


5-2015

Development of Novel Chemical Techniques to Address Biological Questions

Johnathan C. Maza
College of William and Mary

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

 Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Biochemistry Commons](#), and the [Chemical Actions and Uses Commons](#)

Recommended Citation

Maza, Johnathan C., "Development of Novel Chemical Techniques to Address Biological Questions" (2015). *Undergraduate Honors Theses*. Paper 144.
<https://scholarworks.wm.edu/honorstheses/144>

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.

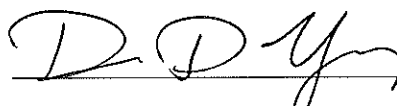
Development of Novel Chemical Techniques to Address Biological Questions

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

by

Johnathan Charles Maza

Accepted for Honors



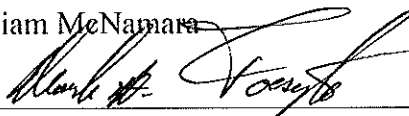
Dr. Douglas Young, Committee Chair



Dr. Lisa Landino



Dr. William McNamara



Dr. Mark Forsyth

Williamsburg, VA

May 5, 2015

Development of Novel Chemical Techniques to Address Biological Questions

Johnathan C. Maza

Purcellville, Virginia

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

Chemistry Department

The College of William and Mary

May 5, 2015

ABSTRACT

The current biological toolkit has been vital in advancing our understanding of the world. That being said, the toolkit has limitations. As such, chemical biologists have been developing novel means to probe biological systems using chemical techniques. Bioorthogonal chemistry represents a new avenue to address biological questions that cannot be answered using current techniques. Herein, we describe a novel technique to probe proteins-of-interest using unnatural amino acid (UAA) mutagenesis. We have found that our UAAs allow us to access bioorthogonal chemistries for the conjugation of fluorophores to UAA-containing proteins. Additionally, we have extended these findings towards the application of protein immobilization. Finally, we used microwave technology to investigate novel means to transform bacterial cells with exogenous DNA.

TABLE OF CONTENTS

Chapter 1: Introduction to Unnatural Amino Acids and Bioconjugations	7
1.I. Unnatural Amino Acids	11
1.I.A. Incorporation of Unnatural Amino Acids	13
1.II. Bioorthogonal and Click Chemistries	18
1.II.A. Maleimide-Thiol Chemistry	19
1.II.B. Cu(I)-1,3-Dipolar Cycloaddition or Cu(I) “Click” Chemistry	20
1.II.C. Glaser-Hay Coupling Chemistry	23
Chapter 2: Development of Unnatural Amino Acid Technology for Bioconjugation Applications	26
2.I. Introduction	26
2.II. Results and Discussion	28
2.II.A. Synthesis of Unnatural Amino Acids	29
2.II.B. Incorporation of Unnatural Amino Acids	34
2.II.C. Bioorthogonal Chemistry Experiments	36
2.II.D. Bioorthogonal Chemistry Applied to Protein Immobilization	43
2.III. Experimental	47
2.IV. Future Work	58
2.IV.A. Condensation Chemistry	58
2.IV.B. Continued Immobilization Trials	60
2.V. Conclusion	60
Chapter 3: Development of Rapid Microwave-Mediated and Low-Temperature Bacterial Transformations	62
3.I. Introduction	62
3.II. Results and Discussion	65
3.III. Experimental	72
3.IV. Conclusion	74
Appendix A: NMR Spectra	75
Acknowledgements	87
References	88

TABLE OF SCHEMES

Scheme 1. General Reaction for Addition of a Thiol across a Maleimide Double Bond ..	19
Scheme 2. General Reaction Scheme for the Cu(I)-1,3-Dipolar Cycloaddition, or the Copper “Click”	21
Scheme 3. Comparison between Cu “Click” and Strained Octyne [3+2] Cycloaddition ¹⁷	22
Scheme 4. General Reaction of the Glaser-Hay Coupling	24
Scheme 5. Final Overall Reaction Scheme for the Synthesis of UAAs	30
Scheme 6. Elimination Byproduct of the Bromotyrosine Synthesis	31
Scheme 7. Synthesis of Alkynyltyrosine Derivatives Using Propargyl Alcohol	32
Scheme 8. Synthesis of 6-Bromo-1-hexyne	33
Scheme 9. General Reaction for the Cu(I)-1,3-Dipolar Cycloaddition (or Copper(I) “Click”)	37
Scheme 10. Initial Click Trials Using pAzF-GFP and Alexafluor 488	37
Scheme 11. General Schema for the Glaser-Hay Coupling	39
Scheme 12. General Reaction Schema for the Nucleophilic Substitution of Bromo-GFP by a Reactive Thiol Group	42
Scheme 13. General Scheme of the Oxime Ligation	58
Scheme 14. General Reaction Scheme for Synthesis of Aminoxy-UAAs	60

TABLE OF FIGURES

Figure 1. Protein immobilization techniques.....	10
Figure 2. The 20 canonical amino acids	12
Figure 3. Unnatural amino acids.	13
Figure 4. Mechanism for genetic incorporation of unnatural amino acids.....	16
Figure 5. Double-sieve selection of aminoacyl-tRNA synthetase	18
Figure 6. UAA residue context dependence on immobilization.....	28
Figure 7. Example of UAA suppression of the TAG codon.....	35
Figure 8. Expression of GFP-151 mutants with azide tethered tyrosine UAAs	39
Figure 9. Expression of GFP-151 mutants with alkyne tethered tyrosine UAAs	41
Figure 10. Oxidation states of Glutathione.....	42
Figure 11. Expression of GFP-151 mutants with bromo tethered tyrosine UAAs	43
Figure 12. Immobilization of an alkyne dye via Glaser-Hay coupling.....	44
Figure 13. Expression of GFP-151 mutants with azido tethered tyrosine UAAs	46
Figure 14. Validation of DNA transformation.	71

TABLE OF TABLES

Table 1. Overall Yields For the Synthesis of all Deprotected UAAs	34
Table 2. Microwave Assisted Transformation Efficiencies	67
Table 3. Freeze-Thaw Transformation Efficiencies.....	70

CHAPTER 1: INTRODUCTION TO THE UNNATURAL AMINO ACIDS AND BIOCONJUGATIONS

In the sequential arrangement of amino acids that form proteins, nature has achieved elegant macromolecules with properties chemists have yet to fully replicate.¹ Based on their biological functions, proteins can both bind small molecules with incredibly high affinities and catalyze reactions with ease that often require harsh conditions *in vitro*. Based on these characteristics, harnessing the power of proteins would allow scientists to develop novel diagnostic and catalytic tools that would advance science. Additionally, these unique properties of proteins make them responsible for the myriad of cellular functions that enable living systems to exist. As such, molecular biology has been revolutionized by the ability to study and observe proteins *in vivo*. Thus, the study of proteins and their development towards novel technologies is a crucial area of science that constantly requires additional innovations.

For example, a serious issue facing researchers wishing to fight cancer is discovering new ways to accurately detect the presence of cancer within patients. *Nature* recently published a report detailing the use of fluorescent probes as a means of live imaging cancer cells during surgery based on the presence and detection of certain proteins.² The results of such technology will enhance removal of malignant tissue while decreasing relapse rates and damage to patient's healthy tissue. Accordingly, fluorescent molecular imaging is proving to be a key player in the war against cancer, and while it has

come a long way, there are limitations that must be overcome before such specialized applications become widespread.

Common fluorescent-imaging techniques rely on either small fluorescent dyes or fluorescent proteins to serve as labels. These methods often require either the diffusion of an excess of fluorescent molecules into the sample organism or extensive genetic manipulation of the organism's genome to fuse fluorescent proteins to the protein of interest.³ Immunolabeling techniques, which rely on the very specific binding of an antibody-fluorophore complex to a specific cellular structure, are often confounded by high background fluorescence due to non-specific binding. Genetic tagging of proteins relies on the extensive manipulation of the organism's genome, and involves the incorporation of a large fluorescent protein moiety into the protein of interest. By altering the protein's structure, genetic tagging risks rendering the protein inactive or altering its normal function. In addition, the fluorescence signal produced is constant and cannot be controlled by the researcher. Thus, the vital importance of studying proteins in molecular biology necessitates that scientists evolve new means to image proteins.

Another emerging field involving proteins is the development of protein biochips. These biochips contain dense spots, or microarrays of thousands of diverse or homogenous proteins immobilized on a solid support.^{4,5} They offer the potential to harness the unique power of protein chemistry. For example, such technology can allow for the rapid and high-throughput screening of thousands of macromolecules of biological interest, allowing for the rapid discovery of new drug candidates. Similar to the *Nature* paper above, such protein-based technology can be directly applied to diagnostics, such as detecting cancer

markers in blood.^{6,7} Due to their high affinity for their substrate, protein chips would allow for the rapid and sensitive detection of biomolecules and toxins.

Despite their promise, protein-based chips are still a nascent technology, being highly limited by current immobilization techniques. Many of these involve simple adsorption onto a surface, which may not provide a robust chip for utilization.⁵ In this technique, proteins are adsorbed onto a surface via ionic or hydrophobic interactions. Because of the potential for multiple interacting sites, the result is often a heterogeneous layer with a random protein orientation. This is disadvantageous, as the protein may not be correctly oriented to perform the desired function. Covalent bonding to the support is often the most ideal mechanism to ensure an irreversible attachment. By reacting a functionalized support with key amino acid residues, such as lysine or cysteine residues that respectively contain amine and thiol functionalities, more robust attachments can be made. However, these amino acid side chain bonding strategies can occur at multiple sites, resulting in proteins that may either be conformationally inactive or also inappropriately oriented to bind molecules of interest.⁵ Bioaffinity based attachment, in which key protein structures or properties are exploited to bind a high affinity ligand, such as another protein (i.e. antibody), is an option that can allow for the binding of proteins in specific orientations. However, this technique relies on non-covalent binding, and is thus less robust. Furthermore, it limits the number of proteins that can be bound to the surface, since proteins must have particular structures or properties in order to interact with the ligand.

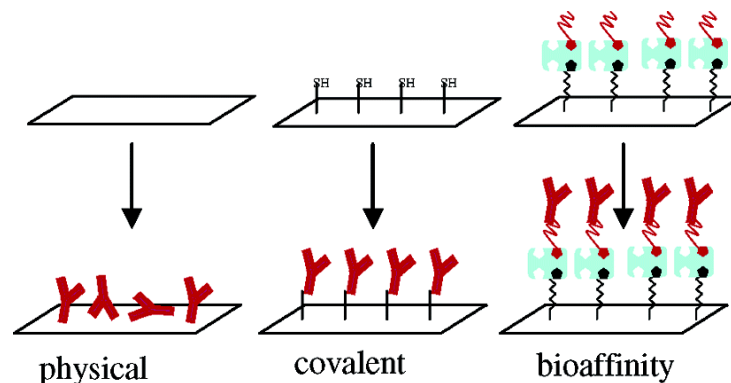


Figure 1. There are numerous methods for immobilizing proteins on a solid support. Covalent attachment of the protein represents the most robust technique, although it is often complicated by the presence of multiple reaction partners on a single peptide. Adapted Rusmini et al.⁵

Before this biochip technology can become widespread, scientists must solve these immobilization problems. Additionally, the need to find new ways to probe protein structure and function also creates the need to develop new ways to manipulate protein structure and function. A potential alternative to both of these areas involves the application of unnatural amino acid technology.

Unnatural amino acids (UAAs) allow scientists to expand the rather limited biochemical functionality found in the 20 canonical amino acids and introduce novel reactivity not typically found in proteins.^{3,8} Incorporated via suppression of a TAG stop codon in translation, which can be incorporated into the DNA at a desired residue position of the target protein, “chemical handles” can be introduced on the surface of the protein. These functional handles then serve as sites of bioorthogonal reactivity. This strategy avoids the issue of multiple reaction partners on a single protein (which is typical when trying to conjugate on a cysteine/lysine residue), and allows the chemist to access new chemistries typically not available from the canonical 20 amino acids. Furthermore, compared to the strategies discussed above, this technique opens up the potential to exploit any protein that the chemist desires in a robust way.

I. UNNATURAL AMINO ACIDS

Within nature, there exist twenty canonical amino acids. These amino acids serve as building blocks for the plethora of proteins. Cellular proteins have numerous functions, as they catalyze the replication of our genetic information, provide structural support to our cells, and serve as defense mechanisms for our immune system, to name a few. The key to the proper function of these proteins is the arrangement of their amino acid sequence. By arranging the 20 amino acids in different combinations, proteins are able to capitalize on the unique structural and biochemical properties of each residue. The appropriate sequences allow for the protein to assemble properly, while also providing key catalytic functionality should it be required by the protein. Given the large number of functions proteins serve within an organism, as well as the myriad of ways in which they could be combined, it is reasonable to assume that evolution would have dictated the need for amino acids with diverse functional groups. However, a quick glance at the natural 20 amino acids reveals only four main chemical functionalities: nonpolar, polar, acidic, and basic.^{3,9}

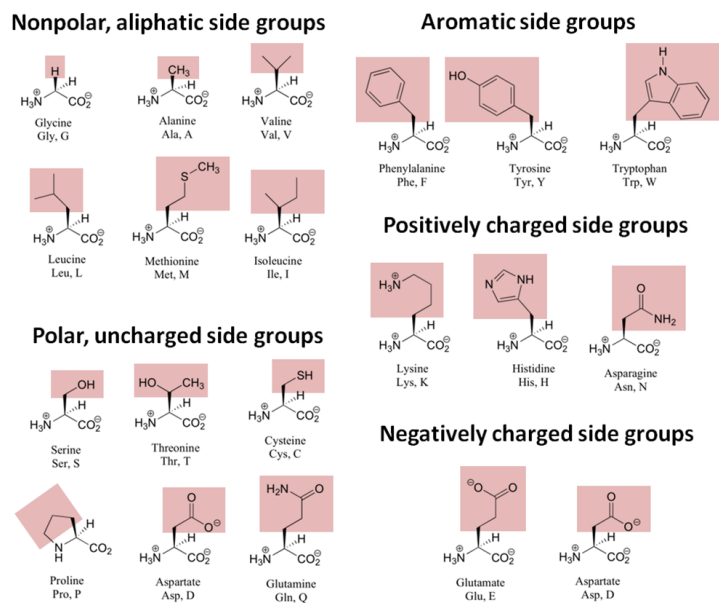


Figure 2. There are 20 naturally occurring amino acids. Only five elements are represented in the canonical amino acids, as well as a very limited set of chemical functionality. Adapted from chemwiki.ucdavis.edu

Advances in chemical biology have led to the development of UAAs that can be synthesized with novel biochemical properties not available in the 20 canonical amino acids. Such syntheses are relatively simple, as amino acids can be readily purchased with their carboxyl and amino ends masked with protecting groups. The result is a selectively reactive R-group that can be chemically modified. In particular, my research has focused on derivativizing the amino acid tyrosine (Tyr) that has a phenol functionality that can be activated by the addition of base. The deprotonated hydroxyl group can then undergo S_N2 reactions with molecules that possess good leaving groups. The result is a UAA that can serve as a starting point to introduce a myriad of novel biochemical functionality into proteins.

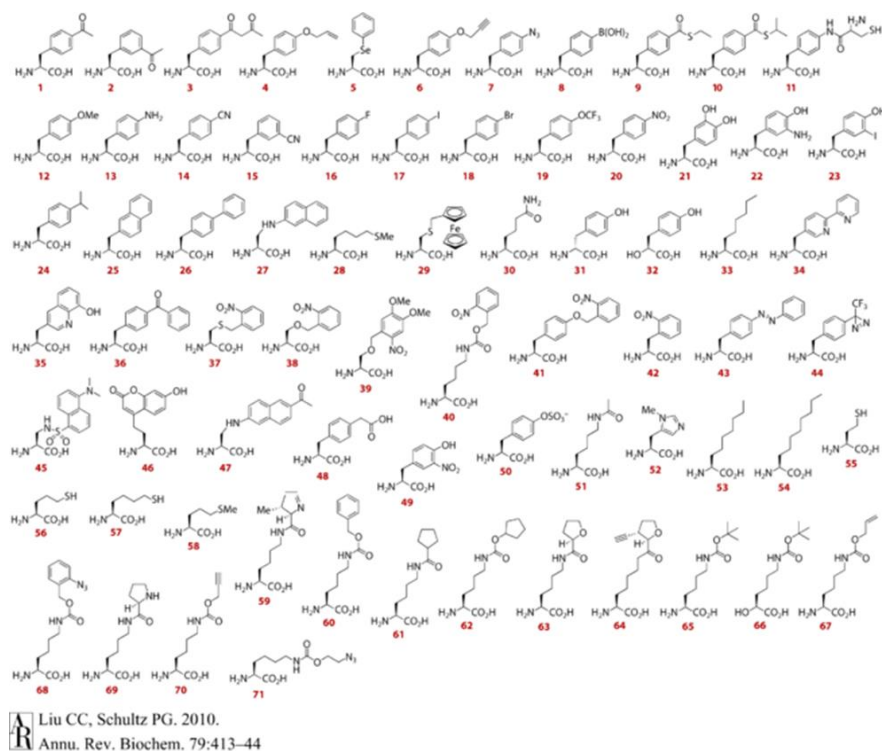


Figure 3. Unnatural amino acids allow chemists to access novel functionality and elements not present in the 20 canonical amino acids.

I.A. Incorporation of Unnatural Amino Acids

While the synthesis of UAAs is feasible, in order to capitalize on this technology it is necessary to be able to insert these UAAs into proteins of interest. Methods have been developed to site-specifically incorporate these UAAs into proteins via insertion in response to a TAG stop codon in the genome, giving researchers very precise control over the placement and number of these UAAs in a protein's amino acid sequence.³

Early methodology relied on incorporating UAAs *in vitro*. By utilizing cell-free protein assembly machinery, researchers were able to incorporate unnaturals into specific sites of the protein via a nonsense codon. The associated tRNA synthetase for this nonsense codon was then chemically aminoacylated with the UAA of interest. The result is a system that ensured the incorporation of the UAA into a single, predefined residue of the polypeptide. In particular, the amber suppressor codon (TAG) has been employed, as its

frequency in the *Escherichia coli* genome, the organism of choice for protein expressions, is relatively low. This ensures the incorporation of a single UAA into a single site.^{10,11} While this approach allowed for the site-specific incorporation of a UAA into a protein's primary sequence, the use of *in vitro* technology had serious limitations. Many proteins require the aid of additional proteins *in vivo* before they can acquire the correct conformation. As such, while this technique could impact the formation of small peptides, polypeptides could not be generated in a robust manner. This severely limited the biological applicability of UAAs. Furthermore, the synthesis of mutated proteins *in vitro* limits the use of this technology beyond biochemical studies. For example, novel methods of protein imaging *in vivo* could not be developed since this technique required cell-free machinery.

An alternative strategy for UAA incorporation that allowed for the *in vivo* incorporation of UAAs involved the complete replacement of a particular amino acid in the organism's proteome with a UAA. In this non-site specific UAA incorporation technique, the canonical amino acids methionine, leucine, isoleucine, phenylalanine, tryptophan, or proline are replaced by an amino acid that resembles of these amino acids in structure.¹² Typically, cells transformed with a protein of interest are grown in complete media, pelleted, and the media removed prior to protein expression. Upon the induction of protein expression, the cells are resuspended in a new media that lacks one of these canonical amino acids. In place of the amino acid, an unnatural amino acid is included in the broth. By synthesizing amino acids that resemble the removed amino acid, scientists can trick the cell into incorporating their UAA in place of the canonical amino acid in the protein of interest. Tirrell and Finn demonstrated the utility of this approach by

synthesizing methionine analogs. In place of the thioether functionality, the group included azide and alkyne functionalities in their UAAs. As a result, the group was able to show that in place of methionine, *E. coli* induced to express a bacteriophage protein capsid would generate capsids labeled with either the azide or alkyne functionality in the residues that were originally methionine. While this approach also allows researchers to incorporate UAAs in a semi-site specific fashion, the frequency of UAA placement in a protein can be highly variable, due to the redundancy inherent in the 20 amino acid lexicon. The result can be complete dysfunction of the protein if the presence of multiple UAAs alters the protein's structure too much. Additionally, this technique impacts the organism's proteome, introducing the UAA in place of the canonical amino acid in every protein expressed. The result could be highly deleterious for the organism, which is not beneficial for *in vivo* applications.¹²

Since these early *in vitro* studies, more recent work has focused on generating technology that would allow for the *in vivo* incorporation of UAAs. Additional challenges come into play when considering such techniques. First, as mentioned above, the incorporation of the UAA must be site specific, and it must occur only in the residue sites that have been previously determined by the researcher. Should the UAA be incorporated into additional non-specified sites, the result could be inactivation of the protein of interest, or subsequent death of the organism. Additionally, the tRNA machinery employed must act solely on the UAA of interest. By translating unnatural amino acid technology towards *in vivo* incorporation, the added problem of having to deal with three elements arises. These elements include the UAA, the tRNA, and the aminoacyl-tRNA synthetase (aaRS) that is responsible with charging the tRNA with the UAA. In order for successful incorporation

to occur, the tRNA synthetase should only pair with the UAA and with the corresponding suppressor tRNA. This independence relative to the endogenous system is often referred to as orthogonality. Should such orthogonality not be maintained, endogenous amino acids could be incorporated into proteins in place of the UAA and vice versa. Finally, the UAA must be cell permeable and not cytotoxic in order to be incorporated.^{3,13}

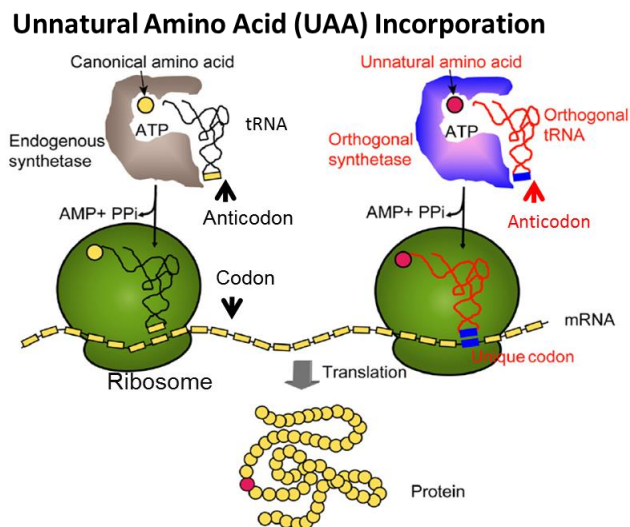


Figure 4. Mechanism for genetic incorporation of unnatural amino acids. An orthogonal synthetase must be introduced that does not cross-react with any natural tRNA or amino acid. This is coupled with a new tRNA that harbors an anticodon for the TAG codon. When the ribosome encounters the TAG codon, a new amino acid is incorporated, rather than the discontinuation of translation. Adapted from Chem. Biol. **2009** 16, 323-336.

