4-2014

Synthesis of a Novel Unnatural Amino Acid for Protein Incorporation and Click Mediated Conjugation

Christopher A. Farley
College of William and Mary

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

Part of the Chemistry Commons

Recommended Citation
https://scholarworks.wm.edu/honorstheses/67

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.
Synthesis of a novel unnatural amino acid for protein incorporation and click mediated conjugation

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

by

Christopher Alexander Farley

Accepted for ________________________________

______________________________________
Douglas Young, Committee Chair

______________________________________
Christopher Abelt

______________________________________
Robert Hinkle

______________________________________
Matthew Wawersik

Williamsburg, VA
April 28, 2014
**Table of Contents**

Acknowledgements iii

List of Figures iv

Abstract 1

Background

- Molecular biology 2
- Unnatural amino acids (UAAs) 5
- Click reaction 10
- Photocaging groups 12

Experimental

- Synthesis of a Clickable Photoreactive Unnatural Amino Acid 16
- Synthesis of Other Unnatural Amino Acids
  - (p-propargyloxy)phenylalanine 19
  - Ethylene-extended 20
  - Malonate-derived 21
- Incorporation of 1 into GFP 22
- Protein click reaction 23

Results and Discussion 24

Conclusion 35

Future Directions 36

References 38
Acknowledgements

I would like to thank Dr. Doug Young for his guidance and patience towards my efforts in this project and this thesis. I also thank the Young lab for their help and support, particularly Val Tripp for teaching me chemical techniques, and Johnathan Maza and Jackie McKenna for their help with the biological component of my work.
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>List and structures of the 20 canonical amino acids</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Diagram of the steps of translation</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Illustration of UAA incorporation</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Scheme for aaRS evolution</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>GFP structure; within the β-barrel is a unique fluorophore that confers its spectroscopic properties</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>Mechanism of Cu(I)-catalyzed [3+2] Huisgen reaction</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Mechanism of cysteine-maleimide conjugation</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>General decaging mechanism</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Mechanism of photoinitiated peptide cleavage</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>(O-nitrobenzyl)tyrosine (ONBY)</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>UAA 1</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>UAA 10</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>Proposed UAA with an ethylene-extended propargyl group</td>
<td>15</td>
</tr>
<tr>
<td>14</td>
<td>Synthetic scheme for UAA 1</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>Mechanism of nitration</td>
<td>25</td>
</tr>
<tr>
<td>16</td>
<td>Resubmission to nitration conditions</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>Mechanism of bromination of primary alcohol</td>
<td>27</td>
</tr>
<tr>
<td>18</td>
<td>Table and graphical depiction of fluorescent intensities of different aaRSs</td>
<td>30</td>
</tr>
<tr>
<td>19</td>
<td>General scheme for the incorporation of 1 into a protein, conjugation, and UV-controlled release</td>
<td>37</td>
</tr>
<tr>
<td>20</td>
<td>Illustration of general antibody structure</td>
<td>37</td>
</tr>
</tbody>
</table>

**Table 1.** Codon table and the corresponding amino acids 4

**Table 2.** Well coordinates of expression products of different aaRSs with and without UAA supplementation 23
Abstract

Unnatural amino acids (UAAs) contain side chains, or R groups, that are not found in the 20 canonical amino acids. These noncanonical groups afford the capability to incorporate powerful chemical capabilities in proteins that are ordinarily unavailable with the naturally-occurring amino acids. Among the most useful moieties to incorporate into proteins are functional groups that can undergo Huisgen [3+2] cycloadditions, or “click,” reactions. This reaction occurs between azides and alkynes, and its mild conditions and high regioselectivity and reactivity make it an ideal process for bioconjugation. Photoreactivity is another useful characteristic that can be conferred to UAAs. Photolabile caging groups can inhibit the function of a protein until brief irradiation with UV light induces an intramolecular rearrangement and its displacement, reestablishing normal function. In this thesis, we propose a synthesis to incorporate both of these moieties into a single UAA.
Background

Twenty canonical amino acids comprise the core of all proteins in living organisms. An enormous diversity in both structure and function is achieved across the expansive proteome despite a relatively miniscule pool of building blocks. Individual amino acids are polymerized to proteins via a dehydration reaction of the amine group of one amino acid with the carboxylic acid group of another to form an amide bond, also known as a peptide bond. The difference between each of the amino acids is the side chain, or R group. These groups contain a range of chemical functionalities including nonpolar substituents, such as aromatic rings and alkyl chains, polar substituents, such as hydroxyl and amine-containing moieties, as well as basic and acidic substituents (Fig 1).

