merino, E. Synthesis of azobenzenes: the coloured pieces of molecular materials. *Chem Soc Rev* **40**, 3835–3853 (2011).
42. Padilla, M. S. & Young, D. D. Photosensitive GFP mutants containing an azobenzene unnatural amino acid. *Bioorg. Med. Chem. Lett.* **25**, 470–473 (2015).
43. Wang, M., Funabiki, K. & Matsui, M. Synthesis and properties of bis(hetaryl)azo dyes. *Dyes and Pigments* **57**, 77–86 (2003).
44. Lampkowski, J. S., Villa, J. K., Young, T. S. & Young, D. D. Development and Optimization of Glaser–Hay Bioconjugations. *Angew. Chem. Int. Ed.* **54**, 9343–9346 (2015).
45. Kalia, J. & Raines, R. T. Advances in Bioconjugation. *Curr Org Chem* **14**, 138–147 (2010).
46. Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat Chem Biol* **7**, 876–884 (2011).
47. Hermanson, G. T., *Journal*, 1 online resource.
48. P. Khare, A. Jain, A. Gulbake, V. Soni, N. Jain and S. Jain, *Critical Reviews in Therapeutic Drug Carrier Systems*, **26**, 119-155 (2002).

49. Duncan, R., Ringsdorf, H. & Satchi-Fainaro, R. Polymer therapeutics--polymers as drugs, drug and protein conjugates and gene delivery systems: past, present and future opportunities. *J Drug Target* **14**,337–341 (2006).
50. Lo, K. K.-W., Lau, J. S.-Y., Ng, D. C.-M. & Zhu, N. Specific labelling of sulfhydryl-containing biomolecules with redox-active N-(ferrocenyl)iodoacetamide. *J. Chem. Soc., Dalton Trans.* 1753–1756 (2002).
51. Wang, Q. *et al.* Bioconjugation by Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J. Am. Chem. Soc.* **125**, 3192–3193 (2003).
52. Nguyen, D. P. *et al.* Genetic Encoding and Labeling of Aliphatic Azides and Alkynes in Recombinant Proteins via a Pyrrolysyl-tRNA Synthetase/tRNACUA Pair and Click Chemistry. *J. Am. Chem. Soc.* **131**, 8720–8721 (2009).
53. Plourde, G. L. & Spaetzel, R. R. Regioselective Protection of the 4-Hydroxyl of 3,4-Dihydroxy-benzaldehyde. *Molecules* **7**, 697–705 (2003).
54. Young, D. D., Torres-Kolbus, J. & Deiters, A. Microwave-assisted synthesis of unnatural amino acids. *Bioorganic & Medicinal Chemistry Letters* **18**, 5478–5480 (2008).
55. The Biology Project, *Antibody Structure*, The University of Arizona, June 12<sup>th</sup> 2000,  
<http://www.biology.arizona.edu/immunology/tutorials/antibody/structure.html>
56. Faden, R. R., Chalkidou, K., Appleby, J., Waters, H. R. & Leider, J. P. Expensive Cancer Drugs: A Comparison between the United States and the United Kingdom. *Milbank Quarterly* **87**, 789–819 (2009).

57. Goonetilleke, K. S. & Siriwardena, A. K. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *European Journal of Surgical Oncology (EJSO)* **33**, 266–270 (2007).
58. Huisgen, R. 1,3-Dipolar Cycloadditions. Past and Future. *Angew. Chem. Int. Ed. Engl.* **2**, 565–598 (1963).
59. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective ‘Ligation’ of Azides and Terminal Alkynes. *Angewandte Chemie* **114**, 2708–2711 (2002).
60. Himo, F. *et al.* Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *J. Am. Chem. Soc.* **127**, 210–216 (2005).
61. Kolb, H. C. & Sharpless, K. B. The growing impact of click chemistry on drug discovery. *Drug Discovery Today* **8**, 1128–1137 (2003).
62. Appukkuttan, P., Dehaen, W., Fokin, V. V. & Van der Eycken, E. A Microwave-Assisted Click Chemistry Synthesis of 1,4-Disubstituted 1,2,3-Triazoles via a Copper(I)-Catalyzed Three-Component Reaction. *Org. Lett.* **6**, 4223–4225 (2004).
63. Halliwell, B. & Gutteridge, J. M. Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem J* **219**, 1–14 (1984).
64. Hunt, J. V., Dean, R. T. & Wolff, S. P. Hydroxyl radical production and autoxidative glycosylation. Glucose autoxidation as the cause of protein damage in the experimental glycation model of diabetes mellitus and ageing. *Biochem. J.* **256**, 205–212 (1988).