Early studies with *E. coli* models attempted to evolve an endogenous tRNA synthetase/tRNA pair that was orthogonal to a UAA of interest utilizing *E. coli*'s translational machinery. However, such tactics were wrought with misaminoacylation, and subsequently terminated.¹⁴ Instead, researchers focused on examining the protein translation machinery of other organisms, which would not interact with the endogenous aaRS/tRNA pairs but still allow efficient incorporation of the UAA. *In vitro* studies found that a tyrosine aaRS/tRNA pair from the archaea *M. jannaschii* was the first of such pairs used to successfully incorporate a UAA in an *E. coli* model.¹⁵

With acceptable aaRS/tRNA machinery in hand, researchers next focused on a means to alter the specificity of the pair to recognize different UAAs. This was done through a double-sieve selection method. First, a library of t-RNA synthetases was constructed based off the aaRS of *M. jannaschii*. Within this library, five to six residues within the binding site of the aaRS were genetically mutated to be one of the 20 amino acids, generating a library size on the order of 10^6 to 10^7 mutant aaRS. This library size is still viable for conventional double transformation procedures. Next, the library underwent a positive selection. In this step, *E. coli* were transformed with the library of mutated synthetases and an enzyme encoding for a mutated-chloramphenicol degrading enzyme that contained an early TAG residue. Following double-transformation, the transformed cells were grown on a plate in the presence of a novel UAA and chloramphenicol. Survival implies the organism was able to use the foreign aaRS machinery to charge the suppressor tRNA and thus produce a functional chloramphenicol degrading enzyme.¹⁶ However, suppression does not guarantee that the aaRS charged the suppressor TAG using the UAA. As such, those cells that grow were then scraped off the plate and underwent a plasmid extraction. Following isolation on gel-electrophoresis, the sieved plasmid library is then subjected to a negative selection. In the negative selection round the toxic barnase gene is used. Barnase is a bacterial ribonuclease, and will degrade the cell's RNA, preventing protein synthesis and leading to cell death. The gene was mutated to harbor three stop codons, and grown in the absence of unnatural amino acid. If the organism can use the exogenous aaRS/tRNA pair to charge endogenous amino acids and suppress the three stop codons, it will produce the barnase protein and die.¹⁶ Thus, the end of this two round procedure selected for organisms that possessed an aaRS/tRNA machinery that could

selectively incorporate the UAA to some degree. By iterating this process for multiple rounds with the UAA of interest, it was possible to evolve an aaRS/tRNA pair that works selectively for the UAA of interest. This is the evolution procedure used today.

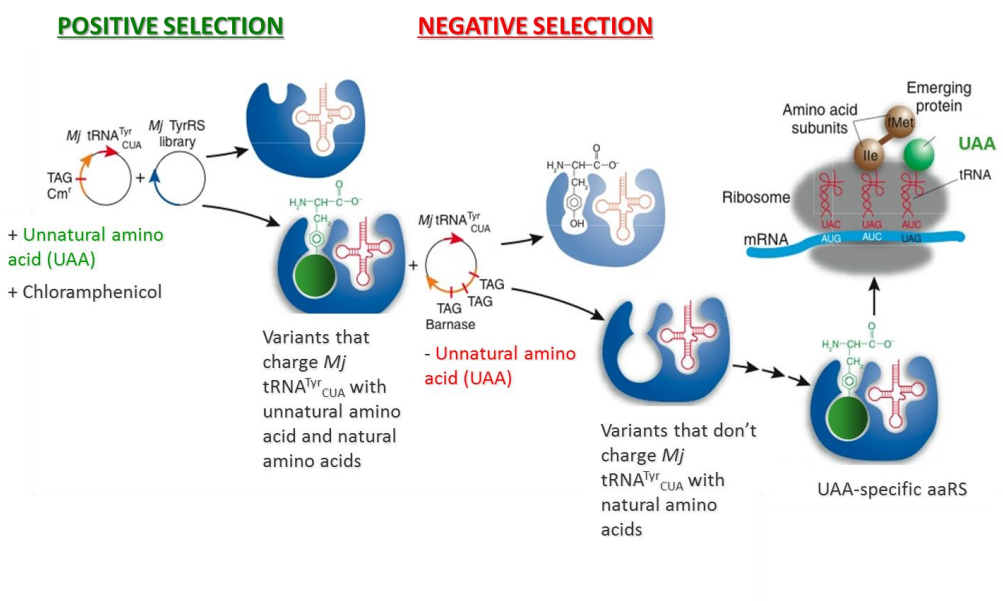


Figure 5. General scheme for the selection of an aminoacyl-tRNA synthetase that recognizes and incorporates a specific unnatural amino acid.

II. BIOORTHOGONAL AND CLICK CHEMISTRIES

A growing area of interest in chemical biology is the ability to perform chemistry, particularly conjugation reactions, within living systems. In order for such chemistry to be feasible though, certain conditions must be met by the chemical reaction employed. Most importantly, the chemistry must not interfere with biological processes. That is, the reaction must be completely chemoselective, as well as rely on reactivity absent from endogenous biological chemical functionalities. Furthermore, the reaction must proceed at relatively rapid rates under physiological conditions (typically at 37°C and ~pH 7) as well as being inert to the myriad of functionalities encountered in biological systems.¹⁷

Should the conditions above be met, the chemistry is then deemed bioorthogonal. Bioorthogonality implies the reaction is capable of being performed in a biological system

in a selective manner without interfering with the already ongoing physiological processes. However, when considering the numerous functionalities found within living systems, as well as the limitations of concentrations and temperatures, the task of identifying bioorthogonal chemistries seems daunting.

II. A. Maleimide-Thiol Chemistry

One of perhaps the earliest attempts at performing chemical reactions in a living organism was the conjugation reaction between a thiol and a maleimide.^{18,19} The technique has become a common approach for the labeling of reactive thiol groups in proteins, typically in the form of cysteine residues. In this reaction, the thiol acts as a nucleophile, adding across the double bond of maleimide to form a thioether. Furthermore, the reaction demonstrates a selectivity for thiols versus amines (a potential confounder), as the reaction with amines can only be accessed at higher pH ranges. Finally, hydrolysis of the maleimide substrate to an unreactive product is minimized at pH ranges below 8, well within the limits of a typical cellular environment.¹⁸

Scheme 1. General Reaction for Addition of a Thiol across a Maleimide Double Bond



While the reactions between maleimides and thiols represent chemistry that can be performed in a cellular context, the chemistry is not truly bioorthogonal due to the presence of naturally occurring thiol groups in proteins. Current uses of the reaction involve labeling thiol-containing proteins with either a maleimide-derivatized biotin group (to increase molecular weight for SDS-PAGE isolation or for streptavidin purification) or a maleimide-containing fluorescent probe. That being said, the maleimide functionality could be

integrated into a protein via UAA-technology, allowing for novel means of protein coupling with proteins containing reactive thiol groups or new labeling methods utilizing thiol-containing dyes.

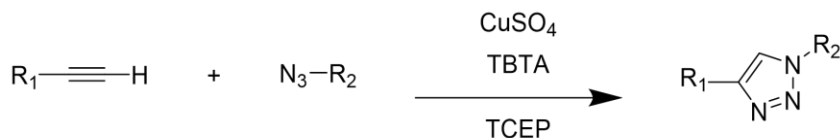
II. B. Cu-1,3-dipolar Cycloaddition or Cu(I) “Click” Chemistry

Click chemistry, a subset of reactions that can proceed to completion in relatively mild conditions to produce a single product with minimal to no side reactions, has emerged as a means for performing bioorthogonal chemistry. In particular, the [3+2] cycloaddition of azides and alkynes, termed a Huisgen-1,3-dipolar cycloaddition, has flourished as a means to tackle the complexity of bioorthogonality.²⁰ Not only do these reactions produce a stable, covalent triazole moiety, thus ensuring efficient and stable conjugation between the two functionalities *in vivo* where bonds may readily decompose, but the azide and alkyne functionalities are also not found in biological systems. Furthermore, the two molecules react selectively with each other, thus side reactions with other biological molecules is not a concern. Based on these facts, the 1,3-dipolar cycloaddition using azides and alkynes is an optimal chemistry for the development of bioconjugation reactions. However, the initial reaction reported in the early 20th century required temperatures upwards of 98°C, well above the temperature range of physiological viability.²⁰

The viability of this chemistry was restored when the Sharpless and Meldal groups concurrently, yet separately, developed a copper(I) catalyzed variant of the cycloaddition.^{21,22} Reduction of copper(II) to copper(I) in the presence of a terminal alkyne allows for the formation of an activated copper acetylide intermediate, which has an increased reactivity with azides to form the triazole ring product. This copper-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC) exhibited an increased reaction rate of

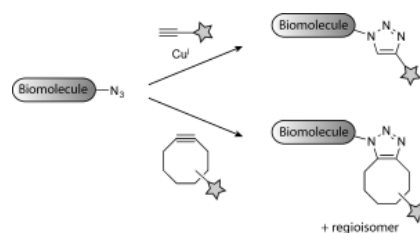
seven orders of magnitude, as well as being capable of reacting at room temperature and even 4°C.²³ The fact that the reaction can be made even more favorable by the addition of tetradentate ligands, such as tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA), that help stabilize the copper(I) catalyst resulted in this chemistry rising to the forefront of bioconjugation reactions.²⁴

Scheme 2. General Reaction Scheme for the Cu(I)-1,3-Dipolar Cycloaddition, or the Copper “Click”



While the CuAAC reaction afforded desirable temperature ranges and reaction rates, it possessed a single hindrance to its use in biological systems. The reaction's dependence on copper was ultimately its greatest flaw, as copper is cytotoxic to most organisms. Bacterial cells have been shown to halt cell division following exposure to 100 μM CuBr, and mammalian cells can only endure very low concentration of copper (below 500 μM) for 1 hour.¹⁷ As such, the Bertozzi group sought to develop a copper free variant that may proceed with similar favorability by capitalizing on ring strain. In particular, cyclooctyne, the smallest ring that can contain a triple bond, was used to activate the alkyne towards the cycloaddition via strain. The result was a cycloaddition, however, the reaction proceeded at rates considerably slower than CuAAC.²⁵ By addition of electron withdrawing fluorines, the Bertozzi lab was able to generate a difluorinated cyclooctyne (termed DIFO) capable of undergoing at copper free click reaction at rates comparable to CuAAC. Thus, the cyclooctyne reagent is nontoxic and can be derivitized to include additional functionality, such as conjugated fluorescent moieties.²⁶

Scheme 3. Comparison between Cu "Click" and Strained Octyne [3+2] Cycloaddition¹⁷



CuAAC and strain promoted click chemistries (SPAAC) both serve as useful reactions for the bioconjugation of proteins bearing UAAs with other molecules. Because the placement of UAAs within a protein can be site-specifically controlled, UAAs can be inserted into proteins in unique residues that minimize changes to the protein's folding while allowing for maximal exposure of the UAA. These surface exposed residue sites allow for UAAs to serve as unique biochemical handles for the targeted promotion of these bioconjugation reactions. If UAAs can be synthesized with azide or alkyne functionality, they could be correspondingly clicked to another molecule bearing the appropriate cognate handle. Such a strategy would allow for the directed control of bioorthogonal chemistry towards proteins of interest.

Indeed, both CuAAC and SPAAC have already begun to receive attention as novel methodologies to probe biological systems. Research on intracellular pathogens is often complicated due to the difficulty of probing the pathogen in their endosomal environment. Methods such as antibody labeling are complicated by the need to pass through multiple membrane structures. As such, Siegrist *et al.* demonstrated that UAA analogs of D-alanine, which harbor an alkyne functional group, can be used to probe peptidoglycan dynamics of intracellular bacterial pathogens *in vivo*, a technique that was not previously accessible using more traditional antibody labeling approaches.²⁷ In particular, the authors demonstrated that the incorporation of the alkyne or azide derivitized D-alanine UAAs

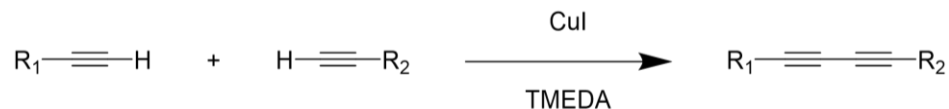
allowed the researchers to perform CuAAC reactions using the orthogonal probe on a fluorescent dye. Because the D-alanine analogs are used exclusively by bacteria in the synthesis of their peptidoglycan layer, the researchers were able to label intracellular pathogens, such as *Listeria monocytogenes*, in an *in vivo* model and then observe peptidoglycan dynamics in real time.

Beyond UAAs, CuAAC chemistry can also be extended to other biomolecules, like sugars. *Mycobacterium tuberculosis*, the causative agent of tuberculosis, uses the non-mammalian disaccharide, trehalose, as a precursor for essential cell wall glycolipids, as well as other metabolites.²⁸ However, not many strategies currently exist to visualize the metabolism of this sugar essential to *M. tuberculosis*. As such, the Bertozzi group generated an azide-trehalose derivative, which they could subsequently ligate to alkyne-functionalized probes via CuAAC chemistry. This allowed the group to explore the fate of the trehalose analogs, generating novel techniques for the investigation of the mycobacterial trehalome and how this sugar is incorporated into the cell membrane.

II. C. Glaser-Hay Coupling Chemistry

As demonstrated by the click chemistries, it is apparent that reactions involving terminal alkynes are advantageous for use in biological systems due to the relative absence of triple bonds within these systems. The Glaser-Hay coupling is a reaction that involves conjugating two terminal alkynes via a copper catalyst to afford a conjoined diyne product^{29,30} Salts of copper(I) are typically used as the catalyst, and it has been proposed that the copper(I) complexes with the terminal alkyne, activating the functionality towards the formation of a carbon-carbon bond with another terminal alkyne.^{31,32}

Scheme 4. General Reaction of the Glaser-Hay Coupling



The coupling of terminal alkynes was first described by Carl Glaser in 1869.²⁹ However, the initial reaction was not suited for biological settings. Glaser's initial reaction, however, involved an explosive copper(I)-phenylacetylide catalyst that required tedious isolation following the reaction.²⁹ Then, in 1962, Allan Hay published results using the bidentate ligand N,N,N',N'-tetramethylethylenediamine (TMEDA) in conjunction with the copper catalyst. This methodology improved reaction times, reduced temperature requirements, and reduced the amount of copper necessary to allow the reaction to proceed.³⁰ By conjugating the Cu(I) catalyst, it is proposed that the bidentate ligand TMEDA stabilizes the Cu(I) species, opening up the complex to favor addition of the activated alkyne species to one end of the copper species.^{31,32} The Glaser-Hay reaction today, in which a Cu(I) salt in the presence of TMEDA in an organic solvent is added to two terminal alkynes, allows for the rapid and facile synthesis of conjugated diynes. Furthermore, recent work has demonstrated the reaction can proceed in good yield under biological temperatures (4°C to 37°C), and has enhanced reaction rates.³³ While little work has been done to take advantage of these unique features of this chemistry, these characteristics make the Glaser-Hay coupling ideal for the utilization of the bioorthogonal alkyne moiety for bioconjugation applications.

III. CONCLUSION

UAA-mutagenesis technology allows the researcher to create site-specific functional “handles” on proteins of interest. Herein we propose applying UAA-

mutagenesis for the development of novel biological probes and immobilization techniques. By combining this technique with bioorthogonal chemistries, we hope to manipulate biological systems and proteins in new ways that can further advance our society and understanding of the world.

CHAPTER 2: DEVELOPMENT OF UNNATURAL AMINO ACID TECHNOLOGY FOR BIOCONJUGATION APPLICATIONS

I. INTRODUCTION

Green fluorescent protein (GFP) is a 26kDa protein isolated from jellyfish.³⁴ Unique to GFP's photochemical properties is a tripeptide chromophore composed of Ser65-Tyr66-Gly67.³⁴ The protein contains a β -barrel structure, as well as an internal α -helix within the barrel core. Within this internal α -helix is Ser65-Tyr66-Gly67. When the protein has folded correctly, these three amino acids are in the correct conformation to undergo a series of cyclization, dehydration, and oxidation reactions, which result in the formation of the GFP chromophore. This chromophore is further stabilized by the hydrophobic interior of the protein, allowing the protein to fluoresce when it is struck by the appropriate wavelength of light.

By inserting TAG mutations into the GFP DNA, it is possible to generate GFP mutants with UAAs at predefined residue positions.³⁵ These residues can be surface exposed, such as 3, 133, and 151, or they can even be part of the chromophore itself, such as residue 66. The placement of the UAA grants GFP with either new surface-exposed chemical functionality or new emission spectra. Furthermore, the unique structural properties that allow GFP to fluoresce also make GFP an ideal reporter protein for work involving UAAs. The successful suppression of the TAG stop codon via the incorporation of a UAA will generate a functional and fluorescent mutant GFP. However, if a UAA is not incorporated, for example, due to poor UAA-recognition by the orthogonal aaRS

machinery, then the protein translation will terminate early and will not fold correctly. Because of the stringent folding-conditions required for the formation of the GFP fluorophore, this improper folding results in lack of fluorescence. Thus, the observable fluorescence of GFP makes it an ideal reporter protein for this work involving GFP.

Recently, our lab has shown that UAA-based mutagenesis allows for the site-specific immobilization of proteins, GFP specifically, on a sepharose resin.³⁵ This technology hinges on the use of an azide-derivatized UAA which was incorporated on the surface of green fluorescent protein (GFP) via TAG suppression. By performing a Cu(I)-catalyzed 1,3-dipolar cycloaddition between an azide and an alkyne, the azide-derivatized GFP reacts with the alkyne-derivatized resin, forming a covalent triazole ring between the two in a pre-defined position (see Figure 1). Furthermore, immobilization did not impede protein fluorescence. Rather, we showed that immobilization actually aided in the maintenance of protein stability even in the presence of harsh organic solvents. Therefore, this technology confers protein-based technologies a wider range of use in harsh organic solvents and extreme environments. While this demonstrated a proof-of-concept, further research is required to better understand immobilization conditions and apply the UAA technology further.

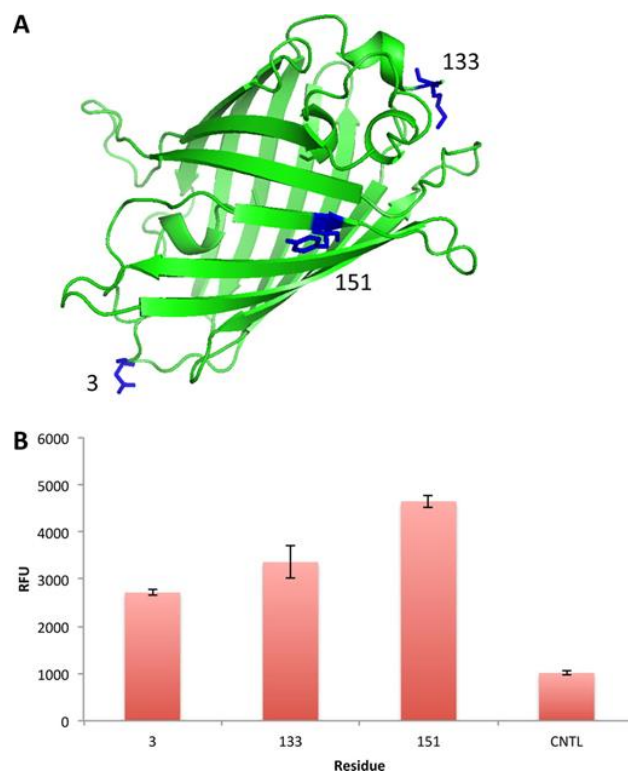


Figure 6. UAA residue context dependence on immobilization. (A) Various GFP pAzF mutants were expressed with the UAA at different positions (highlighted in blue), adapted from PDB: 1EMA.(39) (B) Fluorescence data demonstrating maximal immobilization with the pAzF-GFP-151 mutant, suggesting some importance of UAA context within the protein for the immobilization. All error bars represent standard deviations from three independent experiments.³⁵

II. RESULTS AND DISCUSSION

Initial investigations explored how UAAs could be utilized to optimize protein immobilization, and thereby generate novel protein-based technologies in the future. Two key research aims were assessed to optimize our immobilization technique. First, new bioorthogonal chemical reactions to immobilize proteins were investigated. In our previous reports, only the Cu(I)-catalyzed 1,3-dipolar cycloaddition was probed as a viable reaction to immobilize GFP; however, there currently exist a wide-range of bioorthogonal reactions which rely on less harsh catalysts and enhanced rates. In particular, we were interested in examining the Glaser-Hay coupling reaction between terminal alkynes as a prime candidate for conjugation with biological agents. This reaction is advantageous due to its rapid

reaction rates (4 hours), low temperature requirements (anywhere between 37-4 degrees Celsius), and milder conditions (CuI and TMEDA).³⁶

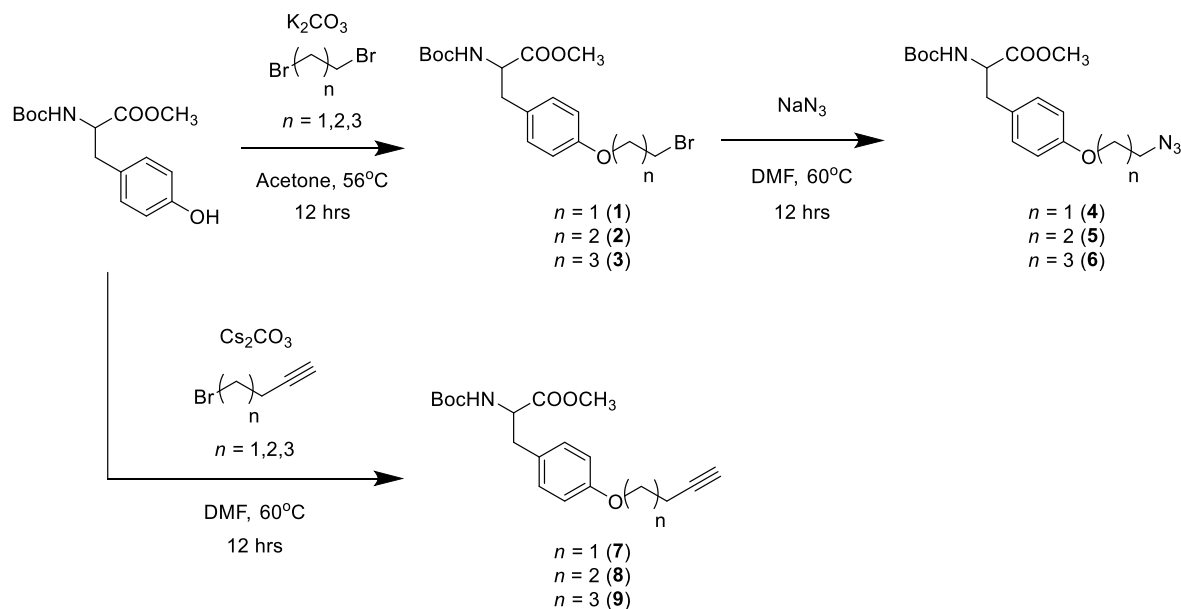
Additionally, our previous studies immobilizing protein demonstrated that the UAA residue location on the protein dictated its resin immobilization efficiency.³⁵ Thus, the spatial configuration of the UAA handle is vital for its reactivity with its conjugation partner. While it is possible to alter where the UAA can be incorporated, this requires extensive genetic manipulation. An alternative approach is to simply vary the distance of the UAA reactive group from the protein's surface. By extending the distance from the protein, the handle is less hindered by the protein's steric bulk, which may increase its propensity to react with its orthogonal partner. Thus, the research also aimed to determine the ideal distance of the functional handle from the protein for efficient reactivity.

In order to probe novel bioorthogonal reactions utilizing UAA technology, a variety of UAAs with new functional groups were synthesized and designed with variable tether lengths to separate the reactive species from the bulk of the protein. These groups included an azide, an alkyne, and a bromide respectively, all of which extended from a derivatized tyrosine (see Scheme 1). These functionalities afford access to the Cu(I)-catalyzed 1,3-dipolar cycloaddition, the Glaser-Hay coupling, and a nucleophilic substitution respectively.