The sequence of amino acids dictates the protein’s overall structure, as the protein adopts a conformation that minimizes the Gibbs free energy. Non-covalent interactions between side chains and other interactions involving the peptide backbone drive the formation of secondary structure of the protein. The most common of these structures, α-helices and β-sheets, form from hydrogen bonding between a lone pair on the peptide carbonyl and the amide hydrogen. The amino acids that comprise α-helices tend to contain straight-chain R groups, such as methionine and alanine, whereas β-sheets form with bulky amino acids like tryptophan or phenylalanine. Not all of the protein adopts a secondary structure, however, and instead forms the “irregular structure.”
Secondary and irregular structures are entropically driven to fold into the tertiary structure such that hydrophobic-hydrophilic interactions are minimized. For most proteins, the environment is predominately aqueous, so the nonpolar amino acids are sequestered to the interior of the proteins. Approximately 25% of all proteins are bound to the cell membrane, resulting in these proteins positioning hydrophobic side chains to the exterior surface as a mechanism to interact with the nonpolar aliphatic chains of the phospholipids. These proteins
also have regions that are not exposed to plasma membrane, and therefore can have polar amino acids exposed on the surface to interact with water.

Proteins are products of gene expression. In this process, DNA is transcribed to messenger RNA (mRNA) which provides the template for the amino acids to be linked together to form a polypeptide. The template is read in groups of three nucleotides called codons, each of which corresponds to either an amino acid or for a translational stop. Codons are translated by transfer RNAs (tRNAs), which are molecules with two important functions. Firstly, tRNAs contain an anticodon, which is the complement to a particular codon. The second important function of the tRNA is that it binds a specific amino acid. Amino acids are covalently bound, or charged, to a tRNA by an aminoacyl-tRNA synthase (aaRS). Each amino acid has one corresponding aaRS, even though in some cases there are degenerate codons that encode the same amino acid since there are 64 unique codons, but only 20 amino acids (Table 1).

Table 1: Codons and their associated amino acids

http://www.biogem.org/codon.jpg
Translation of proteins is catalyzed on the ribosome, a large two-unit complex of both ribosomal RNA (rRNA) as well as protein.\(^1\) There are three main sites for protein synthesis in the ribosome: the exit, peptidyl, and aminoacyl sites. These are commonly abbreviated as the E, P, and A sites. As the mRNA is processed through the ribosome, charged tRNAs enter the complex and form a complementary hydrogen bond to the template mRNA in the A site. The ribosome then catalyzes the nucleophilic attack of the A amino acid’s amine to the carboxylic acid end of the amino acid in the P site. The condensation reaction dissociates the peptide from the P site tRNA to the A site, growing the peptide chain by one amino acid residue. Immediately following the bond formation, the P site tRNA moves to the E site, the A tRNA moves to the P site, and a new charged tRNA can enter the vacant A site. The discharged E site tRNA is ejected every time a peptidyl transfer occurs, upon which time it is almost immediately recharged with another amino acid. Translation is terminated when a stop codon on the mRNA reaches the A site and a release factor binds to it, causing the dissociation of the ribosomal subunits (Fig 2).

A common technique in biology to identify the function of a biological component is to determine the consequences of its elimination or alteration from the system. Oftentimes, if a protein is removed the resulting consequence will elucidate the function of the protein. Instead of completely removing a protein, another technique that has surfaced in recent years is the utilization of unnatural amino acids (UAAs) to modify the function of a protein. UAAs defy the limit nature has placed on biology by expanding the chemical functionality accessible to proteins beyond the 20 canonical amino acids. The ability to introduce novel side chains containing such functionalities as metals, alkynes, fluorophores, and others, allows for new methods to study proteins.\(^2\)
Figure 2: Diagram of the initiation, elongation, and termination steps of translation

The first challenge to overcome is to charge tRNA molecules with UAAs. Nature produces remarkably few errors in charging the canonical amino acids to their corresponding tRNAs. As such, a solution is to introduce both orthogonal tRNAs and a corresponding aaRS. Orthogonality of these components is achieved by introducing them from an exogenous source, such as *Methanococcus jannaschii*. The difference in species source prevents cross-reactivity with the naturally-encoded aaRSs and tRNAs in *E. coli*.

The second major challenge is the development of an aaRS than can recognize and charge a UAA to the tRNA. UAAs can contain unique structures that are unfamiliar to natural aaRS proteins, so directed evolution must be performed to develop an aaRS. A useful aspect to evolving an aaRS is the ability to modulate its specificity. Natural aaRSs have evolved to be highly selective to a single amino acid, and are able to discriminate features as minute as the single methylene difference between glutamate and aspartate. Due to the exogenous nature of the UAA, there are fewer selective pressures in the development of unnatural aaRSs, leading to promiscuity in the UAAs it might recognize. This modulation allows related UAAs to be aminoacylated to a tRNA by the same aaRS.