65. Kocha, T., Yamaguchi, M., Ohtaki, H., Fukuda, T. & Aoyagi, T. Hydrogen peroxide-mediated degradation of protein: different oxidation modes of copper- and iron-dependent hydroxyl radicals on the degradation of albumin. *Biochim. Biophys. Acta* **1337**, 319–326 (1997).
66. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
67. van Berkel, S. S. *et al.* Metal-free triazole formation as a tool for bioconjugation. *Chembiochem* **8**, 1504–1508 (2007).
68. Blackman, M. L., Royzen, M. & Fox, J. M. The Tetrazine Ligation: Fast Bioconjugation based on Inverse-electron-demand Diels-Alder Reactivity. *J Am Chem Soc* **130**, 13518–13519 (2008).
69. Song, W., Wang, Y., Qu, J., Madden, M. M. & Lin, Q. A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angew. Chem. Int. Ed. Engl.* **47**, 2832–2835 (2008).
70. Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **39**, 1272–1279 (2010).
71. Porcelli, M. *et al.* Non-thermal effects of microwaves on proteins: thermophilic enzymes as model system. *FEBS Letters* **402**, 102–106 (1997).
72. CEM, COOLMATE, <http://www.cem.com/coolmate.html>
73. Roversi, E., Monnat, F., Vogel, P., *Helv. Chim. Acta* **85**, 733 (2002).
74. Ormö, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).

75. Wong, W.-Y. Recent Advances in Luminescent Transition Metal Polyynes. *J Inorg Organomet Polym* **15**, 197–219 (2005).
76. Shi Shun, A. L. K. & Tykwinski, R. R. Synthesis of Naturally Occurring Polyynes. *Angewandte Chemie International Edition* **45**, 1034–1057 (2006).
77. Lu, W., Zheng, G., Haji, Aisa, A. & Cai, J. First total synthesis of optically active panaxydol, a potential antitumor agent isolated from *Panax ginseng*. *Tetrahedron Letters* **39**, 9521–9522 (1998).
78. Nakayama, S. *et al.* TX-2152: A conformationally rigid and electron-rich diyne analogue of FTY720 with in vivo antiangiogenic activity. *Bioorganic & Medicinal Chemistry* **16**, 7705–7714 (2008).
79. Slepko, A. D., Hegmann, F. A., Eisler, S., Elliott, E. & Tykwinski, R. R. The surprising nonlinear optical properties of conjugated polyynes. *The Journal of Chemical Physics* **120**, 6807–6810 (2004).
80. Glaser, C. Beiträge zur Kenntniss des Acetylnylbenzols. *Ber. Dtsch. Chem. Ges.* **2**, 422–424 (1869).
81. Hay, A. S. Oxidative Coupling of Acetylenes. III. *J. Org. Chem.* **27**, 3320–3321 (1962).
82. Tripp, V. T., Lampkowski, J. S., Tyler, R. & Young, D. D. Development of solid-supported Glaser-Hay couplings. *ACS Comb Sci* **16**, 164–167 (2014).
83. Lampkowski, J. S., Durham, C. E., Padilla, M. S. & Young, D. D. Preparation of asymmetrical polyynes by a solid-supported Glaser–Hay reaction. *Org. Biomol. Chem.* **13**, 424–427 (2014).

84. Stadtman, T. C. Selenocysteine. *Annual Review of Biochemistry* **65**,83–100 (1996).
85. Polycarpo, C. *et al.* An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc Natl Acad Sci U S A* **101**, 12450–12454 (2004).
86. Wessjohann, L. A. & Schneider, A. Synthesis of Selenocysteine and Its Derivatives with an Emphasis on Selenenylsulfide (-Se-S-) Formation. *Chemistry & Biodiversity***5**, 375–388 (2008).
87. Rakauskaitė, R. *et al.* Biosynthetic selenoproteins with genetically-encoded photocaged selenocysteines. *Chem. Commun. (Camb.)* **51**,8245–8248 (2015).
88. Rita Cardoso, B., Silva Bandeira, V., Jacob-Filho, W. & Franciscato Cozzolino, S. M. Selenium status in elderly: Relation to cognitive decline. *Journal of Trace Elements in Medicine and Biology* **28**, 422–426 (2014).