II.A. Synthesis of Unnatural Amino Acids

To optimize the immobilization process, UAAs were synthesized with various numbers of methylene units extending from a tyrosine precursor. Derivatives containing 2, 3, and 4 methylenes of the azide, alkyne, and bromide functionalities previously mentioned are readily accessible via an S_N2 reaction (Scheme 1).

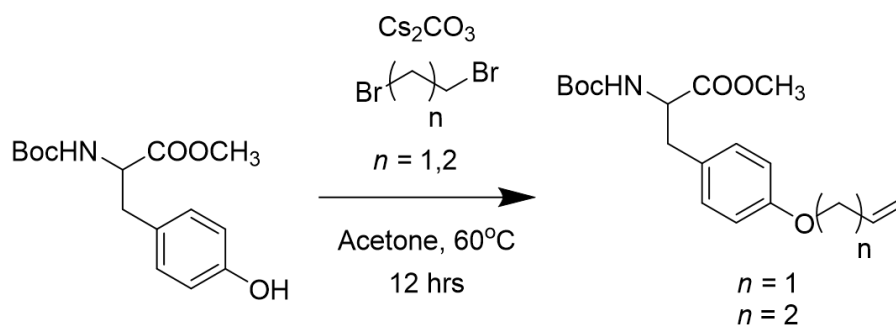
Scheme 5. Final Overall Reaction Scheme for the Synthesis of UAAs



Synthesis of the nine UAAs began from a common N-Boc/OMe protected tyrosine precursor. The goal was to synthesize a bromotyrosine derivative via a substitution reaction with either dibromoethane, dibromopropane, or dibromobutane, all of which are commercially available. These bromo-tyrosines would then have alkyl chains of varying lengths, and possess a bromine that could serve as a good leaving group in future reactions. Then, reaction with either an azide or alkynyl nucleophile would yield novel reactivity to the tyrosine derivatives. However, synthesis of the bromo-tyrosine series proved more difficult than initially anticipated.

Initial reactions between the protected tyrosine and the dibromobutane in dimethylformamide (DMF) yielded the desired tyrosine derivative **3**. However, reactions involving the dibromopropane and dibromoethane in DMF did not proceed as expected towards the desired tyrosine derivatives **2** and **1** respectively. Instead, starting material or the alkene tyrosine derivative, the product of an elimination reaction, was obtained for both reactions.

Scheme 6. Elimination Byproduct of the Bromotyrosine Synthesis

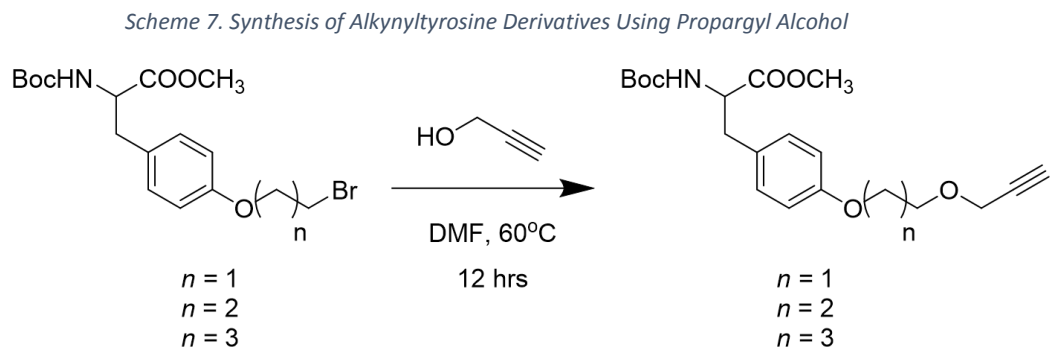


Initially, the concern was that the DMF was not dry, and the presence of water in the solvent may have led to elimination of the bromide. However, switching to dry solvent systems did not improve the reaction, nor did it reduce the amount of elimination product obtained. To optimize the reactions, we attempted adding a catalytic amount of potassium iodide to both the dibromopropane and dibromoethane reactions. The iodide can serve as a catalyst in this nucleophilic substitution, replacing the bromine with iodine, a better leaving group, which should facilitate the S_N2 addition by the hydroxyl group of tyrosine. However, only starting material was recovered from these reactions, even when reaction times and temperatures were varied. Ultimately, a literature search revealed a similar compound had been successfully prepared using acetone as the solvent system.^{37,38} Thus, all three reactions were attempted in acetone, with greater success. All three bromotyrosine derivatives were successfully synthesized using acetone as the solvent system in the presence of base and potassium iodide as a catalyst. We hypothesize that the additional polar amine functionality in the DMF allowed for increased solvent-mediated stability of the free bromine via dipole-dipole interactions following an elimination reaction. Thus, DMF aided in the formation of the eliminated product. Furthermore, electronics are clearly involved, as the shorter the alkyl chain became, the closer in proximity the large, electronegative bromine groups were, increasing the propensity to eliminate. This was

observed by a higher rate of elimination product using dibromoethane when compared to dibromopropane. The use of acetone, a less polar solvent, decreases the solvent-mediated dipole-dipole interactions with the free bromine, and thus disfavors the elimination reaction. This was observed as the bromotyrosine derivatives were able to be successfully synthesized under these modified conditions.

Once the bromotyrosine compounds **1** to **3** had been synthesized, they were then divided and reacted with either sodium azide or propargyl alcohol. Reaction with sodium azide in DMF resulted in a nucleophilic substitution reaction between the azide and the bromotyrosine, affording the azidotyrosine derivatives **4** through **6** following purification and NMR confirmation. The other set was reacted with propargyl alcohol in DMF in order to perform a nucleophilic substitution between the hydroxyl group and the bromotyrosine, resulting in the formation of the alkynyltyrosine product.

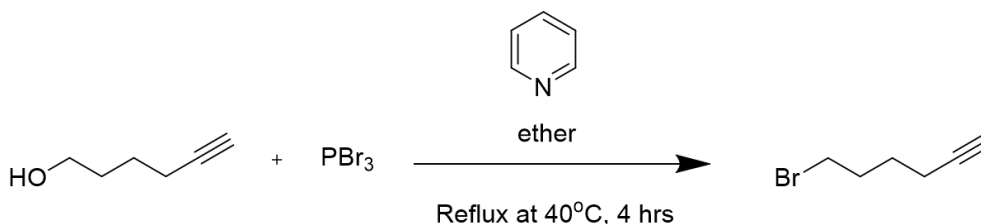
The alkyne-derivitized tyrosine series was not successfully synthesized under these conditions, and instead yielded either the previously described elimination product or the starting material.



Initially, potassium tert-butoxide (pK_a ~17) was used as the activating base in the reaction. We hypothesized that the propargyl alcohol (pK_a ~13.6) may not be fully deprotonated under these conditions, perhaps reducing the overall yield of the reaction.¹ As such, a small

amount of sodium hydroxide was used as the activating base with little success. Ultimately, due to the inability to obtain derivatized tyrosine product, an alternative synthesis strategy was adopted. Reaction of 4-bromo-1-butyne and 5-bromo-1-pentyne with protected tyrosine in acetone with a catalytic amount of potassium iodide was allowed to progress overnight. Following column chromatography, the desired alkyne products **7** and **8** were obtained, with the proper length alkyl chains (either 3 or 2 methyl) for proper comparison to the azido- and bromo- derivatives of similar lengths. In addition, 6-bromo-1-hexyne could be readily synthesized by reacting 6-hydroxy-1-hexyne with phosphorus (III) bromide, followed by immediate reaction with the protected tyrosine in acetone. Again, following column chromatography the desired compound **9** was isolated.

Scheme 8. Synthesis of 6-Bromo-1-hexyne



The synthesis described above results in the derivatized tyrosines still harboring protecting groups. To prevent reactivity of the amino group of the amino acid, an acid labile tert-butyloxy (Boc) group was employed. Additionally, to protect the carboxylic acid functional group of the amino acid backbone, a base labile methyl ester was used. Following successive stirring in 1M lithium hydroxide to remove the methyl ester and 2% TFA in DCM to remove the Boc group, the deprotected UAAs were concentrated under vacuum and analyzed by ¹H- and ¹³C NMR. Table 1 summarizes the overall deprotected UAA yields for this general reaction protocol.

Table 1. Overall Yields For the Synthesis of all Deprotected UAAs

X	n	% Yield
	1 (10)	5.3% (p2BrY)
	2 (11)	8.2% (p3BrY)
	3 (12)	10.8% (p4BrY)
	1 (13)	32.0% (p2AzY)
	2 (14)	58.0% (p3AzY)
	3 (15)	43.3% (p4AzY)
	1 (16)	2.6% (p2yneY)
	2 (17)	18.7% (p3yneY)
	3 (18)	3.1% (p4yneY)

II.B. Incorporation of Unnatural Amino Acids

Following the Schultz methodology for the incorporation of UAAs (see Chapter 1), the deprotected unnatural amino acids were then incorporated into green fluorescent protein (GFP) on the surface exposed residue 151, using transformed *E. coli*. GFP, in this case, was an ideal reporter protein, as successful incorporation of the UAAs could be monitored by fluorescence. Because the incorporation of UAAs is based on the suppression of a TAG (amber stop codon) mutation in the DNA encoding the protein, *E. coli* cultured in the absence of the UAA will terminate protein translation early, resulting in an incomplete protein and an absence of fluorescence (see Figure 7).

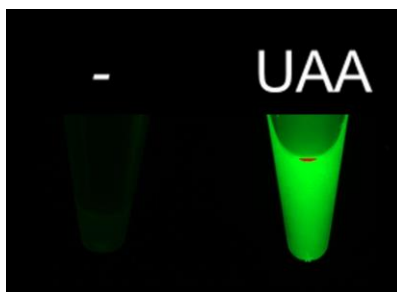


Figure 7. Successful suppression of the TAG stop codon results in translation of Green Fluorescent Protein (GFP), and the presence of fluorescence. In the absence of a UAA, GFP translation is terminated early, resulting in a protein that cannot fold correctly to produce fluorescence.

Furthermore, because residue 151 in the wt GFP is a tyrosine residue, we hypothesized that addition of tyrosine-derivatives **10** to **18** would not disrupt the overall folding of the protein. All UAAs were successfully incorporated into GFP using the pCNF synthetase. This synthetase was originally evolved for the para-cyanophenylalanine UAA, however, the double-sieve selection methodology discussed in chapter 1 was not pursued as rigorously. As a result, recent work has demonstrated that the pCNF synthetase is a highly promiscuous synthetase.³⁹ Recent work has demonstrated that this synthetase is capable of charging at least twenty UAAs with a tyrosine or phenylalanine core and with variable functional groups. Based on this work, it was hypothesized that this synthetase could recognize and charge compounds **10** to **18** to the TAG-tRNA for incorporation into GFP, obviating the need to do synthetase evolutions for each individual compound.

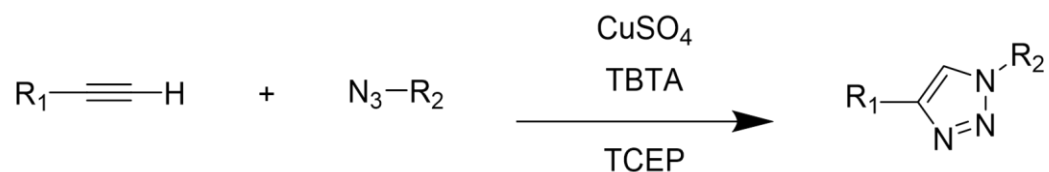
Because the pCNF synthetase was evolved to recognize pCNF, it was hypothesized that the p2AzY-, p3AzY-, and p4AzY-UAA series would be most easily incorporated due to the structural similarity between the cyano group and azide group in size and polarizability. This result was observed, with the p2AzY producing a larger quantity of mutant-GFP, and the p4AzY producing the lowest quantity of mutant-GFP within the azido series. These qualitative observations make sense based on the evolution of the synthetase. Because the synthetase was evolved for a UAA that contained a cyano in the para position

on the ring of phenylalanine, it makes sense that the increased tether distance may have more difficulty fitting into the binding pocket relative to smaller derivatives. The steric bulk of p4AzY may not allow it to fit as well into the binding pocket of the pCNF-synthetase, preventing the UAA from being charged to the TAG-tRNA, and leading to early termination of GFP translation (and thus a decrease in observed fluorescence following protein purification). Thus, a tradeoff may exist between the advantages of distancing the functional group from the core of the protein and the reduced protein yields of these same compounds. The bromotyrosine UAA series had the next highest protein yields, and followed a similar trend. The alkynyltyrosine UAA series had the lowest protein yields, with very little protein obtained throughout the series of derivatives. The enhanced yields of the bromo-series over the alkynyl-series is possibly due to similarity in size and shape between the azide groups and the bromo groups of the tyrosine derivatives, whereas the alkyne functionality was unable to fit as easily in the synthetase binding pocket. All proteins were purified via a Ni-NTA affinity chromatography due to the presence of a 6X histidine tag, and eluted in a solution with increased imidazole concentrations.

II.C. Bioorthogonal Chemistry Experiments

With the UAA-mutated GFP in hand, it was then possible to examine the capabilities of the various functional handles for bioorthogonal reactivity. Initial studies involved the p2AzY-, p3AzY-, p4AzY-GFP series. To test for bioorthogonal reactivity, we exposed the azido-GFPs to a fluorescent dye containing an alkynyl functional group (Alexafluor 488). We expected the formation of a triazole ring between the dye and the azido-GFP in the presence of the Cu(I)-1,3-dipolar cycloaddition (Copper “click”) conditions.

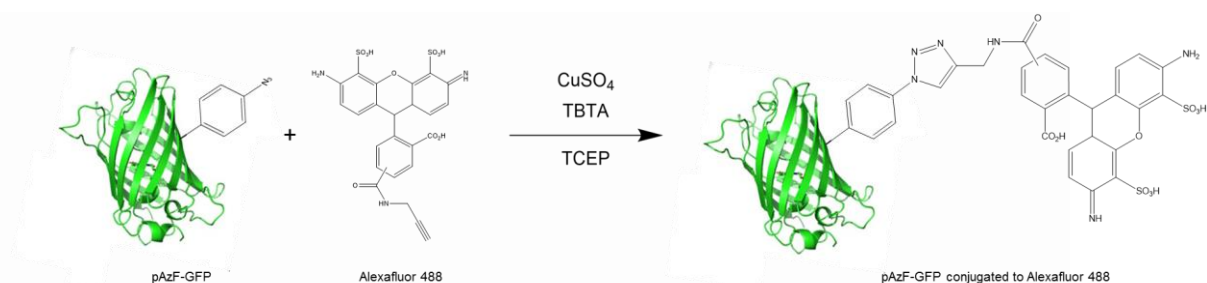
Scheme 9. General Reaction for the Cu(I)-1,3-Dipolar Cycloaddition (or Copper(I) "Click")



Thus, when the GFP was denatured, a fluorescent band due to the presence of the conjugated dye should still appear at the appropriate molecular weight band for GFP (~26 kDa) via SDS-PAGE in only those proteins exposed to the copper click catalyst (CuSO_4). Furthermore, because the dye is small, it should not impact the migration of the GFP protein on the gel.

Initial optimization trials for the copper "click" involved reactions between p-azidophenylalanine (pAzF), a commercially available UAA with no tether group between the aromatic ring and the azide moiety, and the fluorescent dye in the copper click conditions based on literature conditions were not successful, as indicated by the absence of fluorescence at 26 kDa on SDS-PAGE.^{40,41}

Scheme 10. Initial Click Trials Using pAzF-GFP and Alexafluor 488



We next attempted varying the CuSO_4 concentrations (either 1M, 50mM, or 1mM) to no avail. Furthermore, we attempted using either sodium ascorbate or tris(2-carboxyethyl)phosphine (TCEP) as reducing agents, at various concentrations as well. This did not produce positive results either. Finally, based on a literature precedence a chelating

ligand to help increase the stability and reactivity of the copper(I) species was added. Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA), which is a tetradentate ligand was selected to aid in the reaction.^{40,41} Initial click attempts in the presence of this ligand revealed promising results, however, none were replicable. Ultimately, it was realized that purification of proteins required transferring the protein from the imidazole buffer to phosphate buffer solution (PBS).⁴¹ Thus, solvent exchange of the pAzF-GFP to PBS via a concentrator column resulted in a successful click reaction between the pAzF-GFP and the alkyne-dye in the presence of CuSO₄ (50 mM), TCEP (50 mM), and TBTA (20 mM). We hypothesize that the presence of imidazole in the elution buffer must have coordinated with the copper (I) species, preventing it from coordinating with the alkyne functional group and catalyzing the formation of the triazole ring with the azido-handle of the mutant-GFP.

Figure 8 demonstrates the successful copper click reaction between p2AzY-, P3AzY-, and p4AzY-GFP and the alkyne-derivitized Alexafluor 488. All mutant proteins were subjected to an experimental trial in the presence of the copper click catalyst system (CuSO₄, TCEP, TBTA) and a control in presence of the alkyne fluorophore but in the absence of the copper click catalyst system. This control allowed us to rule out non-specific interactions as the mechanism for binding between the dye and the mutant-GFP. After running overnight at 4°C, the proteins were denatured at 98°C for 15 mins and then analyzed via SDS-PAGE. As indicated by the top gel, all protein expressions using the azido-UAA series were successful, as indicated by the presence of a band at 26 kDa, the molecular weight of GFP. Furthermore, the relative widths of the band confirm the visual observations discussed above: UAAs p2AzY and p3AzY had higher GFP yields than the bulkier p4AzY. The bottom lane demonstrates the successful bioconjugation between the

azido-GFPs and the alkyne-derivitized Alexafluor 488. Because the GFP is denatured, only the fluorophore will fluoresce, with fluorescence at a molecular weight of 26 kDa indicating that the fluorophore is conjugated to the GFP. Furthermore, no coupling was observed in the absence of the copper catalyst system, even in the presence of the fluorophore, indicating that bioorthogonal copper click chemistry is responsible for the conjugation between the fluorophore and the dye.

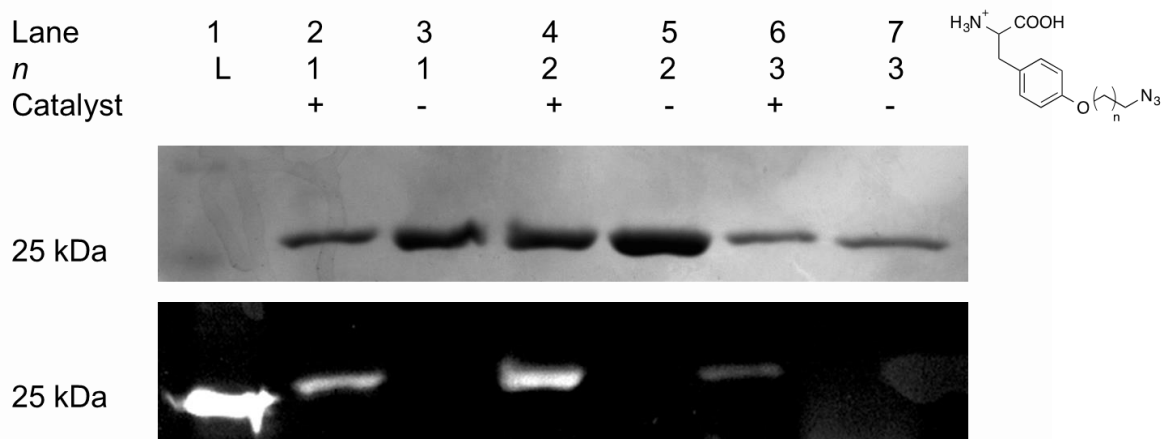


Figure 8. Expression of GFP-151 mutants with azide tethered tyrosine UAAs. The top gel indicates the successful incorporation of all three azides in the series as an appropriate band is observed for GFP at ~26 kDa when stained with coomassie blue. Each azido-GFP mutant also was able to undergo a successful bioconjugation with an AlexaFluor-488 alkyne. No coupling was observed in the absence of copper catalyst despite the inclusion of the fluorophore alkyne. This is apparent when observing fluorescent bands in the bottom gel.

Interestingly, the p3AzY handle appears to have the best conjugation efficiency, compared to the more sterically hindered p2AzY and the more flexible p4AzY handles.

Next, we attempted to perform the Glaser-Hay reaction between the p2yneY-, p3yneY-, p4yneY-GFP, and the alkyne-fluorescent dye.

Scheme 11. General Schema for the Glaser-Hay Coupling

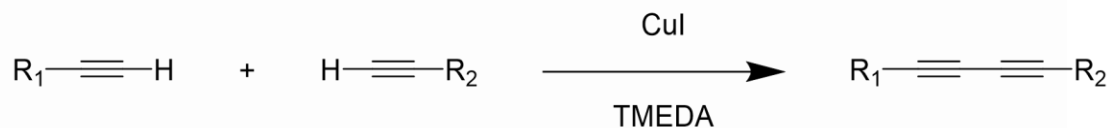


Figure 9 indicates that our initial attempt using CuI (50mM) and tetramethylethylenediamine (TMEDA, 50 mM) as a catalyst system was successful. All mutant proteins were subjected to an experimental trial in the presence of the Glaser-Hay catalyst system (CuI, TMEDA) and a control in presence of the alkyne fluorophore but in the absence of catalyst system. This control allowed us to rule out non-specific interactions as the mechanism for binding between the dye and the mutant-GFP. After running overnight at 4°C, the proteins were denatured at 98°C for 15 mins and then analyzed via SDS-PAGE. As indicated by the top gel, all protein expressions using the alkynyl-UAA series were successful, as indicated by the presence of a band at 26 kDa, the molecular weight of GFP. The bottom lane demonstrates the successful bioconjugation between the alkyne-GFPs and the alkyne-derivitized Alexafluor 488. Because the GFP is denatured, only the fluorophore will fluoresce, with fluorescence at a molecular weight of 26 kDa indicating that the fluorophore is conjugated to the GFP. Furthermore, no coupling was observed in the absence of the Glaser-Hay catalyst system, even in the presence of the fluorophore, indicating that bioorthogonal Glaser-Hay chemistry is responsible for the conjugation between the fluorophore and the dye. Even more interesting is the presence of a slight band at ~52 kDa in only the experimental samples (Lanes 2, 4, and 6), which suggests that a homodimerization between alkynyl-GFPs in the presence of the Glaser-Hay system occurred. Even more interesting is the relative intensities of this band, with the p4yneY-GFP have the darkest band, and thus the most homodimerization, and the p2yneY-GFP having the lightest band, and thus the least homodimerization. This would be expected, as the increased distance off of the protein afforded by p4yneY UAA would

reduce the steric hindrance of the protein, potentially increasing the propensity of this functional handle to react.

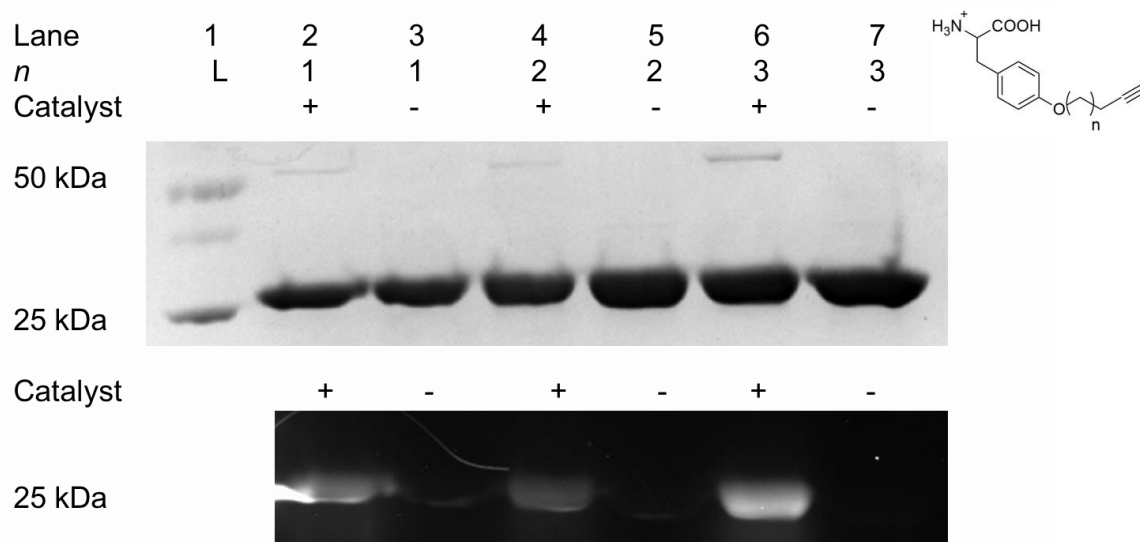


Figure 9. Expression of GFP-151 mutants with alkyne tethered tyrosine UAAs. The top gel indicates the successful incorporation of all three alkynes in the series as an appropriate band is observed for GFP at ~26 kDa when stained with coomassie blue. Each alkyne-GFP mutant also was able to undergo a successful bioconjugation with an AlexaFluor-488 alkyne. No coupling was observed in the absence of copper catalyst despite the inclusion of the fluorophore alkyne. This is apparent when observing fluorescent bands in the bottom gel. In addition, the presence of bands at ~52 kDa suggests homodimerization in the presence of copper catalyst as the alkyne-GFP mutants conjugate with themselves.

The final reaction we sought to perform on our mutant protein was a nucleophilic substitution between a nucleophilic dye and our bromo-mutated GFPs. Due to the presence of many nucleophilic functional groups within biological settings, this reaction is not truly bioorthogonal because cross-reactivity is likely. Nevertheless, we reasoned that successful nucleophilic substitution with a dye was an ideal means to demonstrate the successful incorporation of the bromo-UAAs and the generation of the p2BrY-, p3BrY-, p4BrY-GFPs. Furthermore, we wanted the reaction to still have some biological contextualization.

Glutathione (GSH) is a tripeptide composed of glutamine, cysteine, and glycine.¹ Thanks to the reactive nature of the cysteine's thiol group, GSH readily undergoes substitution chemistry *in vivo* as a means to modulate protein oxidation states. In solution, GSH can exist in one of two forms: the reduced glutathione (GSH) form, in which the thiol

group is protonated, or in the oxidized dimeric form (GSSG), in which the cysteine residues form a disulfide bond. *In vivo*, the reversible exchange of glutathione between these two forms allows the tripeptide to modulate protein oxidation states.

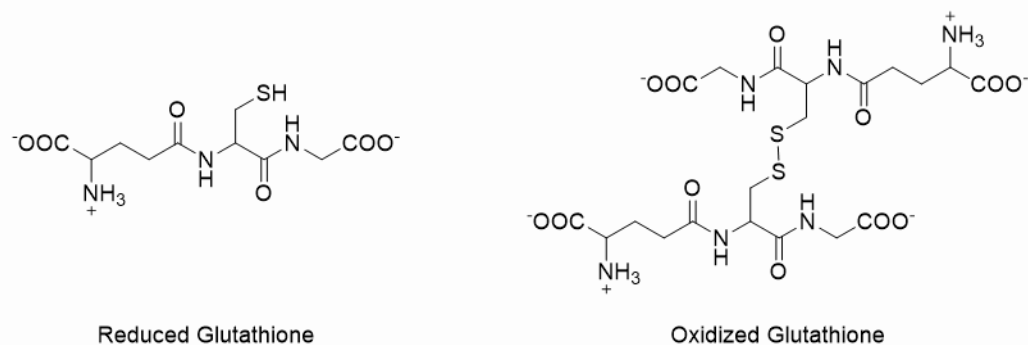


Figure 10. Glutathione (GSH) can exist in either a reduced or oxidized form, based on the oxidation state of the cysteine's thiol group. Only the reduced form of GSH can undergo a nucleophilic substitution reaction with our bromo-GFP mutants.

Thankfully, the Landino lab at William and Mary has already established a precedent for generating GSH-labeled dyes.⁴² By reacting fluorescein or dansyl dyes with the free amino end of GSH, the Landino group has been able to generate labeled-GSHs that are still capable of performing redox chemistry with proteins of interest. Thus, we reasoned that the dansyl-labeled glutathione (d-GSH, see Figure 5), would be an ideal label to undergo a nucleophilic substitution reaction with our bromo-GFP series.

Scheme 12. General Reaction Schema for the Nucleophilic Substitution of Bromo-GFP by a Reactive Thiol Group



We obtained d-GSH (~3 mM) from the Landino lab as a nucleophile in reaction with our bromo-GFP in PBS at 4°C overnight. However, this reaction did not work. We hypothesized that the d-GSH could have oxidized, and thus no free thiol group was free to serve as nucleophile to displace the bromide. As a solution, we attempted the reaction, this time in the presence of 20mM imidazole to serve as a mild base to help activate the thiol,

and 50 mM TCEP as a reducing agent to reduce any oxidized d-GSH. This reaction was performed at 4°C overnight. However, the reaction results were inconclusive, as a weak fluorescence at the band associated with denatured GFP was observed, but it was also somewhat present in the wild-type control, as shown in Figure 3. We hypothesize that due to the lower affinity of the synthetase for the bromotyrosine UAAs, the GFP yield is considerably reduced. Furthermore, the d-GSH may have been oxidized due to improper handling, further reducing the amount of d-GSH that may have reacted. These two factors combined can lead to a considerably decreased yield of d-GSH-labeled GFP, which would increase our signal to noise ratio on the SDS-PAGE making detection difficult. Future studies will attempt this substitution again, this time with more concentrated p2BrY-, p3BrY-, and p4BrY-GFPs and a freshly prepared stock of d-GSH.

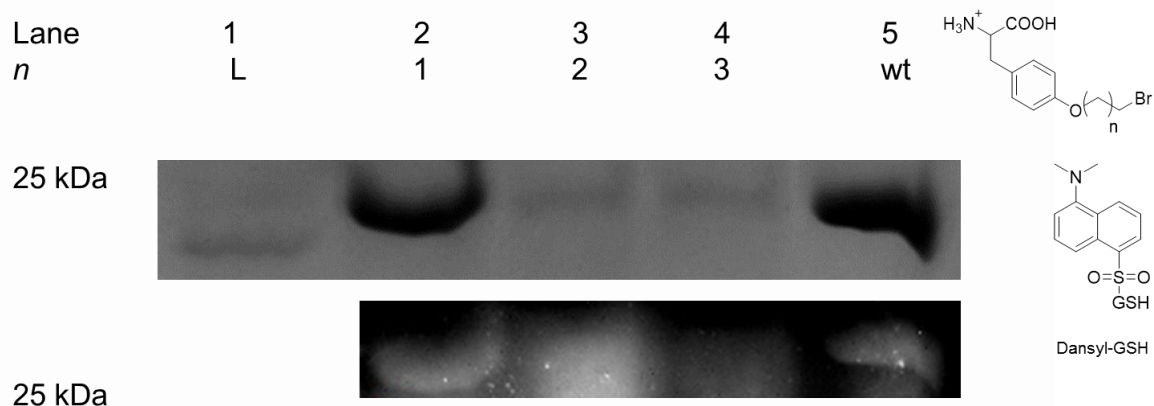


Figure 11. Expression of GFP-151 mutants with bromo tethered tyrosine UAAs. The top gel indicates the successful incorporation of all three bromotyrosines in the series as an appropriate band is observed for GFP at ~26 kDa when stained with coomassie blue. The bromo-GFP mutants were able to undergo a nucleophilic substitution reaction with dansyl-labeled glutathione to varying degrees.

II.D. Bioorthogonal Chemistry Applied Towards Protein Immobilization

With successful bioorthogonal reaction conditions in hand, we next sought to apply the chemistry for immobilization purposes. As mentioned previously, the most robust and ideal mechanisms for protein immobilization involve covalent linkage. However, this

technique is often confounded by the potential for multiple reaction sites on a given protein, which could render proteins that are oriented incorrectly for their function. UAA mutagenesis allows for the development of a bioorthogonal handle in a predetermined site.⁴³ This predefined functional group can then be exploited via bioorthogonal chemistries for the generation of a covalent linkage between the UAA-mutant protein and the solid support at a predefined site.³⁵ This strategy prevents the complications of having multiple reaction sites as well as improperly oriented proteins, as the number and positions of UAAs can be controlled by the scientist. As such, we sought to apply UAA-based technologies to generate covalent linkages between UAA-mutated GFP and a solid support in a site-specific manner.

For reaction with the azido- and alkynyl-GFP mutants, an epoxy-activated sepharose resin was derivatized with propargyl alcohol in the presence of trimethylamine (TEA) by stirring overnight at room temperature.³⁵ This afforded a resin with an alkyne functionality, which was confirmed via the successful coupling of the resin with the alkyne-dye under Glaser-Hay coupling conditions.

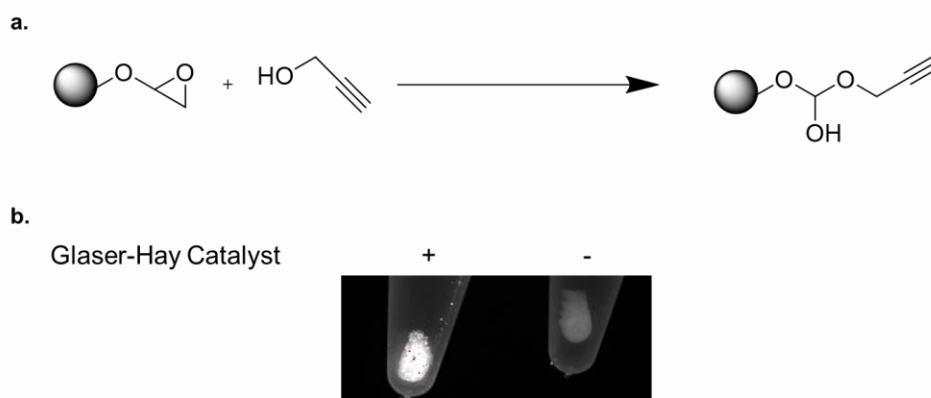


Figure 12. A. The reaction scheme for the formation of an alkyne derivatized resin. B. An alkyne containing dye was immobilized on the alkyne resin via the Glaser-Hay coupling, demonstrating that the reaction performed in A. was successful.

Next, we reacted the alkyne-derivatized resin to the various azido-tyrosine mutated GFPs under the previously established click conditions. We reasoned that by extending the tethering distance of the functional handle of the UAA, and thus the site of bioorthogonal reactivity, we could decrease steric hindrance due to the protein's bulk and improve immobilization efficiency as measured by an increase in covalently-bound GFP on the resin. As such, the azidoethyl-, azidopropyl-, and azidobutyl-mutated GFPs were exposed to copper click conditions (50 mM CuSO₄, 50 mM TCEP, and 20 mM TBTA) at fixed concentration of GFP (roughly 0.250 mg/mL). As a positive control, the p-azidophenylalanine mutated GFP, which has no tether between the aromatic ring and the azide functional group, was also subjected to the same copper click conditions. Our lab has previously immobilized the p-azidophenylalanine mutated GFP on the alkyne-functionalized sepharose resin using similar conditions.³⁵ As a negative control, wild type GFP, in which there is no azido-functional group present, was also subjected to the same copper click conditions. This control provided a means to rule out non-specific interactions as the reason for immobilization rather than the selective copper click reaction. Figure 13. demonstrates that our copper click-based immobilization strategy was effective. Interestingly, the tether distance did factor into the efficiency of the immobilization reaction, as increased distance from the protein did not necessarily lead to improved immobilization (as measured by relative fluorescence intensity on a microplate reader).

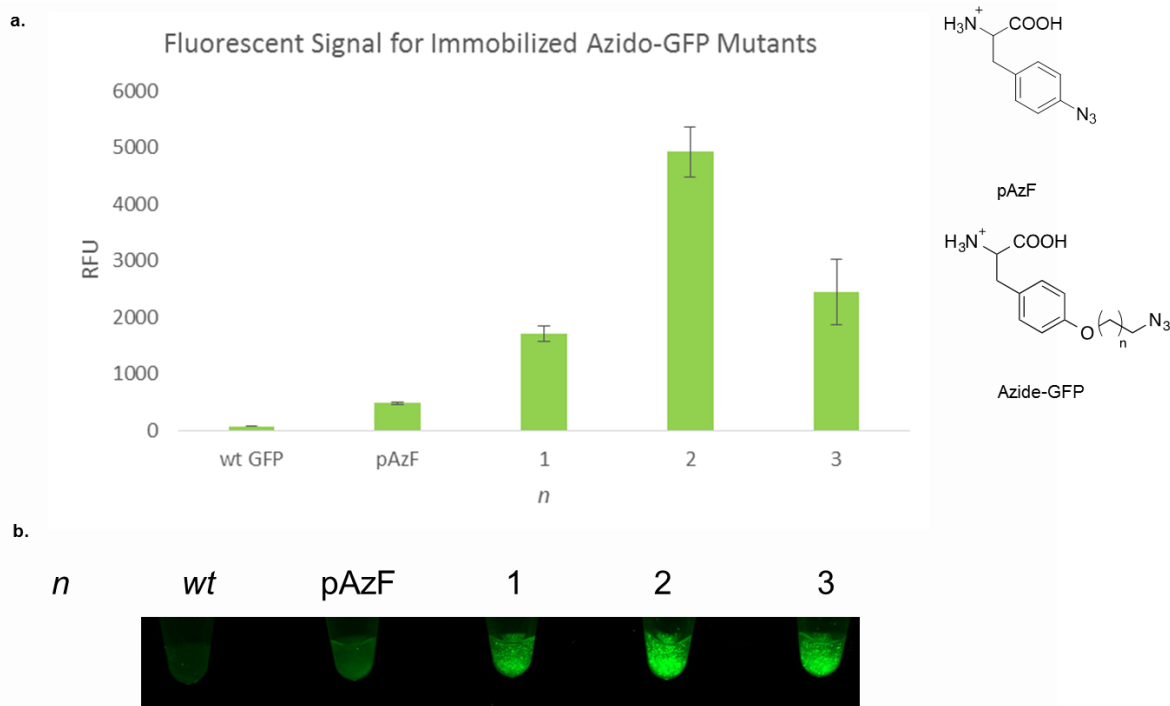


Figure 13. Expression of GFP-151 mutants with either an azide tethered tyrosine UAAs or a *p*-azidophenylalanine. **a.** The graph demonstrates the successful immobilization of GFP at a fixed concentration of 0.246mg/mL of protein. Fluorescence signal read at 528nm indicates that the *n*=2 GFP-151 azide tethered variant performed best for immobilization. **b.** The bottom image displays the immobilized resins as viewed on a 368nm plate reader. Overall, the tethered variants appeared to have a greater immobilization efficiency, suggesting improved access to the functional handle thanks to the greater distance from the protein.

Next we sought to immobilize the alkyne-functionalized GFP on the same alkyne-functionalized sepharose resin using the Glaser-Hay coupling conditions. However, this strategy did not prove as effective. Due to the reduced yields of the alkyl-tyrosine mutant GFP series, we could not obtain sufficient protein concentrations in order to obtain observable immobilization results. Current work is focusing on resynthesizing these alkyl-tyrosine UAAs for larger protein expressions, which will hopefully yield greater quantities of protein for future immobilization studies.

Current work is also attempting to generate an amine functionalized resin using the epoxy-activated sepharose resin and ethylenediamine. Once made, we plan on using the amine group of this resin as a nucleophile in the substitution reaction of the bromo-tyrosine mutated GFP series for immobilization of these mutant GFPs as well.

III. EXPERIMENTAL

Synthesis of *p*-Bromo-ethyl-*Boc*-*OMe*-Tyrosine (1). A solution of *Boc*-Tyrosine-*OMe* (0.500 g, 1 eq, 1.693mmol) was created in acetone (5 mL) in a flame-dried vial. To this solution, potassium carbonate (0.702 g, 3 eq, 5.079 mmol) was added, followed by the addition of dibromoethane (292 μ L, 2 eq, 3.383 mmol). The reaction stirred overnight at 56°C. It was then cooled to room temperature, filtered into a roundbottom and the precipitate in the vial washed with ethyl acetate. The filtrate was then rotovapped. Column chromatography (silica gel, gradient of 9:1 hexanes/ethyl acetate for 5 fractions followed by 7:1 hexanes/ethyl acetate for 10 fraction followed by 5:1 hexanes/ethyl acetate) afforded the desired product as a clear oil (36 mg, 0.089 mmol, 5.3% yield). ^1H NMR (400 MHz, CDCl_3): δ 7.02 (d, J = 8 Hz, 2H), 6.82 (d, J = 8 Hz, 2H), 4.96 (d, J = 8 Hz, 1H), 4.53 (d, J = 8 Hz, 1H), 4.01 (t, J = 4 Hz, 2 H), 3.70 (s, 3 H), 3.51 (t, J = 8 Hz, 2 H), 2.94 (m, 2 H), 1.41 (s, 9 H)

Synthesis of *p*-Bromo-propyl-*Boc*-*OMe*-Tyrosine (2). A solution of *Boc*-Tyrosine-*OMe* (0.500 g, 1 eq, 1.693mmol) was created in acetone (5 mL) in a flame-dried vial. To this solution, potassium carbonate (0.702 g, 3 eq, 5.079 mmol) was added, followed by the addition of 1,3-dibromopropane (345 μ L, 2 eq, 3.383 mmol). The reaction stirred overnight at 56°C. It was then cooled to room temperature, filtered into a roundbottom and the precipitate in the vial washed with ethyl acetate. The filtrate was then rotovapped. Column chromatography (silica gel, gradient of 7:10 hexanes/ethyl acetate for 10 fractions followed by 5:1 hexanes ethyl acetate for 10 fractions) afforded the desired product as a clear oil (18 mg, 0.283 mmol, 33.5% yield). ^1H NMR (400 MHz, CD_3OD): δ 6.99 (d, J = 8 Hz, 2 H), 6.78 (d, J = 8 Hz, 2 H), 5.07 (d, J = 4 Hz, 1 H), 4.59 (d, J = 4 Hz, 1 H), 4.01 (t, J = 4 Hz, 2

H), 3.65 (s, 3 H), 3.53 (t, J = 4 Hz, 2 H), 2.97 (m, 2 H), 2.24 (quintet, J = 4 Hz, 2 H), 1.37 (s, 9 H)

Synthesis of *p*-Bromo-butyl-*Boc*-*OMe*-Tyrosine (3). A solution of *Boc*-Tyrosine-*OMe* (0.500 g, 1 eq, 1.693 mmol) was created in acetone (5 mL) in a flame-dried vial. To this solution, potassium carbonate (0.702 g, 3 eq, 5.079 mmol) was added, followed by the addition of 1,4-dibromobutane (369 μ L, 2 eq, 3.383 mmol). The reaction stirred overnight at 56°C. It was then cooled to room temperature, filtered into a roundbottom and the precipitate in the vial washed with ethyl acetate. The filtrate was then rotovapped. Column chromatography (silica gel, gradient of 7:1 hexanes/ethyl acetate for 10 fractions followed by 5:1 hexanes/ethyl acetate) afforded the desired product as a white crystal (342 mg, 0.795 mmol, 46.9% yield). ^1H NMR (400 MHz, CDCl_3): δ 7.01 (d, J = 8 Hz, 2 H), 6.79 (d, J = 8 Hz, 2 H), 4.97 (d, J = 8 Hz, 1 H), 4.52 (d, J = 8 Hz, 1 H), 3.95 (t, J = 4 Hz, 2 H), 3.69 (s, 3 H), 3.47 (t, J = 4 Hz, 2 H), 3.01 (m, 2 H), 2.04 (m, 2 H), 1.92 (m, 2 H), 1.40 (s, 9 H)

Synthesis of *p*-Azido-propyl-*Boc*-*OMe*-Tyrosine (5). Dry dimethylformamide (DMF, 5 mL) was added to a vial containing protected-bromo-propyl tyrosine (synthesized as indicated above, 60 mg, 1 eq, 137 μ mol) to create a solution containing the bromo-tyrosine. This bromo-tyrosine solution was then transferred to a flame dried vial, into which was added sodium azide (11 mg, 1.2 eq, 165 μ mol). The reaction stirred overnight at 56°C. The organic layer was mixed with brine and then extracted with methylene chloride (DCM, 10 mL) and three times. This organic layer was then back extracted with brine (10 mL). This organic layer was then dried with magnesium sulfate, filtered, and rotovapped. Column chromatography (silica gel, 1:1 hexanes ethyl acetate) afforded the desired product as a crystal (32 mg, 0.085 mmol, 61.7% yield). ^1H NMR (400 MHz, CDCl_3): δ 7.03 (d, J

= 12 Hz, 2 H), 6.82 (d, J = 12 Hz, 2 H), 4.96 (d, J = 8 Hz, 1 H), 4.53 (d, J = 8 Hz, 1 H), 4.01 (t, J = 4 Hz, 2 H), 3.71 (s, 3 H), 3.46 (t, J = 8 Hz, 2 H), 3.02 (m, 2H), 2.03 (quintet, J = 4 Hz, 2 H), 1.41 (s, 9 H). ¹³C (400 MHz, CDCl₃): δ 172.39, 157.71, 155.07, 130.31, 128.20, 114.51, 79.89, 64.42, 54.49, 52.19, 48.22, 37.44, 28.77, 28.28

Synthesis of *p*-Azido-butyl-*Boc*-*OMe*-Tyrosine (6). The protected bromo-butyl tyrosine (252 mg, 1 eq, 0.586 mmol) was dissolved in a solution of dimethylformamide (12 mL) in a flame-dried vial. Sodium azide (76 mg, 2 eq, 1.17 mmol) was then added to the solution. The reaction was stirred overnight at 65°C. The organic layer was mixed with brine and then extracted with methylene chloride (DCM, 10 mL) and three times. This organic layer was then back extracted with brine (10 mL). This organic layer was then dried with magnesium sulfate, filtered, and rotovapped. Column chromatography (silica gel, 3:1 hexanes ethyl acetate) afforded the desired product (132 mg, 0.349 mmol, 57% yield). ¹H NMR (400 MHz, CDCl₃): δ 7.03 (d, J = 8 Hz, 2 H), 6.84 (d, J = 8 Hz, 2 H), 4.97 (d, J = 8 Hz, 1 H), 4.54 (d, J = 8 Hz, 1 H), 3.95 (t, J = 8 Hz, 2 H), 3.72 (s, 3 H), 3.36 (t, J = 8 Hz, 2 H), 3.05 (m, 2 H), 1.80 (m, 4 H), 1.43 (s, 9 H). ¹³C (400 MHz, CDCl₃): δ 172.40, 157.90, 130.28, 114.47, 67.07, 54.51, 52.18, 51.15, 37.43, 28.28, 26.48, 25.74

Synthesis of *p*-Alkynyl-ethyl-*Boc*-*OMe*-Tyrosine (7). To a flame dried vial, 5 mL of DMF and a stir bar were added. Boc-Tyrosine-OMe (295 mg, 1 eq, 1.000 mmol) was added without stirring followed by the slow addition of potassium tert-butoxide (337 mg, 3 eq, 3.000 mmol). The reaction immediately turned dark green. This mixture was allowed to stir at 100°C for 10 mins, then 4-Bromo-1-butyne was added (94 μL, 133 mg, 1 eq, 1.000 mmol). The mixture immediately turned a lighter, more brownish green. The mixture was then capped and allowed to stir at 80 °C overnight. Following reaction, 10 mL of brine was

then added to the reaction followed by the addition of 10 mL of ether. The aqueous layer was then drained and the ether layer was washed with brine (10 mL portions, 3 times). The ether layer was collected and then the aqueous layer was back extracted with 10 mL of ether. The ether layer was then dried and rotovapped down. Addition and subsequent removal of DCM resulted in the formation of precipitate that was the desired product.

Synthesis of *p*-Alkynyl-propyl-*Boc*-*OMe*-Tyrosine (8). Boc-Tyrosine-OMe (114 mg, 2 eq, 0.3855 mmol) was added to a flame-dried vial. Cesium carbonate (254 mg, 3 eq, 0.5783 mmol) was then added, and the solids were dissolved in dry dimethylformamide (3 mL). This mixture was then stirred at 100°C for 20 mins to activate the tyrosine starting material. 5-Bromo-1-pentyne (20 μ L, 1 eq, 0.1927 mmol) was then added to the mixture, as well as a spatula tip of potassium iodide. The reaction was then stirred overnight at 100°C. The reaction was then cooled to room temperature and washed with brine (10 mL) and ether (10 mL). The ether layer was then washed three times with 10 mL portions of brine. The brine layer was then backextracted with 10 mL of ether. The organic layers were combined, dried with magnesium sulfate, filtered, and rotovapped. Column chromatography (silica gel, 5:1 hexanes/ethyl acetate) to yield the desired product as a white crystal (3 mg, 0.008 mmol, 4.3% yield). ^1H NMR (400 MHz, CDCl_3): δ 7.02 (d, $J = 12$ Hz, 2 H), 6.82 (d, $J = 12$ Hz, 2 H), 4.95 (d, $J = 8$ Hz, 1 H), 4.53 (d, $J = 8$ Hz, 1 H), 4.03 (t, $J = 4$ Hz, 2 H), 3.71 (s, 3 H), 3.02 (m, $J = 8$ Hz, 1 H), 2.39 (t, $J = 4$ Hz, 2 H), 1.97 (m, $J = 8$ Hz, 2 H), 1.55 (s, 1 H), 1.41 (s, 9 H). ^{13}C (400 MHz, CDCl_3): δ 172.42, 157.93, 130.26, 127.94, 114.54, 83.46, 79.89, 68.84, 66.05, 54.50, 52.18, 37.43, 28.28, 28.16, 21.06, 15.15

Synthesis of *p*-Alkynyl-butyl-*Boc*-*OMe*-Tyrosine (9). Add 10 mL of ether to a flame-dried round bottom flask. To this, add pyridine (18 μ L, 16.2 mg, 0.21 mmol) and 6-hexyne-

1-ol (342 μ L, 300 mg, 3 eq, 3.06 mmol). A reflux bath was set up at 40°C. Phosphorous (III) bromide (97 μ L, 276 mg, 1 eq, 1.02 mmol) was then slowly added through the reflux condenser, producing a white gas and precipitate in the process. The reaction was then refluxed for 2 hours. The reaction was then cooled in an ice bath and diluted with 25 mL of ether. The organic mixture was then washed twice with a solution of saturated sodium bicarbonate (2 portions of 20 mL) and a single wash with brine (20 mL). The organic layer was then quickly condensed, to yield a clear, pungent oil that was the 6-Bromo-hexyne product (295 mg).

A solution of Boc-Tyrosine-OMe (1.089g, 2 eq, 3.69 mmol) in acetone (1 mL) was then made in a flame-dried vial. The 6-Bromohexyne product (295 mg, 1 eq, 1.84 mmol) was then dissolved in 4 mL of acetone, and all of this was transferred to the vial containing the tyrosine solution. To this mixture, cesium carbonate was added (1.802g, 3 eq, 5.53 mmol). The reaction was then stirred overnight at 60°C. The reaction was then cooled to room temperature and filtered into a new round bottom. The precipitate was washed with ethyl acetate and the collected filtrate was condensed. Column chromatography (silica gel, gradient of 7:1 hexanes/ethyl acetate for 10 fractions, 5:1 hexanes/ethyl acetate for 20 fractions, and 1:1 hexanes/ethyl acetate for 10 fractions) afforded the desired product as a white crystal (62 mg, 0.165 mmol, 9.0% yield). ^1H NMR (400 MHz, CDCl_3): δ 7.01 (d, J = 8 Hz, 2 H), 6.81 (d, J = 8 Hz, 2 H), 4.95 (d, J = 8 Hz, 1 H), 4.53 (d, J = 8 Hz, 1 H), 3.95 (t, J = 8 Hz, 2 H), 3.70 (s, 3 H), 3.02 (m, 2 H), 2.27 (td, J = 8 Hz, 4 Hz, 2 H), 1.96 (t, J = 4 Hz, 1 H), 1.90 (m, 2 H), 1.71 (m, 2 H), 1.41 (s, 9 H). ^{13}C (400 MHz, CDCl_3): δ 172.42, 158.04, 130.74, 127.81, 114.49, 84.07, 68.62, 67.19, 52.17, 37.44, 28.29, 25.02, 18.14.

General Protocol for Removal of Methyl Protecting Group. To deprotect the N-terminus of the unnatural amino acids, 500 μ L of dioxane and 500 μ L of 1M lithium hydroxide were added to a vial containing the unnatural amino acid on ice. The vial was then stirred at room temperature for 2 hours. The dioxane was evaporated off and 6M hydrochloric acid was added dropwise until a pH of 4 was obtained. The product was then extracted with ethyl acetate into a vial, and the organic layer was dried with magnesium sulfate and filtered into a new vial. The filtrate was then rotovapped.

General Protocol for Removal of tert-Butyloxy Protecting Group. To the methyl-free unnatural amino acid, 1 mL of 2% trifluoroacetic acid (200 μ L trifluoroacetic acid into 10 mL of DCM) was added on ice to deprotect the C-terminus of the unnatural amino acid. The vial was stirred at room temperature for 1 hour, and the DCM was then rotovapped off, affording the final deprotected unnatural amino acid, typically in the form of a white crystal.

Synthesis of *p*-Bromoethyltyrosine (p2BrY, 10). According to the general protocol outlined above, **1** was deprotected to yield a white crystalline powder (26 mg, 0.090 mmol, ~100% yield). $^1\text{H NMR}$ (400 MHz, CD_3OD): δ 7.11 (d, $J = 8$ Hz, 2H), 6.82 (d, $J = 8$ Hz, 2H), 4.23 (t, $J = 4$ Hz, 2 H)

Synthesis of *p*-Bromopropyltyrosine (p3BrY, 11). According to the general protocol outlined above, **2** was deprotected to yield a white crystalline powder (21 mg, 0.069 mmol, 24.5% yield). $^1\text{H NMR}$ (400 MHz, CD_3OD): δ 7.10 (d, $J = 8$ Hz, 2 H), 6.81 (d, $J = 8$ Hz, 2 H), 3.57 (t, $J = 4$ Hz, 2 H), 2.23 (quintet, $J = 4$ Hz, 2 H).

Synthesis of *p*-Bromobutyltyrosine (p4BrY, 12). According to the general protocol outlined above, **3** was deprotected to yield a white crystalline powder (58 mg, 0.183 mmol,

23.0% yield). ¹H NMR (400 MHz, CD₃OD): δ 7.09 (d, J = 8 Hz, 2 H), 6.79 (d, J = 8 Hz, 2 H), 3.94 (t, J = 4 Hz, 2 H), 2.80 (t, J = 8 Hz, 1 H), 1.98 (m, 2 H), 1.86 (m, 2 H)

Synthesis of *p*-Azidopropyltyrosine (p3AzY, 14). According to the general protocol outlined above, **5** was deprotected to yield a white crystalline powder (21 mg, 0.080 mmol, 94.1% yield). ¹H NMR (400 MHz, CD₃OD): δ 7.10 (d, J = 8 Hz, 2 H), 6.81 (d, J = 8 Hz, 2 H), 4.90 (s, 1 H), 4.16 (t, J = 4 Hz, 2 H), 1.97 (m, J = 4 Hz, 2 H). ¹³C (400 MHz, CD₃OD): δ 172.48, 156.15, 128.37, 127.74, 112.49, 109.99, 62.82, 26.91, 25.71

Synthesis of *p*-Azidobutyltyrosine (p4AzY, 15). According to the general protocol outlined above, **6** was deprotected to yield a white crystalline powder (28 mg, 0.100 mmol, 30.0% yield). ¹H NMR (400 MHz, CD₃OD): δ 7.09 (d, J = 12 Hz, 2 H), 6.79 (d, J = 12 Hz, 2 H), 3.93 (t, J = 4 Hz, 2 H), 1.99 (m, 2 H), 1.87 (m, 2 H). ¹³C (400 MHz, CD₃OD): δ 172.44, 156.28, 154.78, 128.33, 127.51, 112.44, 77.53, 65.05, 31.13, 27.80, 26.06, 25.71

Synthesis of *p*-Alkynylethyltyrosine (p2yneY, 16). According to the general protocol outlined above, **7** was deprotected to yield a white crystalline powder (6 mg, 0.026 mmol).

Synthesis of *p*-Alkynylpropyltyrosine (p3yneY, 17). According to the general protocol outlined above, **8** was deprotected to yield a white crystalline powder (9 mg, 0.036 mmol, 18.7% yield). ¹H NMR (400 MHz, CD₃OD): δ 7.53 (d, J = 12 Hz, 2 H), 7.24 (d, J = 12 Hz, 2 H), 4.44 (t, J = 4 Hz, 2 H), 2.76 (td, J = 8 Hz, 4 Hz, 2 H), 2.63 (t, J = 4 Hz, 1 H), 2.34 (m, 2 H). ¹³C (400 MHz, CD₃OD): δ 172.44, 128.29, 112.43, 66.93, 64.26, 26.50, 25.67, 12.74

Synthesis of *p*-Alkynylbutyltyrosine (p4yneY, 18). According to the general protocol outlined above, **9** was deprotected to yield a white crystalline powder (15 mg, 0.057 mmol, 34.5% yield). ¹H NMR (400 MHz, CD₃OD): δ 7.09 (d, J = 8 Hz, 2 H), 6.79 (d, J = 8 Hz, 2 H), 3.93 (t, J = 4 Hz, 2 H), 2.21 (td, J = 8 Hz, 4 Hz, 2 H), 2.18 (t, J = 4 Hz, 2 H), 1.83 (m,

2 H), 1.64 (m, 2 H). ^{13}C (400 MHz, CD_3OD): δ 172.44, 156.34, 128.25, 127.38, 112.39, 81.69, 77.48, 66.74, 65.37, 26.34, 25.66, 23.31, 15.78

General GFP Expression. A pET-GFP-TAG variant plasmid (0.5 μL) was cotransformed with a pEVOL-pAzF plasmid (0.5 μL) into *Escherichia coli* BL21(DE3) cells using an Eppendorf eporator electorporator. The cells were then plated and grown on LB agar in the presence of chloramphenicol (34 mg/mL) and ampicillin (50 mg/mL) at 37° C overnight. One colony was then used to inoculate LB media (4 mL) containing both ampicillin and chloramphenicol. The culture was incubated at 37° C overnight and used to inoculate an expression culture (10 mL LB media, 50 mg/mL Amp, 34 mg/mL Chlr) at an OD600 0.1. The cultures were incubated at 37° C to an OD600 between 0.6 and 0.8 at 600 nm, and protein expression was induced by addition of pAzF (100 A L, 100 mM) and 20 % arabinose (10 A L) and 0.8 mM isopropyl β -D-1-thiogalactopyranoside (IPTG; 10 A L). The cultures were allowed to shake at 30° C for 16-20 h then centrifuged at 5,000 rpm for 10 minutes and stored at -80° C for 3 hours. The cell pellet was re-suspended using 500 μL of Bugbuster (Novagen) containing lysozyme, and incubated at 37° C for 20 minutes. The solution was transferred to an Eppendorf tube and centrifuged at 15,000 rpm for 10 minutes, then the supernatant was poured into an equilibrated His- pur Ni-NTA spin (Qiagen) column with of nickel resin (200 μL) and GFP was purified according to manufacturer's protocol. Purified GFP was analyzed by SDS-PAGE (BioRad 10% precast gels, 150V, 1.5h), and employed without further purification. Protein concentrations were determined both by a BCA assay and fluorescence measurements.

Biological Cu(I)-1,3-dipolar Cycloaddition Reaction. Clicks on azido-tyrosine GFP-mutants were performed in 1.5mL Eppendorf tubes. To the experimental tube, in order, the

following was added: 2 μL of 50mM CuSO_4 as the copper(II) source, 20 μL of 5mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) as the catalyst ligand, 5 μL of azido-tyrosine mutated GFP (with either an ethyl, propyl, or butyl tether) in PBS, 10 μL of 488-Alexafluorophore (alkyne functionalized), 50 mM of tris(2-carboxyethyl)phosphine (TCEP) as the reducing agent, and 11 μL of phosphate buffer solution (PBS, pH = 7.4) to bring to a 50 μL volume. Control reactions were performed with 20 μL azido-tyrosine mutated GFP samples (either ethyl, propyl, or butyl) and 10 μL 488-Alexafluorophore in the absence of the Cu-1,3-dipolar cycloaddition catalysts (CuSO_4 , TBTA, and TCEP) to ensure that successful reactions were not the result of non-specific interactions between the protein and fluorophore. The control reaction was brought to 50 μL volume using PBS. All mixtures were reacted for 10 hours at 4°C. The whole sample volumes were then transferred to concentrator column (10k MWCO, Corning Spin-X) pre wet with PBS and spun at maximum speed in a tabletop centrifuge for 2 minutes. They were then washed with 100 μL portions of PBS and spun for another 2 minutes. This process was repeated for a total of seven washes, and the final spin time was adjusted to bring the samples to a concentration of ~ 25 μL (as indicated by the marking on the column).

Biological Glaser-Hay Coupling. Glaser-Hay couplings on alkyne-tyrosine mutated GFP were performed in 1.5mL Eppendorf tubes. To the experimental tube, the following was added in order: 3 μL of 50mM CuI as the catalyst, 3 μL 50 mM trimethylethylenediamine (TMEDA) as the catalyst ligand, 10 μL of 488-Alexafluorophore (alkyne functionalized), 20 μL of alkyne-tyrosine mutated GFP (with either ethyl, propyl, or butyl tether) in PBS for a total reaction volume of 36 μL . Control reactions were performed with 6 μL PBS, 10 μL 488-Alexafluorophore (alkyne), and 20 μL of alkyne-tyrosine mutated GFP (with either

ethyl, propyl, or butyl tether) to ensure that successful reactions were not the result of non-specific interactions between the protein and fluorophore. All mixtures were reacted for 4 hours at 4°C. The whole sample volumes were then transferred to concentrator column (10k MWCO, Corning Spin-X) pre wet with PBS and spun at maximum speed in a tabletop centrifuge for 2 minutes. They were then washed with 100 μ L portions of PBS and spun for another 2 minutes. This process was repeated for a total of seven washes, and the final spin time was adjusted to bring the samples to a concentration of \sim 25 μ L (as indicated by the marking on the column).

Nucleophilic Substitution of Bromo-Tyrosine Derivatives. Nucleophilic substitutions using dansyl-labeled glutathione (d-GSH, courtesy of the Landino lab) were performed in 1.5 mL Eppendorf tubes. To the experimental tube, the following was added in order: 20 μ L of bromo-tyrosine mutated GFP (with either an ethyl, propyl, or butyl tether) in 20 mM imidazole wash buffer, 10 μ L of d-GSH which provided a reactive thiol for nucleophilic attack, 5 μ L of 50 mM TCEP to help reduce any oxidized d-GSH, 5 μ L of 250 mM imidazole elution buffer to help activate cysteine residues in the d-GSH, and 10 μ L of PBS to bring the reaction to a 50 μ L volume. Control reactions were performed on wild-type GFP using the same conditions above. All reactions were vortexed and then reacted for 24 hours at 4°C. The whole sample volumes were then transferred to concentrator column (10k MWCO, Corning Spin-X) pre wet with PBS and spun at maximum speed in a tabletop centrifuge for 2 minutes. They were then washed with 200 μ L portions of PBS and spun for another 2 minutes. This process was repeated for another wash, and the final spin time was adjusted to bring the samples to a concentration of \sim 25 μ L (as indicated by the marking on the column).

Immobilization of Alkynes onto Epoxy Sepharose 6B Resin. Epoxy-activated 6B Sepharose (GE Healthcare, 200 mg) was added to a filter syringe and washed with distilled water (5 washes, 3 mL). Alkyn-ol (700 μ mol) and coupling buffer (3.5 mL, pH 13.0) was added to a 15 mL tube followed by the resin. The mixture was shaken at room temperature for 16 h. The resin was transferred to a filter syringe and washed with coupling buffer (4mL). The sepharose was transferred to a 15 mL tube and capped with ethanolamine (3.5 mL). The resin was incubated at 30 °C for four hours then washed in a filter syringe with acetate buffer (10 mM) and tris-HCl buffer (pH 4 and pH 8 respectively, 3 alternating washes with 3 mL each).

Immobilization of Azido-Tyrosine Mutated GFP Series. All azido-tyrosine GFP mutants were brought to a final concentration of 0.250 mg/mL using PBS. To a 1.5 mL Eppendorf tube, 15 μ L of each azido-tyrosine GFP mutant was added (with either an ethyl, propyl, or butyl tether). To this, the Cu(I)-1,3-dipolar cycloaddition reaction conditions were then added. In the following order, 5 μ L of 4 mM TCEP, 6 μ L 1 mM CuSO₄, 15 μ L PBS, and 20 μ L 5 mM TBTA were then added to the mix. This liquid mixture was then vortexed and transferred to a 1.5 mL Eppendorf that contained 15 mg preweighed alkyne-derivatized sepharose resin. A positive control consisting of the conditions above using p-azidophenylalanine, which was previously demonstrated to immobilize on the alkyne-derivatized sepharose resin, was set up.³⁵ A negative control consisting of the condition above using wild-type GFP, which should not react with the alkyne-derivatized sepharose resin due to the absence of a reactive azide moiety, was set up. All samples were allowed to react overnight at 4°C. These were then transferred to spin columns by adding 2, 200 μ L portions of PBS to the reaction tube to aid in removal. The reaction was then flowed

through at 1.1 rcf on a table-top microcentrifuge for 1 minute, and washed two times with 100 μ L portions at 1.1 rcf for 1 minute each. The protein-immobilized sepharose resin was then transferred to a clear Eppendorf tube for visualization and fluorescence quantification.

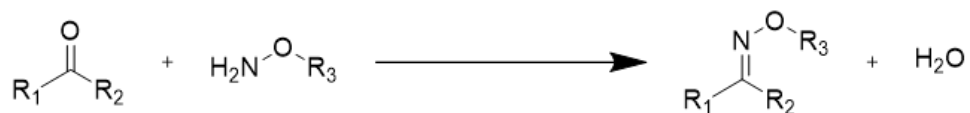
IV. FUTURE WORK

IV. A. Condensation Chemistry

Ketones and aldehydes represent another functional group that biological systems do not inherently possess, and can therefore be manipulated for bioorthogonal chemistry. In particular, condensation reactions involving these groups are ideal, as they can occur in aqueous environments and under physiological pH and temperature ranges.¹⁷

The oxime ligation represents a condensation reaction involving a ketone and a hydroxylamine group. Jencks demonstrated that at neutral pH, the nitrogen attacks the carbonyl compound in a rapid fashion, with the rate-limiting step involving the acid-catalyzed dehydration of the tetrahedral carbon intermediate, releasing a water molecule and forming a carbon-nitrogen double bond.⁴⁴ When the resulting bond ($C=N^1-X^2$) has an O in the X^2 position (as is the case when the nitrogen base is hydroxylamine), the bond formed is called an oxime.

Scheme 13. General Scheme of the Oxime Ligation



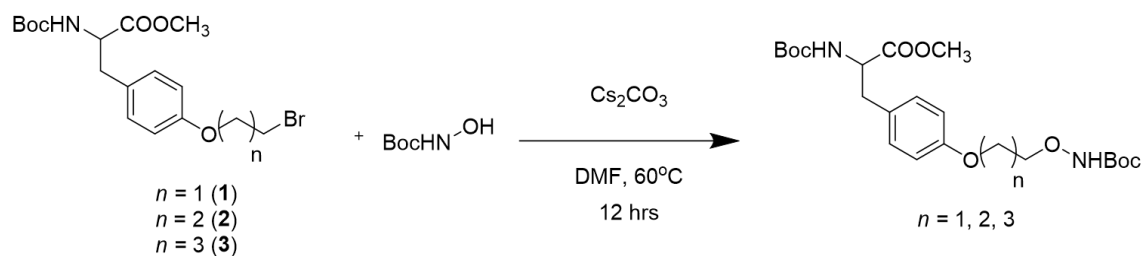
The Raines lab explored the hydrolytic stability of this particular carbon-nitrogen double bond relative to the hydrazone variant ($X^2=NR^3$).⁴⁵ Both bond types have an intrinsic hydrolytic stability, which is important for the bond's stability in the aqueous

environment of a cell. This stability is conferred from the increased negative charge that is imparted on the carbon group bonded to N¹, which reduces its electrophilicity. However, using pH-dependent rate profiles and titrations, the Raines lab found that oxime form is much more stable than the hydrazone form, which they speculated was due to the greater resistance of the oxime to protonation (the first step in the hydrolysis of the imine bond is the protonation of the X² group).

Based on these findings, the oxime ligation between ketones and hydroxylamine groups represents another chemistry that can be exploited for bioorthogonal purposes. For example, the Schutz lab employed the oxime ligation to conjugate an aminoxy-containing dye to a protein containing a ketone UAA.⁴³

While this method worked, aminoxy containing dyes are rare. A more ideal conjugation mechanism would involve the generation of an aminoxy-functionalized UAA to ligate with a ketone-containing dye, which are more readily available. Compound **1** through **3**, the protected bromo-tyrosine series, represents an ideal starting point for the formation of an aminoxy UAA. The bromine can readily be displaced via a Boc-protected hydroxylamine in the presence of base, affording an aminoxy-derivitized tyrosine containing alkyl chains of varying lengths. I have already performed this reaction using compound **3**, and following column chromatography, obtained the p-Aminoxy-butyl-*Boc-OMe*-tyrosine in good yield. Future studies will attempt to perform the oxime ligation using this UAA and a ketone-containing dye.

Scheme 14. General Reaction Scheme for Synthesis of Aminoxy-UAAs



IV. B. Continued Immobilization Trials

As discussed previously, GFP mutated with compounds **16** to **18**, the alkynyltyrosine UAA series, still needs to be immobilized on the alkyne derivitized resin via the Glaser-Hay coupling. Additionally, GFP mutated with compounds **10** to **12**, the bromotyrosine UAA series, still needs to be immobilized. In order to do this, we propose generating an amine derivitized sepharose resin. This amine can then displace the bromine via a nucleophilic substitution reaction, conjugating the GFP to the dye. Future work will attempt to generate the amine derivitized sepharose resin, as well as optimize reaction conditions for the substitution reaction.

V. CONCLUSION

In conclusion, I was able to synthesize nine, novel UAAs with various functional handles and alkyl chains. Furthermore, all UAAs were successfully incorporated into GFP using a polyubiquitous aaRS. The novel functional handles present in the UAAs allowed me to access different bioorthogonal chemistries. Additionally, the presence of alkyl chains of different lengths seemed to successfully distance the site of reaction from the protein surface, having differential effects on the outcome of the different bioorthogonal reactions. Based on these results, we have demonstrated the potential of UAA-mutagenesis to generate unique, site-specific functional handles in proteins for the conjugation of

fluorescent dyes. This technology has downstream molecular imaging applications for the study of protein dynamics *in vivo*.

Additionally, the UAAs developed here could also be used to generate novel sites of immobilization. Our previous immobilization studies only looked at GFP mutated with pAzF, a UAA in which the azido functionality comes directly off of the aromatic ring of the phenylalanine. We reasoned that the efficiency of the copper click reaction in this context was hindered by the bulk of the protein immediately surrounding the functional handle. By extending the functional handles off of the protein surface with alkyl chains of different lengths, we were able to improve immobilization efficiency in the case of the GFP mutated with p2AzY, p3AzY, and p4AzY. The best efficiency came from the p3AzY-mutant, which interestingly paralleled the click efficiencies with the fluorescent dye experiments (see Figures 8 and 13). Thus, the two-carbon alkyl linker may be too sterically hindered for access by the reaction partner. Likewise, the four-carbon alkyl linker may be too flexible, preventing the functional handle from orienting in the proper fashion for reaction with the reaction partner.

Overall, the UAAs described represent a novel means to both probe biological systems as well as improve current immobilization technologies.

CHAPTER 3: DEVELOPMENT OF RAPID MICROWAVE-MEDIATED AND LOW-TEMPERATURE BACTERIAL TRANSFORMATIONS

I. INTRODUCTION

Due to the requisite for substantially high bacterial transformation efficiencies (currently the upper limits of $\sim 10^9$) for the creation of a synthetase plasmid library for the evolution of aminoacyl-tRNA synthetases, our lab sought to generate a more efficient transformation technology. Ideally, this novel methodology could potentially increase efficiency while decreasing overall cost, in terms of time producing competent cells and technical equipment, of the transformation.

Current bacterial transformation techniques can be summed up as a series of three basic steps.⁴⁶ In the first step, bacteria cells are made competent for the introduction of foreign genetic material, typically in the form of a plasmid. Next, some kind of physical shock, either in temperature or voltage, is applied to the cells to allow for the passage of DNA across the membrane. Finally, the cells are recovered for about an hour before undergoing a selection process.

Typical DNA transformations involve one of two approaches.^{46,47} Perhaps the least technologically sophisticated are chemical transformations. These involve the pretreatment of cells with calcium chloride. As a consequence of its negatively charged phosphate backbone, DNA is repelled by the hydrophobic core of the bacteria's phospholipid cell membrane. The presence of the divalent calcium cation helps to mitigate the negative charge of the DNA, increasing its propensity for passage across the membrane. Next, the

cells are placed on ice, which slows down the random motion of the phospholipid bilayer. The cells then undergo a heat shock for 50 seconds at 42°C. It is speculated that the rapid change in temperature causes “pores” to expand in the membrane, which allow for the passage of the calcium-DNA complex into the cell. chemical transformations have been shown to have transformation efficiencies ranging from 10^5 to 10^9 colony-forming-units (cfu).

Another DNA transformation technique is electroporation, which requires specialized cell stocks and equipment.⁴⁶⁻⁴⁸ In the first step, bacterial cells are prepared for electrical shock treatment via the removal of any conducting ions from the cell media and the concentration of cells in glycerol. Once prepared, the cells are then subjected to high energy electrical pulses (5-10 kV/cm for 5-10 μ s), resulting in the transient formation of pores in the cell membrane for the transport of plasmid DNA. The cells are then allowed to recover from the stress and subjected to a selection procedure. Electroporation techniques have led to transformation efficiencies as high as 10^{10} cfu.^{48,49}

While both approaches have been revolutionary in aiding the progress of molecular biology, they have their limitations. First, both require the costly and somewhat time-extensive preparation of cells prior to transformation.⁴⁶ Second, both require specialized equipment that can either reliably control temperatures (for chemical transformations) or deliver precise electrical pulses (electroporation). Finally, while both techniques have high transformation efficiencies, their upper limits dictate the maximum size of a plasmid library. Particularly in approaches such as ours, where the generation of a larger library would be ideal for the evolution of more suitable synthetases for specific UAAs, this can be huge impediment in the speed at which science occurs.

In particular, our lab is interested in the application of microwave technology for biological approaches. Microwaves are oscillating electromagnetic fields with wavelengths in the range of 1 mm to 1 m and frequencies of 0.3 to 300 GHz.⁵⁰ It is hypothesized that the electric field portion of the microwave irradiation is capable of heating materials through a phenomenon known as “microwave dielectric heating.” According to this hypothesis, the electric portion of the microwave field interacts with the dipolar polarization of a molecule. As the field oscillates, it causes the dipole within the molecule to oscillate as well, in an effort to realign with the electric field. However, due to the frequencies within the microwave region, the electric field oscillates too quickly for the molecule to realign with the field. As a result, the molecule has a phase difference between the orientation of its dipole and that of the electric field, producing heat in the form of energy lost from the dipole due to molecular collisions and friction. This heating is called dielectric heating.

Based on these principles, the key to microwave effects is presence of a dipole moment within the molecule. Due to the large abundance of polar molecules that comprise the cell membrane, especially phospholipids with a polar head group, our lab was interested in investigating the potential of microwave technology towards the application of microwave-mediated bacterial transformations.

Recent reports have already investigated the ability of traditional microwaves (those found in a kitchen for example) to replace a water bath as the “heat shock” during transformations.⁵¹ The literature does report a three-fold increase in transformation efficiency when employing 1-minute pulses at 180 W; however, the cells used in these studies were pretreated with calcium chloride to increase competency. Compared to

conventional microwaves, microwave reactors can deliver focused and narrow pulses of microwave irradiation. Since biological cell membranes are composed of highly polarizable phospholipid molecules, and because this is the main obstacle to the uptake of exogenous DNA, we hypothesized that it may be feasible to further enhance the utility of microwave transformation by replacing the shock with a microwave pulse. Furthermore, this strategy could optimize cell viability, as microwave reactors can be chilled prior to irradiation.⁵⁰ This allows for the transmission of microwave energy without the process of heating, which results in cell death and decreased efficiency overall. Finally, because microwave irradiation directly interacts with cell membrane and other biomacromolecules, it may not be necessary to pretreat cells prior to transformation, increasing the utility of this approach.

Additionally, other studies have demonstrated that a freeze-thaw shock can replace the more traditional heat shock in transformation protocols.⁵² As such, we sought to investigate the potential of liquid nitrogen shock (-198 °C) transformation as a viable alternative to current transformation technologies that does not require pretreated cells.

II. RESULTS AND DISCUSSION

Initial attempts aimed at generating a microwave-based approach for bacterial transformation involved replacing the water-bath heat shock in chemotransformation with pulses of microwave irradiation in a CEM Discover microwave reactor. *Escherichia coli* BL21(DE3) strain of cells were grown to stationary phase (OD₆₀₀ 0.5) or log phase (OD₆₀₀ 3.2) followed by centrifugation and resuspension in either 1:1 LB/water or 1:1 LB/glycerol. Cells were then incubated on ice with a pET plasmid containing ampicillin resistance and subjected to brief pulses (2 or 5s) of microwave irradiation (50, 100, 150, 200 or 300 W)

in the microwave reactor. Following a 1 hr recovery time in LB media, the cells were plated on agar plates containing ampicillin. Unfortunately, none of the cells grew. In order to rule out that the microwave irradiation was killing the cells, the same microwave pulses were performed in the absence of any plasmid and cells were grown on agar plates containing no antibiotic. Interestingly, no differential cellular viability was observed, indicating that the lack of growth on the previous ampicillin plates was indeed due to a failed transformation.

Due to the abundance of polarizable ions in the aqueous media that the cells were grown up in, the samples were found to rapidly heat in the microwave. We hypothesized that this rapid heating caused the double-stranded plasmid to denature, as most DNA requires annealing temperatures between 50-65 °C. As such, linear, single-stranded DNA may have been transferred into the cells, which resulted in a failed transformation. Moreover, the temperatures attained may have promoted cellular death prior to the transformation event.

To overcome this limitation of the microwave transformation, we attempted a similar transformation procedure in a microwave reactor equipped with a CEM CoolMate system. This cooling system uses a microwave-transparent solution to cool the microwave reaction vessel to a temperature well below 0 °C, while still allowing for the input of microwave irradiation into the system. We hypothesized that by cooling the transformation we could prevent the heat cytotoxicity of the shock, while simultaneously generating an oscillating field that could aid in the uptake of DNA across the cell membrane. Using a procedure similar to that outlined above, we attempted bacterial transformations of a pET plasmid containing ampicillin resistance. First, we used the CoolMate system to cool the

transformation to -30 °C (~1 mins), followed by brief (10-30 s) pulses of microwave irradiation (100, 200, or 300 W). After a 1 h recovery period, the cells were plated on agar containing ampicillin. Gratifyingly, colonies were observed, although the overall transformation efficiency remained low. The highest transformation efficiency was observed using 15 s pulses with 300 W of power, yielding efficiencies of 213 cfu.

Table 2.. Microwave Assisted Transformation Efficiencies

Power (W)	Time (s)	Initial Temperature (°C)	Transformation Efficiency
0	10	0	0
0	30	0	0
0	10	- 30	0
100	10	- 30	24
200	10	- 30	94
300	10	- 30	94
300	15	- 30	213
300	30	- 30	101
300	60	- 30	24
300 ^a	30	- 30	0
300 ^b	2	25	0

In the process of attempting various controls in the absence of microwave irradiation at reduced temperatures in order to mimic the microwave temperature profile, a rapid freezing of the cells was attempted using liquid nitrogen (-196 °C). Following recovery, this resulted in a surprisingly high transformation efficiency of 10⁴. A literature search revealed some protocols that had explored rapid freezing as a means to transform cells. Most notably, a report by Takahashi et al attempted a similar freeze-thaw

transformation with non-chemically competent cells, but only obtained efficiencies of ~100 under optimal conditions. Other reports used competent cells, and replaced the heat shock portion of the protocol with the rapid freezing in liquid nitrogen. While these protocols were able to report optimal transformation of $\sim 10^6$, most procedures were still 1-2 orders of magnitude lower. The fact that our approach appeared to generate similar results without the need for the preparation of competent cells made us interested in exploring the potential of low-temperature transformations.

In order to optimize the results observed in our freeze-thaw transformation, several experimental variables were altered including the number of freeze-thaw cycles, the temperature of the transformation, the concentration of divalent cations, and the OD₆₀₀ of the bacterial cells. Overall, the most optimal transformation conditions involved the growth of cell cultures to an OD₆₀₀ of 0.5-0.6, followed by concentration and subjection to two freeze-thaw cycles directly in liquid nitrogen. Interestingly, the presence of 10 mM Mg²⁺, which should help mask the negative charge on the diphosphate backbone of DNA, improving DNA uptake, had no effect on the transformation efficiency. Additionally, the use of chemically competent cells did not increase transformation efficiencies either. These results combined indicate that the mechanism for DNA uptake is not based on the ionic shielding of the charged DNA-backbone. Additionally, the growth phase of the bacteria appeared to have an impact on the transformation efficiency, with cells in early stages of replication and in stationary phase exhibiting lower overall efficiencies compared to cells in log phase. Furthermore, we also observed that increasing the temperature of the freezing above -42 °C resulted in no transformants. Resuspending the cells in a 10% glycerol solution, a cryoprotectant, rather than water also resulted in no transformants. These data

combined suggest that the mechanism of transformation in this case is the rapid formation of ice crystals, which must disrupt the bacterial cell wall. This “hole” in the wall provides a convenient means for the uptake of exogenous DNA. This data was further corroborated by the fact that increasing the number of freeze-thaw cycles >2 resulted in decreased transformation efficiencies. We hypothesize that this is due to the increased damage sustained by the bacterial cell walls.

Table 3. Freeze-Thaw Transformation Efficiencies

Cell OD ₆₀₀ ^a	Freeze-Thaw Cycles	Cation Additive ^b	Freeze Temp. (°C)	Transformation Efficiency
0.1	1	--	-196	5.2 x 10 ³
0.1	2	--	-196	2.9 x 10 ⁴
0.2	1	--	-196	2.8 x 10 ²
0.2	2	--	-196	2.8 x 10 ²
0.4	1	--	-196	8.4 x 10 ⁴
0.4	2	--	-196	2.3 x 10 ⁵
0.6	1	--	-196	2.2 x 10 ²
0.6	2	--	-196	3.7 x 10 ²
0.8	1	--	-196	1.6 x 10 ²
0.8	2	--	-196	3.5 x 10 ⁴
1.0	1	--	-196	5.0 x 10 ³
1.0	2	--	-196	4.8 x 10 ⁴
3.0	1	--	-196	7.2 x 10 ³
0.6	1	--	-94	1.17 x 10 ³
0.6	1	--	-78	1.6 x 10 ²
0.6	1	--	-42	38
0.6	1	--	-30	0
0.6	2	NH ₄ Cl	-196	8.9 x 10 ⁴
0.6	2	MgCl ₂	-196	1.5 x 10 ²
0.6	2	CaCl ₂	-196	1.1 x 10 ⁵

Beyond transformation variables, several different plasmids of varying sizes and harboring different antibiotic resistances were also tested for their ability to undergo low-temperature transformation. Gratifyingly, these transformations also yielded consistently high transformation efficiencies.

To further verify the presence of foreign DNA in the transformants, as well as the retention of cellular function under these conditions, isolation of plasmid and expression of a GFP reporter were successfully performed.

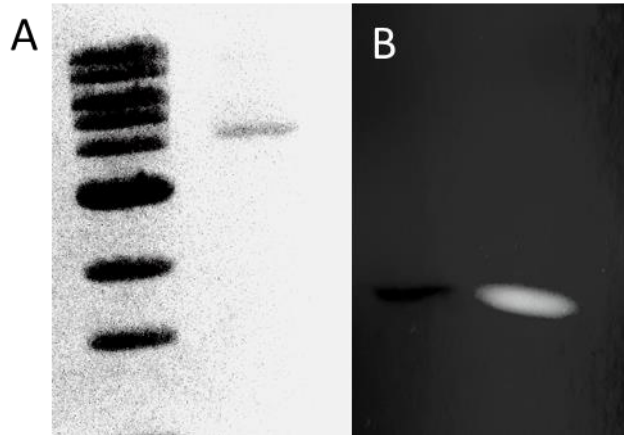


Figure 14. Validation of DNA transformation. A) Following either a liquid nitrogen or a microwave transformation of a pET22-GFP plasmid, a culture of cells were grown, the plasmid was isolated and digested with EcoRI. Analysis of the digest on an 1% agarose gel confirms the presence of the plasmid DNA. B) To further confirm the transformation and the competence of the cells the GFP protein was expressed, purified and analyzed via SDS-PAGE. Successful transfer of plasmid DNA, while retaining cellular function is demonstrated by the presence of GFP fluorescence

Furthermore, the ability to perform simultaneous transformation of two plasmids using this procedure was found possible, albeit with decreased transformation efficiencies ($\sim 10^2$), which is to be expected for the uptake of two separate plasmids.

Overall, while none of the transformation efficiencies obtained in *E. coli* were comparable to electroporation or chemotransformation ($\sim 10^5$ in this case), the fact that this protocol obviates the need to prepare competent cells makes it a highly efficient and useful technique for virtually any laboratory.

Finally, we sought to explore the potential application of both microwave and low-temperature transformation techniques towards the transformation of *Mycobacteria smegmatis*. This mycobacterium represents a more significant challenge for the transformation process, as its cell wall composition is more complex.⁵³ Furthermore, very few efficient mechanisms exist for the transfer of genetic material into this organism.

Although electroporation protocols have been generated, they still remain highly inefficient in comparison.⁵³ Based on mycobacterial involvement in clinical disease, the potential to transfer genetic material into the organism represents a huge advancement in scientific research.

Using identical freeze-thaw conditions as those described above, a pMycVec1 plasmid was employed towards *M. smegmatis* freeze-thaw transformation. Gratifyingly, the transformation yielded an efficiency of $\sim 10^3$ cfu. While this is lower than the efficiency associated with *E. coli*, this is to be expected as the literature demonstrates that even under electroporation conditions, *M. smegmatis* has efficiencies less than *E. coli*. Unfortunately, the optimal microwave mediated transformation conditions for *E. coli* were not transferable to *M. smegmatis*. This is not surprising, however, given the already lower transformation efficiencies.

Thus, we have demonstrated a novel strategy for the efficient and easy transformation of exogenous DNA into *M. smegmatis* that requires no additional preparatory steps.

III. EXPERIMENTAL

Optimized preparation of competent cells. 2xYT media (4 mL) was inoculated with *Escherichia coli* Novagen BL21(DE3) strain of cells and then incubated for 16 h at 37 °C to stationary phase. The starter culture (1 mL) was then added to fresh 2xYT media (100 mL) and incubated with vigorous shaking at 37 °C and monitored by routine optical density measurements on a spectrophotometer at 600 nm to assess bacterial growth. At the desired density, the culture was centrifuged at 4,000 rpm for 20 min. The resulting cell pellet was

resuspending in either deionized water or glycerol, for a final cell optical density of 27.7. The resuspended cells were then aliquoted (100 μ L each) and kept on ice.

Liquid nitrogen transformation. An aliquot of BL21(DE3) cells was kept on ice, and plasmid DNA (1 μ L, 70 μ g) was added. The cells were incubated on ice for 5 mins and then submerged in liquid nitrogen for 20 s. After thawing, the cells were either subjected to a second freeze cycle or recovered in 2xYT media (500 μ L), followed by 1 h of incubation at 37 $^{\circ}$ C. Cells were then plated on Luria-Bertani (LB) plates containing appropriate antibiotics and incubated overnight at 37 $^{\circ}$ C. Colonies were counted to assess transformation efficiency, and individual colonies from these plates were selected to ensure the presence of plasmid DNA.

Microwave CoolMate transformation. An aliquot of BL21(DE3) cells was kept on ice, and plasmid DNA (1 μ L, 70 μ g) was added. The cells were incubated on ice for 5 mins and then placed in a Discover CoolMate vessel. Each sample was cooled in the CoolMate to a specific temperature (typically -30 $^{\circ}$ C) before it was subjected to microwave radiation and simultaneous cooling. The temperature was monitored using a fiber optic probe, and upon the conclusion of a microwave pulse, the sample was removed from the vessel and thawed to room temperature. The cells were recovered in 2xYT media (500 μ L), followed by 1 h of incubation at 37 $^{\circ}$ C. Cells were then plated on LB plates containing appropriate antibiotics and incubated overnight at 37 $^{\circ}$ C. Colonies were counted to assess transformation efficiency, and individual colonies from these plates were selected to ensure the presence of DNA.

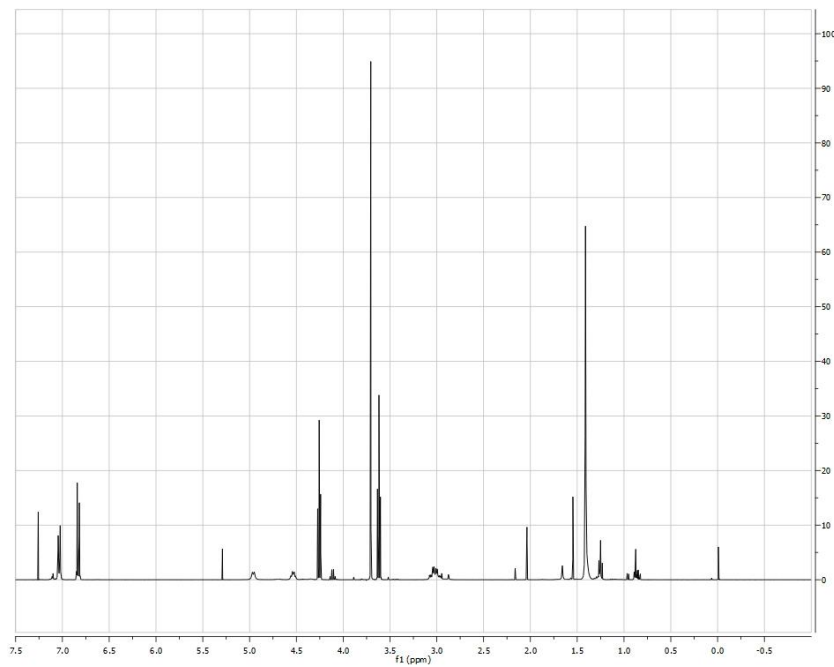
IV. CONCLUSION

Overall we have demonstrated two novel means to introduce foreign DNA into bacterial systems. The freeze-thaw technique in particular obviated the need to generate competent cells, cutting down the amount of time and equipment required to transform cells. Furthermore, this freeze-thaw approach also represents a new means to introduce DNA into hardier bacterial species, such as the *Mycobacteria*.

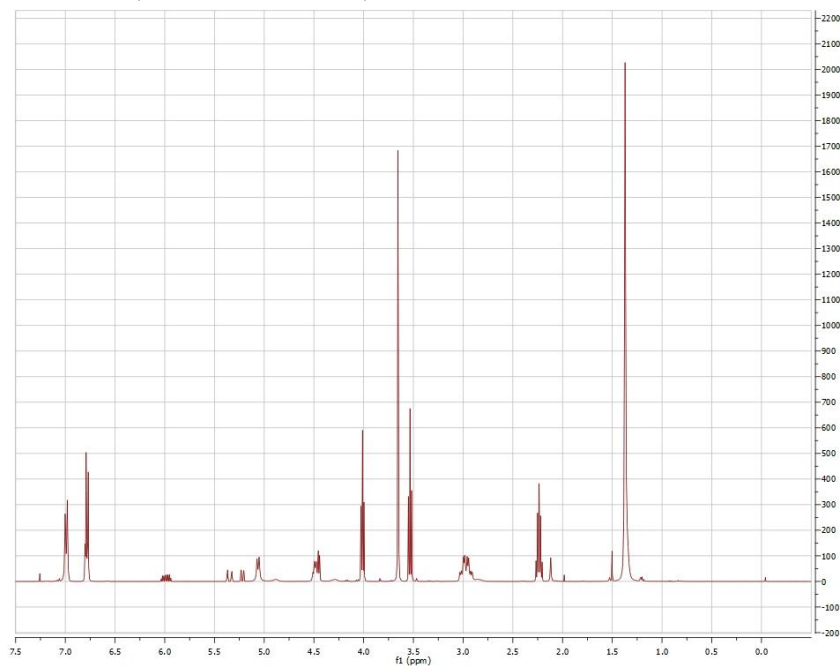
APPENDIX A: NMR SPECTRA

p-Bromo-ethyl-*Boc*-*OMe*-Tyrosine (1)

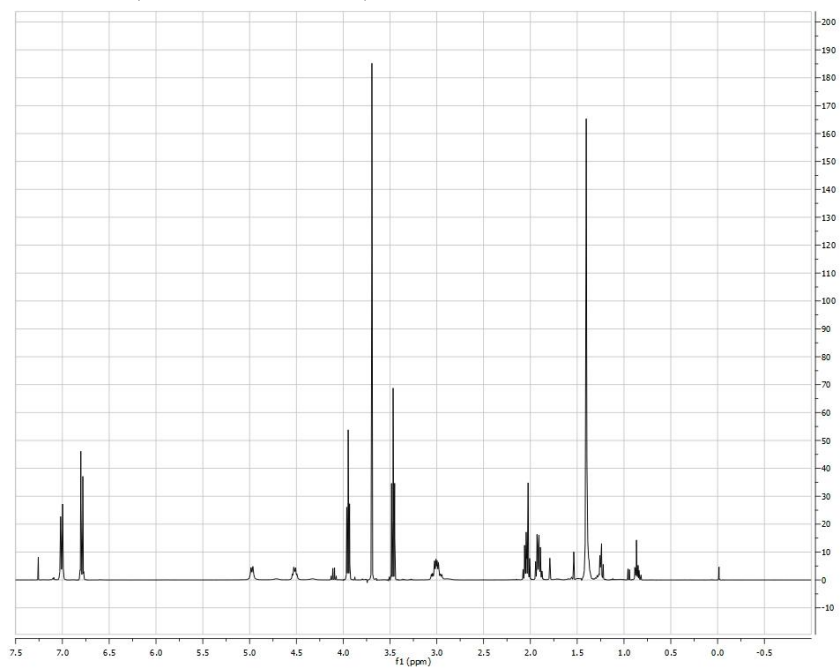
^1H NMR (400 MHz, CDCl_3)



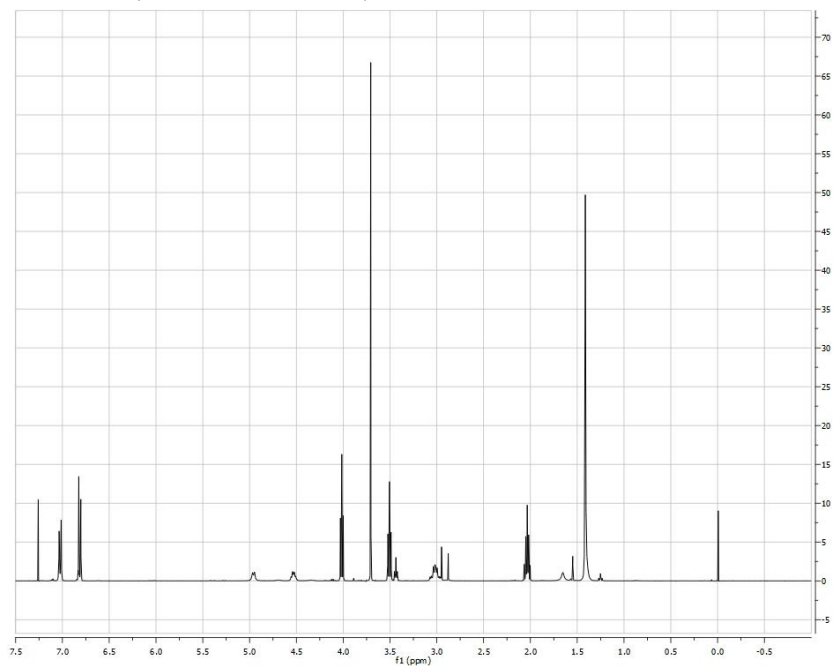
***p*-Bromo-propyl-*Boc*-*OMe*-Tyrosine (2)**
¹H NMR (400 MHz, CDCl₃)



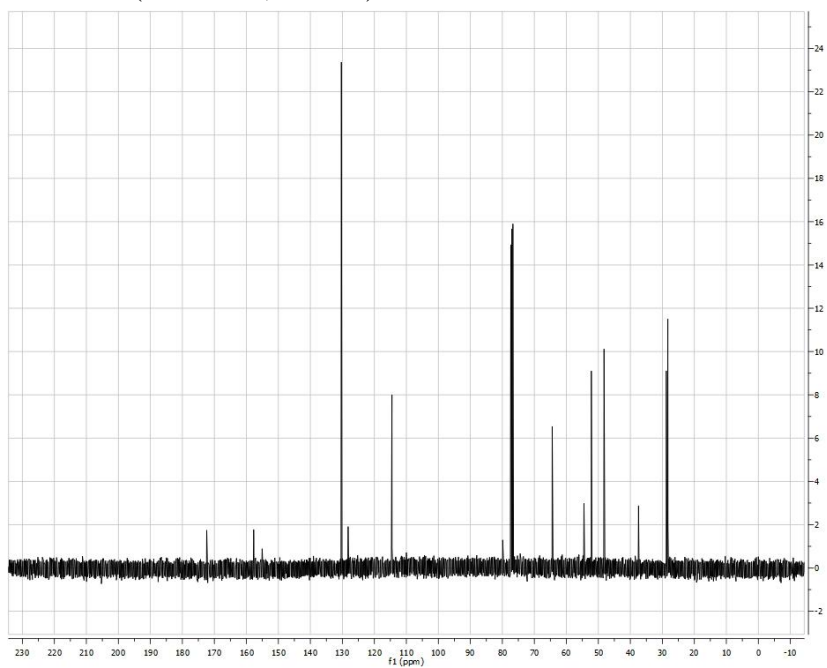
***p*-Bromo-butyl-*Boc*-*OMe*-Tyrosine (3)**
¹H NMR (400 MHz, CDCl₃)



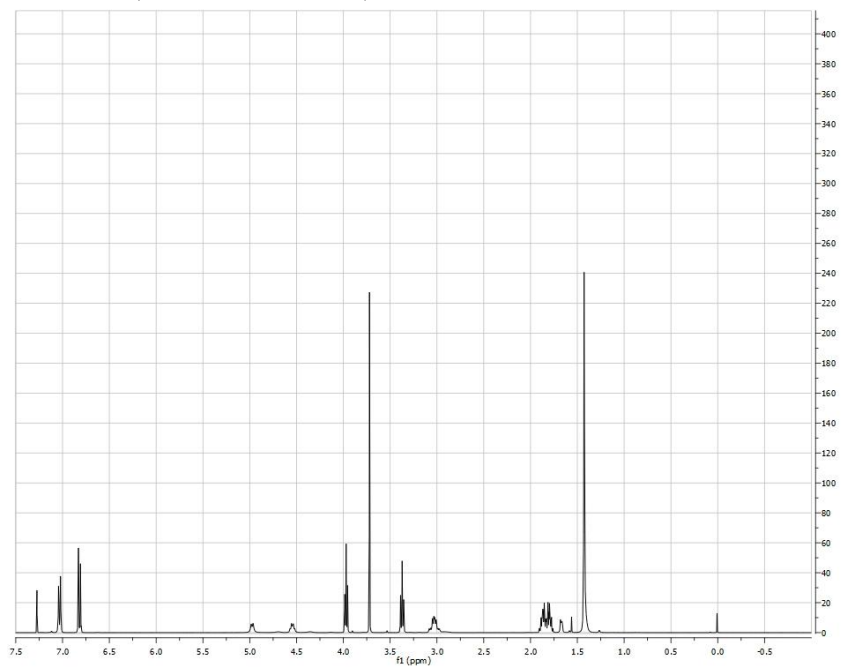
***p*-Azido-propyl-*Boc*-*OMe*-Tyrosine (5)**
¹H NMR (400 MHz, CDCl₃)



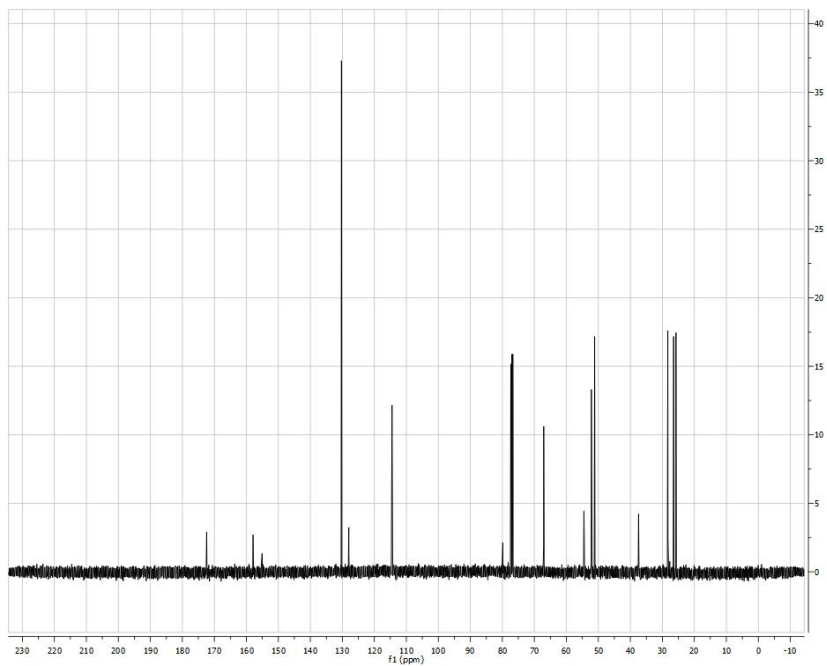
¹³C NMR (400 MHz, CDCl₃)



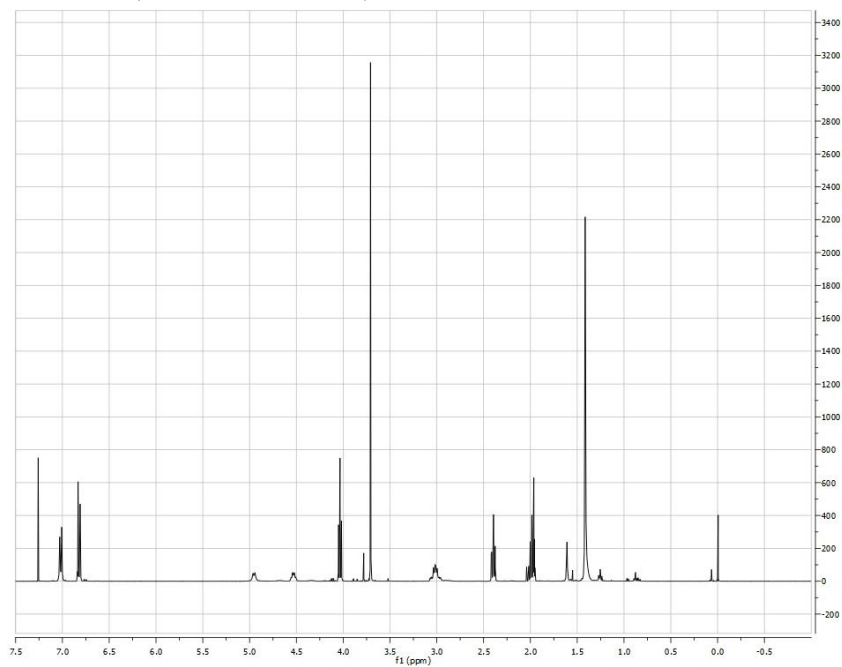
***p*-Azido-butyl-*Boc*-*OMe*-Tyrosine (6)**
¹H NMR (400 MHz, CDCl₃)



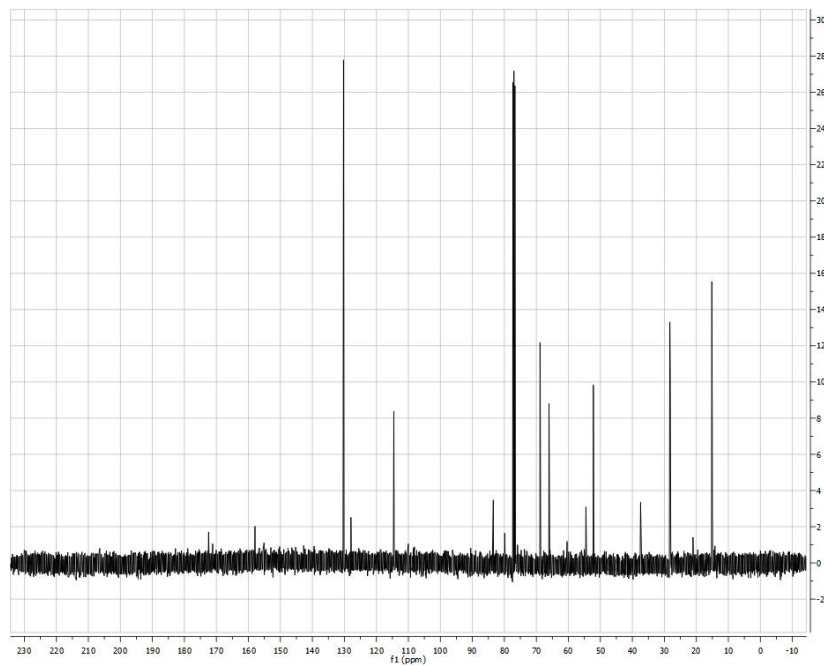
¹³C NMR (400 MHz, CDCl₃)



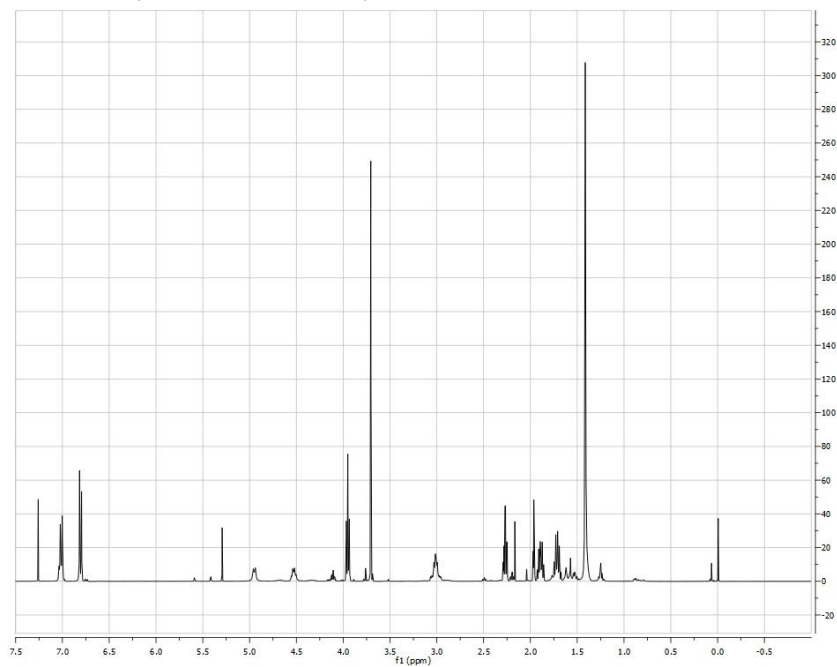
***p*-Alkynyl-propyl-*Boc*-*OMe*-Tyrosine (8)**
¹H NMR (400 MHz, CDCl₃)



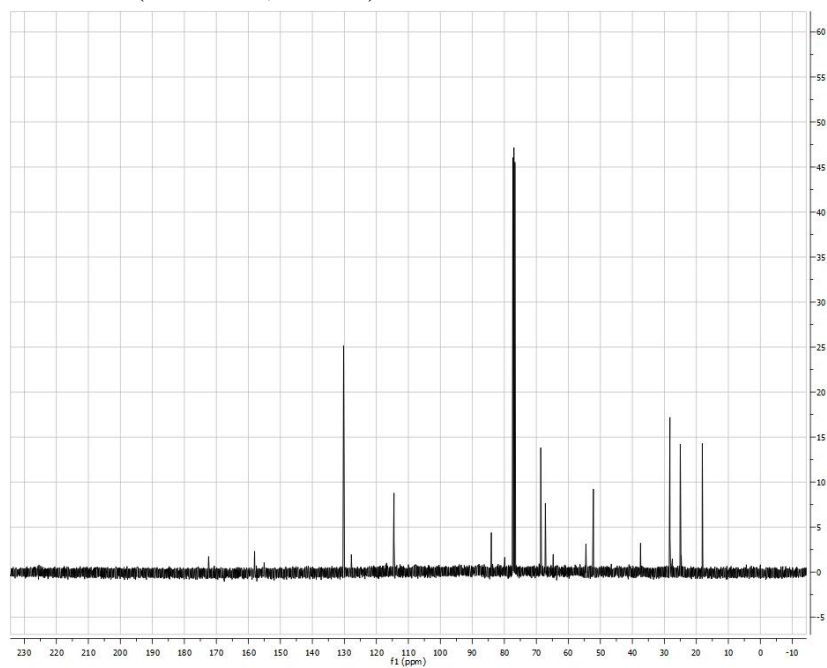
¹³C NMR (400 MHz, CDCl₃)



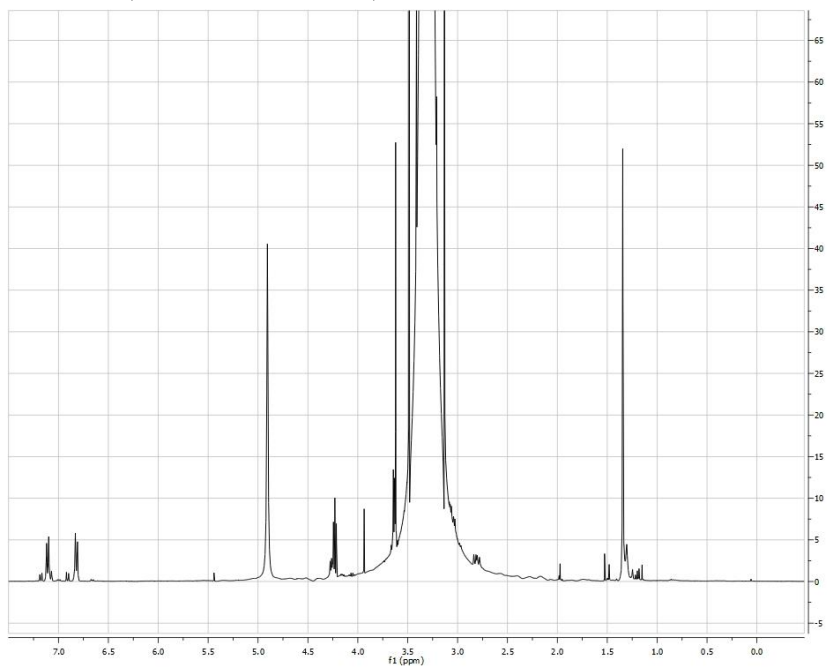
***p*-Alkynyl-butyl-*Boc*-*OMe*-Tyrosine (9)**
¹H NMR (400 MHz, CDCl₃)



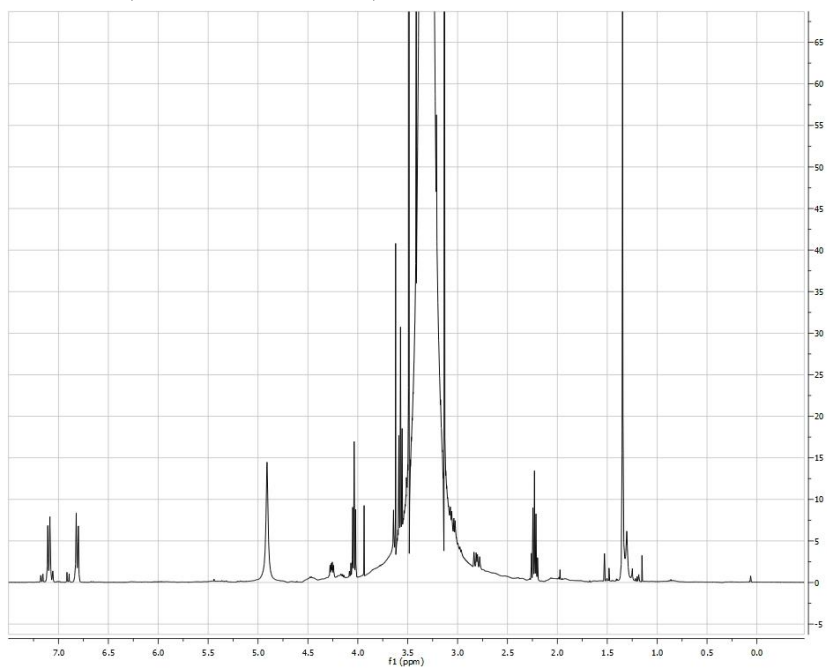
¹³C NMR (400 MHz, CDCl₃)



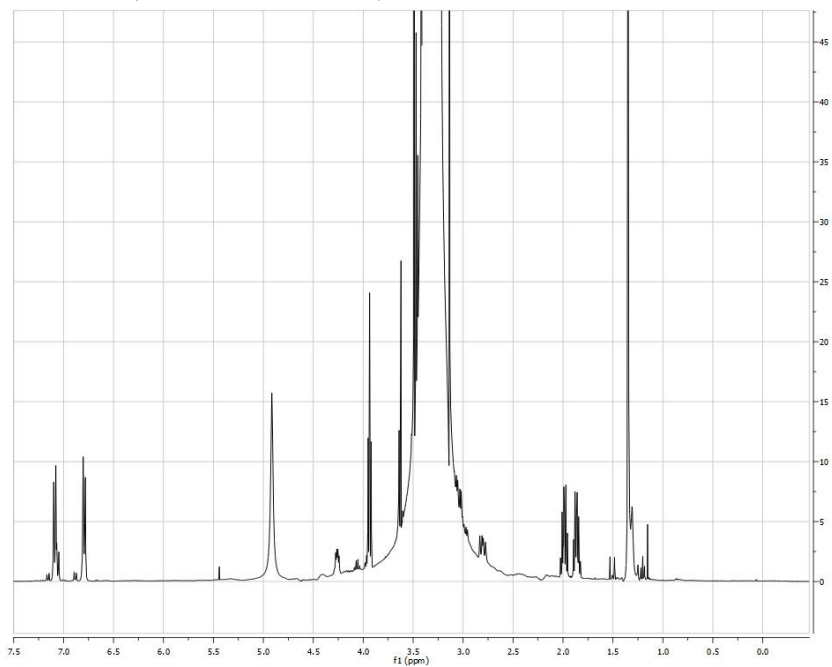
***p*-Bromoethyltyrosine (p2BrY, 10)**
¹H NMR (400 MHz, CD₃OD)



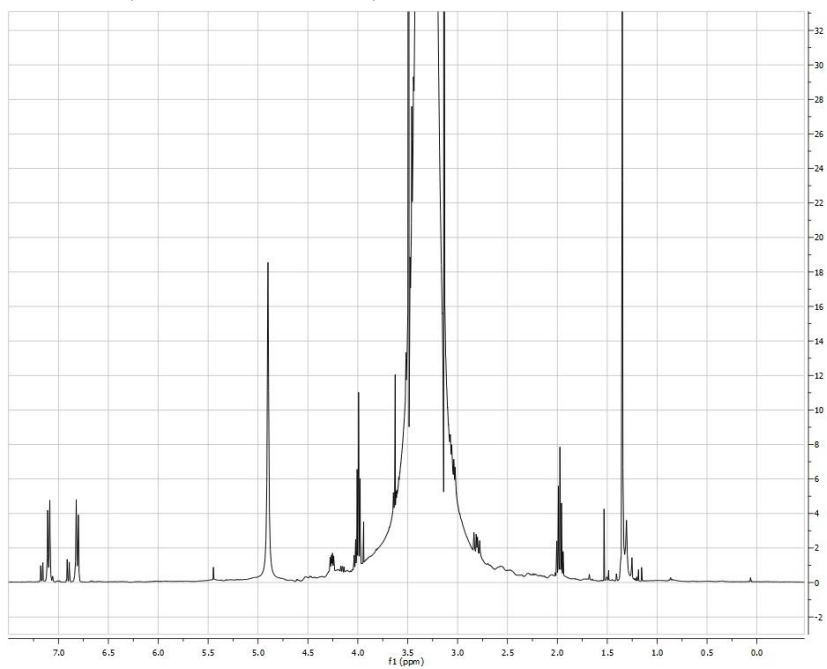
***p*-Bromopropyltyrosine (p3BrY, 11)**
¹H NMR (400 MHz, CD₃OD)



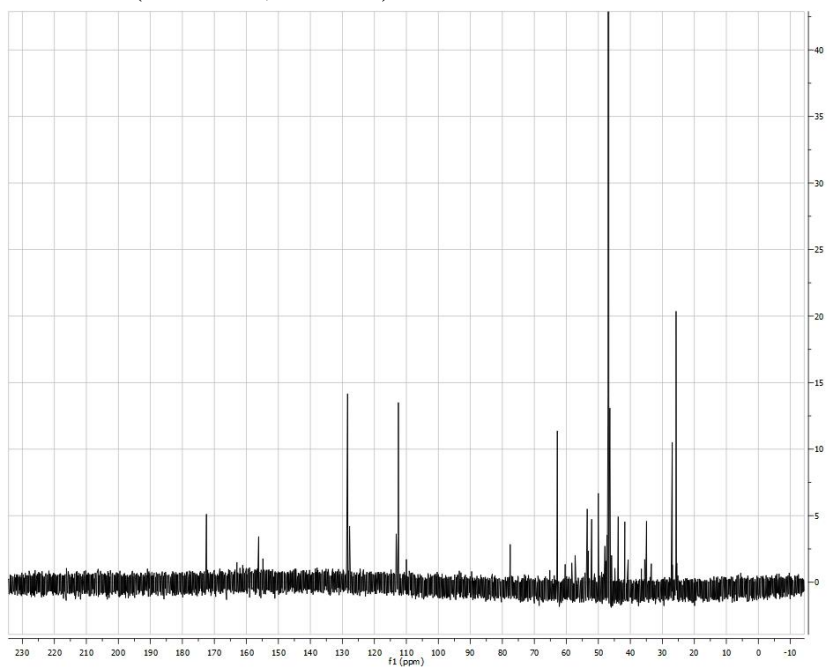
***p*-Bromobutyltyrosine (p4BrY, 12)**
¹H NMR (400 MHz, CD₃OD)



***p*-Azidopropyltyrosine (p3AzY, 14)**
¹H NMR (400 MHz, CD₃OD)

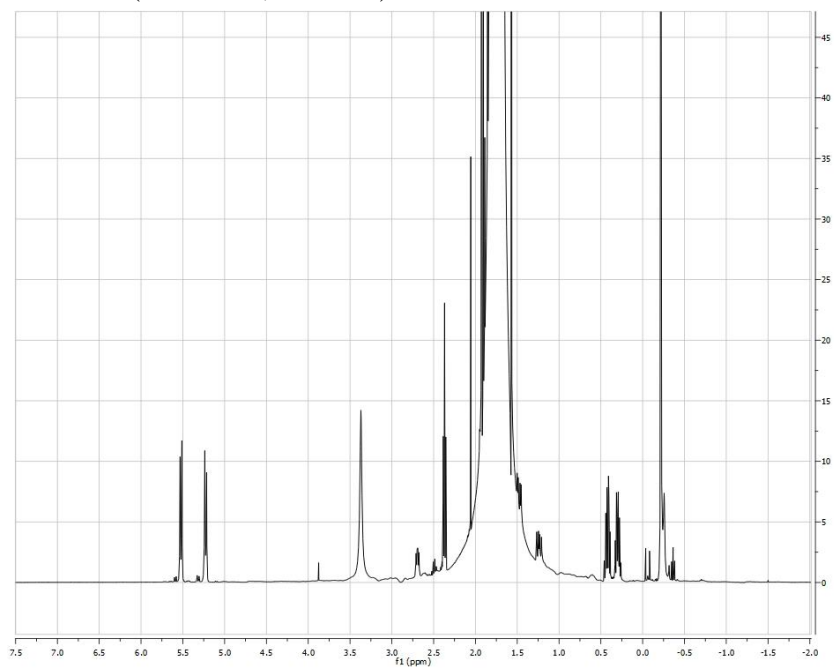


^{13}C NMR (400 MHz, CD_3OD)

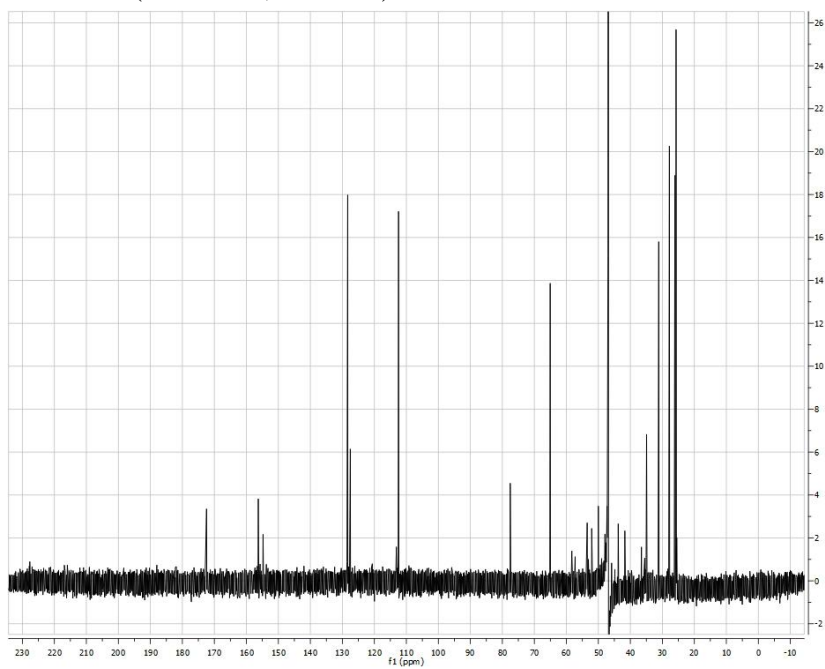


p-Azidobutyltyrosine (p4AzY, 14)

^1H NMR (400 MHz, CD_3OD)

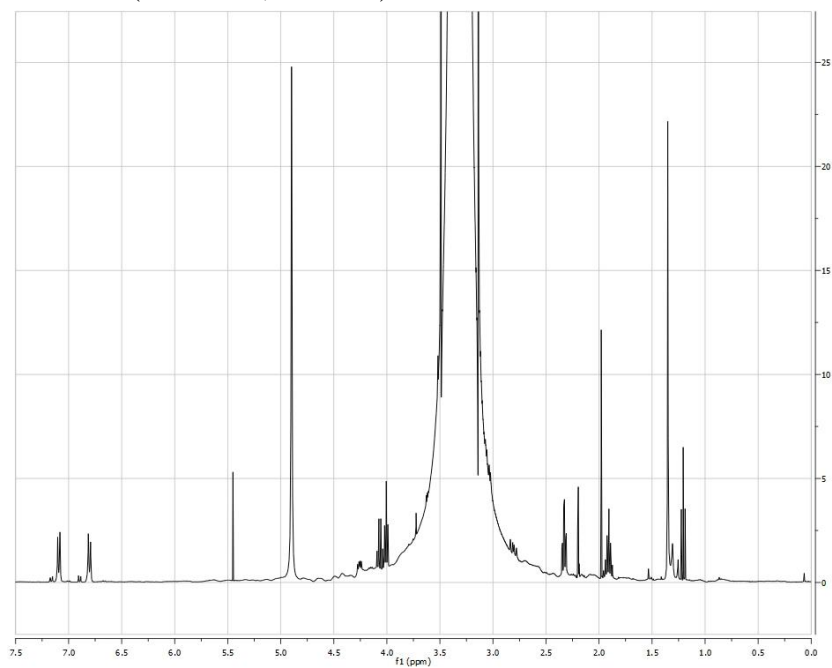


^{13}C NMR (400 MHz, CD_3OD)

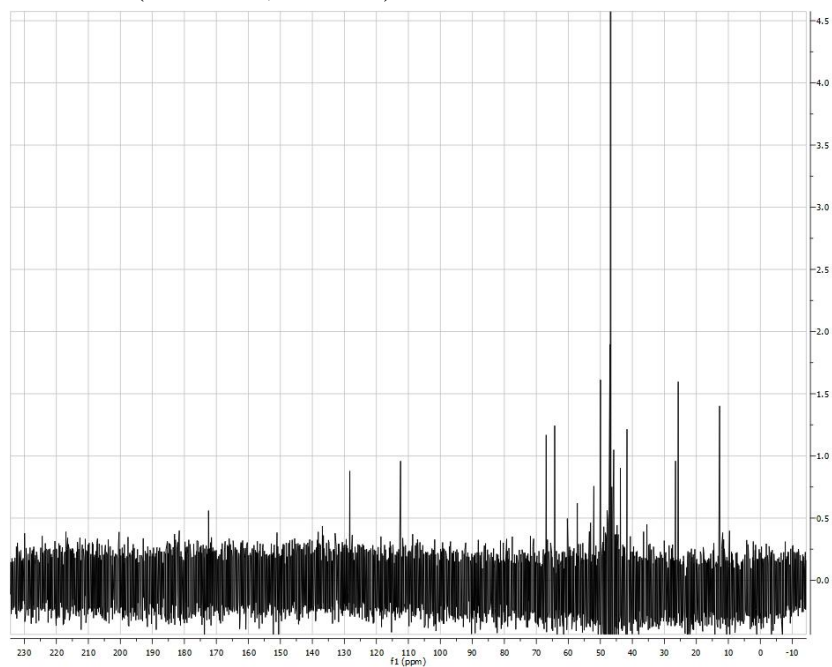


p-Alkynylpropyltyrosine (p3yneY, 13)

^1H NMR (400 MHz, CD_3OD)

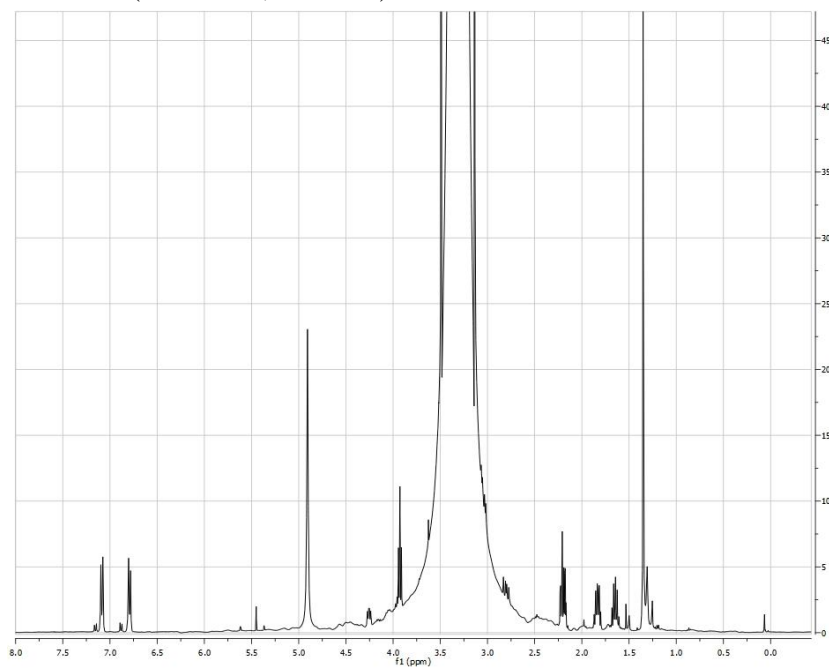


^{13}C NMR (400 MHz, CD_3OD)

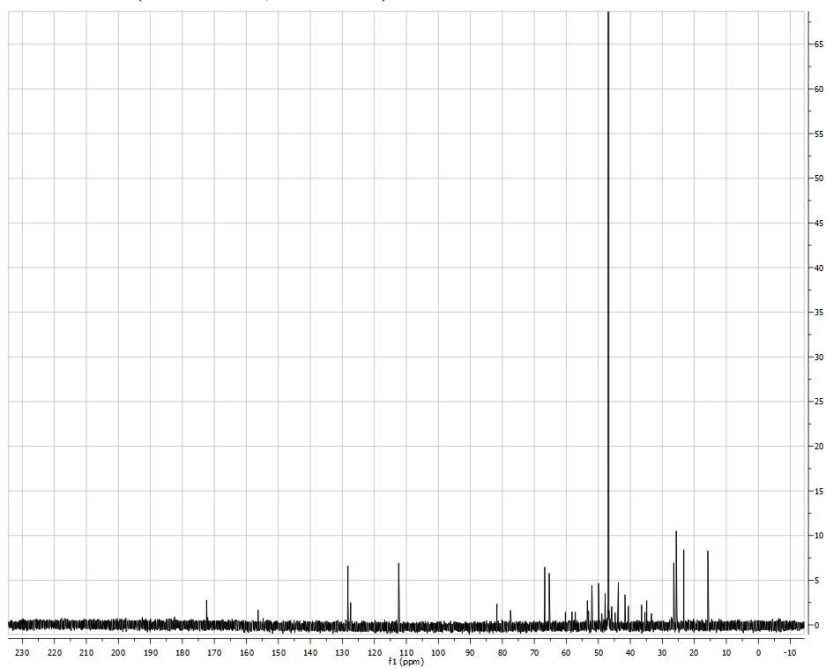


p-Alkynylbutyltyrosine (**p4yneY**, **13**)

^1H NMR (400 MHz, CD_3OD)



^{13}C NMR (400 MHz, CD_3OD)



ACKNOWLEDGEMENTS

I'd like to thank my family, who has supported me in everything that I do.

I'd like to thank Valerie Tripp, who mentored me when I first joined the Young Lab and was fundamental in developing my laboratory technique

I'd like to thank Jackie McKenna, whose work on this project helped jumpstart this project. She was also an amazing hoodmate.

I'd like to thank Matt Freedman, who helped me make and purify various UAAs.

I'd like to thank Benjamin Raliski, who provided constant moral support and was an amazing sounding board when experiments were not working. His ideas helped push this project forward.

I'd like to thank Jordan Villa and Jessica Lampkowski, who helped console me in lab when experiments were not working.

I'd like to thank Ryan Tyler, who was instrumental in keeping me calm throughout this process.

I'd like to thank the members of my Honors committee, Dr. McNamara, Dr. Landino, and Dr. Forsyth, all of whom have been instrumental in helping me better understand how the world works and in inspiring within me an interest in science.

I'd like to thank Dr. Doug Young, who saw potential within me that I didn't even know existed. He constantly believed in me even when I didn't. The opportunity to work under him has been a life-changing experience, and has inspired within me a deep interest in science.

REFERENCES

1. Berg, J. M., Tymoczko, J. L., Gatto, G. J. & Stryer, L. *Biochemistry*. (MacMillan, 2015).
2. Nguyen, Q. T. & Tsien, R. Y. Fluorescence-guided surgery with live molecular navigation--a new cutting edge. *Nat. Rev. Cancer* **13**, 653–62 (2013).
3. Young, T. S. & Schultz, P. G. Beyond the canonical 20 amino acids: Expanding the genetic lexicon. *J. Biol. Chem.* **285**, 11039–11044 (2010).
4. Zhu, H. & Snyder, M. Protein chip technology. *Curr. Opin. Chem. Biol.* **7**, 55–63 (2003).
5. Rusmini, F., Zhong, Z. & Feijen, J. Protein immobilization strategies for protein biochips. *Biomacromolecules* **8**, 1775–1789 (2007).
6. Tan, W. *et al.* Optical protein sensor for detecting cancer markers in saliva. *Biosens. Bioelectron.* **24**, 266–271 (2008).
7. Jokerst, J. V. *et al.* Nano-bio-chips for high performance multiplexed protein detection: Determinations of cancer biomarkers in serum and saliva using quantum dot bioconjugate labels. *Biosens. Bioelectron.* **24**, 3622–3629 (2009).
8. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).
9. Wang, L., Xie, J. & Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 225–249 (2006).
10. Cornish, V. W. & Schultz, P. G. Mutagenesis an. (1995).
11. Martin, a B. & Schultz, P. G. Opportunities at the interface of chemistry and biology. *Trends Cell Biol.* **9**, M24–M28 (1999).
12. Strable, E. *et al.* Unnatural amino acid incorporation into virus-like particles. *Bioconjugate ...* 866–875 (2008). at <http://pubs.acs.org/doi/abs/10.1021/bc700390r>
13. Mehl, R. a. *et al.* Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* **125**, 935–939 (2003).
14. 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).

15. Wang, L., Brock, A., Hererich, B. & Schultz, P. G. Expanding the genetic code of *E. Coli*. *Science* **292**, 498–500 (2001).
16. Wang, L. & Schultz, P. G. A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* **8**, 883–890 (2001).
17. Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew. Chemie - Int. Ed.* **48**, 6974–6998 (2009).
18. Borges, C. R. & Watson, J. T. Recognition of cysteine-containing peptides through prompt fragmentation of the 4-dimethylaminophenylazophenyl-4'-maleimide derivative during analysis by MALDI-MS. *Protein Sci.* **12**, 1567–1572 (2003).
19. Kalia, J. & Raines, R. T. Catalysis of imido group hydrolysis in a maleimide conjugate. *Bioorganic Med. Chem. Lett.* **17**, 6286–6289 (2007).
20. Huisgen, R. Kinetics and Mechanism of 1,3-Dipolar Cycloadditions. *Angew. Chemie Int. Ed. English* **2**, 633–645 (1963).
21. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. Supporting Information for ' A Stepwise Huisgen Cycloaddition Process Catalyzed by Copper (I): Regioselective Ligation of Azides and Terminal Alkynes '. *Angew. Chemie Int. Ed. English* **114**, 2708–2711 (2002).
22. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
23. 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).
24. Chan, T. R., Hilgraf, R., Sharpless, K. B. & Fokin, V. V. Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.* **6**, 2853–2855 (2004).
25. 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. 15046–15047 (2004).
26. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16793–16797 (2007).
27. Siegrist, M. S. *et al.* D-amino acid chemical reporters reveal peptidoglycan dynamics of an intracellular pathogen. *ACS Chem. Biol.* **8**, 500–505 (2013).

28. Swarts, B. M. *et al.* Probing the mycobacterial trehalome with bioorthogonal chemistry. *J. Am. Chem. Soc.* **134**, 16123–16126 (2012).
29. Glaser, C. *Ber. Dtsch. Chem. Ges.* **2**, 422–424 (1869).
30. Hay, A. Oxidative Coupling of Acetylenes. III. *J. Org. Chem.* 3–4 (1962). doi:10.1021/jo01056a511
31. Vilhelmsen, M. H., Jensen, J., Tortzen, C. G. & Nielsen, M. B. The Glaser-Hay reaction: Optimization and scope based on ¹³C NMR kinetics experiments. *European J. Org. Chem.* **3**, 701–711 (2013).
32. Fomina, L., Vazquez, B., Tkatchouk, E. & Fomine, S. The Glaser reaction mechanism. A DFT study. *Tetrahedron* **58**, 6741–6747 (2002).
33. Lampkowski, J. S., Villa, J. K. & Young, D. D. Development and Optimization of Glaser-Hay Bioconjugations. *Angew. Chemie Int. Ed. English (pending Revis.* (2015).
34. Craggs, T. D. Green fluorescent protein: structure, folding and chromophore maturation. *Chem. Soc. Rev.* **38**, 2865–2875 (2009).
35. Raliski, B. K. *et al.* Site-Specific Protein Immobilization Using Unnatural Amino Acids. (2014).
36. Tripp, V. T., Lampkowski, J. S., Tyler, R. & Young, D. D. Development of solid-supported glaser-hay couplings. *ACS Comb. Sci.* **16**, 164–7 (2014).
37. Chen, X.-H. *et al.* Genetically Encoding an Electrophilic Amino Acid for Protein Stapling and Covalent Binding to Native Receptors. *ACS Chem. Biol.* (2014). doi:10.1021/cb500453a
38. Xiang, Z. *et al.* Proximity-enabled protein crosslinking through genetically encoding haloalkane unnatural amino acids. *Angew. Chemie - Int. Ed.* **53**, 2190–2193 (2014).
39. Young, D. D., Jockush, S., Turro, N. J. & Schultz, P. G. Synthetase polyspecificity as a tool to modulate protein function. *Bioorganic Med. Chem. Lett.* **21**, 7502–7504 (2011).
40. Deiters, A., Cropp, T. A., Summerer, D., Mukherji, M. & Schultz, P. G. Site-specific PEGylation of proteins containing unnatural amino acids. *Bioorganic Med. Chem. Lett.* **14**, 5743–5745 (2004).
41. Deiters, A. & Schultz, P. G. In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorganic Med. Chem. Lett.* **15**, 1521–1524 (2005).

42. Landino, L. M. *et al.* Fluorescein-labeled glutathione to study protein S-glutathionylation. *Anal. Biochem.* **402**, 102–104 (2010).
43. Brustad, E. M., Lemke, E. a., Schultz, P. G. & Deniz, A. a. A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *J. Am. Chem. Soc.* **130**, 17664–17665 (2008).
44. Sayer, J. M., Peskin, M. & Jencks, W. P. Imine-Forming Elimination Reactions. I. General Base and Acid Catalysis and Influence of the Nitrogen Substituent on Rates and Equilibria for Carbinolamine Dehydration. *J. Am. Chem. Soc.* **95**, 4277–4287 (1973).
45. Kalia, J. & Raines, R. T. Hydrolytic stability of hydrazones and oximes. *Angew. Chemie - Int. Ed.* **47**, 7523–7526 (2008).
46. Sambrook, J. & Russell, D. W. *Molecular cloning: a laboratory manual*. (Cold Spring Harbor Laboratory Press, 2001).
47. Aune, T. E. V & Aachmann, F. L. Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Appl. Microbiol. Biotechnol.* **85**, 1301–1313 (2010).
48. Neumann, E., Schaefer-Ridder, M., Wang, Y. & Hofschneider, P. H. Gene transfer into mouse lymphoma cells by electroporation in high electric fields. *EMBO J.* **1**, 841–845 (1982).
49. D, H. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580 (1983).
50. Collins, J. M. & Leadbeater, N. E. Microwave energy: a versatile tool for the biosciences. *Org. Biomol. Chem.* **5**, 1141–1150 (2007).
51. Fregel, R., Rodríguez, V. & Cabrera, V. M. Microwave improved *Escherichia coli* transformation. *Lett. Appl. Microbiol.* **46**, 498–499 (2008).
52. Weigel, D. & Glazebrook, J. Transformation of *Agrobacterium* using the freeze-thaw method. *CSH Protoc* **2006**, 1031–1036 (2006).
53. Wards, B. J. & Collins, D. M. Electroporation at elevated temperatures substantially improves transformation efficiency of slow-growing mycobacteria. *FEMS Microbiol. Lett.* **145**, 101–105 (1996).