UAAs can be incorporated into proteins in *E. coli* at site-specific locations by taking advantage of rarely used stop codons. There are three stop codons found in nature, and often codon bias towards one of the three exists within an organism’s genome. In *E. coli*, the UAG stop codon is rarely used. Since this codon is present in few genes, there is a correspondingly low concentration of release factors that bind this sequence to terminate translation. These conditions allow for minimal competition for the orthogonal tRNAs that have been engineered to contain the complementary anticodon to the UAG codon (Fig. 3). Introducing engineered plasmids that contain genes encoding an orthogonal aaRS and its cognate tRNA, antibiotic
resistance, an exogenous protein, and other sequences such as promoters to bacteria facilitates the incorporation of UAAs into proteins.\textsuperscript{4,6}

Figure 3: Illustration of UAA incorporation\textsuperscript{7}

The selection process for developing aaRSs is twofold (Fig. 4). First, a positive selection is performed by transforming cells with a plasmid encoding an aaRS as well as a mutated gene encoding chloramphenicol acetyltransferase\textsuperscript{7} (CAT) with a TAG codon, an enzyme that eliminates the antibiotic properties of chloramphenicol by catalyzing its acetylation. The mutation introduces the codon that is complementary to the orthogonal UAA tRNA. The cells are grown in the presence of the chloramphenicol, so only the cells that possess an aaRS that can charge an amino acid to the tRNA will produce CAT and survive. The aaRS has been mutated at the tRNA binding site to recognize the anticodon AUC as well as at the amino acid binding site to provide the possibility for UAA recognition. The mutants that survive this selection are then
subject to negative selection. In this process, the cells are transformed with the same aaRS as well as a plasmid encoding a mutated barnase gene.\textsuperscript{7} Barnase is a ribonuclease that is toxic to cells if expressed. The cells are grown in the absence of UAAs to select against aaRSs that aminoacylate natural amino acids, since barnase will only be expressed if the aaRS can read through the TAG codon. The selection process is repeated multiple times to ensure that the aaRS is evolved to truly recognize the desired UAA.\textsuperscript{7}

![Diagram](image)

Figure 4: Scheme for aaRS evolution\textsuperscript{7}

One of the simplest ways to test for a UAA’s incorporation is to express it in green fluorescent protein (GFP), a well characterized reporter protein (Fig. 5) that was discovered in jellyfish and emits green light upon irradiation with UV light.\textsuperscript{1} The UV activity makes this protein quite easy to find in cells and tissues, so it is often used as a fusion tag to other proteins of interest in other fields of research. The complementary DNA from which the protein is translated has been sequenced and the protein itself can be expressed under simple conditions,
making it highly attractive for use in basic research. The protein’s DNA can be altered to contain a nonsense codon anywhere a UAA is desired to be encoded. Since the two main outcomes of a UAG codon in an mRNA are UAA incorporation and translational termination, presence of GFP indicates a successful expression with the UAA. Translation has also been shown to proceed through the less-used stop codons in cells in the absence of UAAs, so the presence of GFP must be analyzed to ensure that it is not the product of this background expression.

Figure 5: GFP structure; within the β-barrel is a unique fluorophore that confers its spectroscopic properties

Click reactions in a biological context

The azide-alkyne cycloaddition, commonly referred to as the “click” reaction, was first developed by Huisgen in the 1960s under thermal conditions to produce a mixture of 1,4- and 1,5-triazoles. It was not until 2002 when Sharpless discovered that the use of a Cu(I) catalyst could afford 91% selectivity for the 1,4-substituted product under mild conditions (Fig. 6). He noted that the reaction was robust to many organic synthesis conditions, including in the presence of water and O₂, and was tolerant of many functional groups. For these reasons, along
with the fact that alkynes and azides are virtually nonexistent in biology, this reaction has become a prominent tool in the field of bioconjugation.  

Common examples of bioconjugates are various probes that allow for molecular imaging. These chromophores can be clicked to a number of biomolecules that are engineered to contain either an azide or alkyne functionality. For example, this technology has been used in mice for positron emission tomography (PET), a powerful medical imaging technique. In one such experiment, researchers were able to engineer the peptide A20FMDV2 which selectively binds integrin $\alpha_v\beta_3$, a protein that is present in abnormally higher concentrations in tumor cells. The peptide was modified by the addition of an azide to the N-terminus, which was reacted with an alkyne bearing three positron emitting isotopes. Excellent spatial and temporal resolution was achieved through this process.

![Figure 6: Mechanism of Cu(I) catalyzed click reaction](image)

Another advantage of the use of clickable UAAs is the ability to conjugate different labels to a protein in fluorescence resonance energy transfer (FRET) studies. Click conditions
are facile and do not interfere with cysteine-labeling conditions. The thiol group of cysteine is commonly reacted irreversibly with maleimides in protein chemistry (Fig. 7). This is because cysteine residues are present in many proteins, albeit in low numbers, and the thiol group makes it the most nucleophilic amino acid. The low abundance of cysteine makes its conjugation to fluorophores or other molecules a highly specific process. When both cysteine-maleimide and azide-alkyne conjugation can be achieved in the same protein, conditions for FRET experiments can be prepared with ease. The use of clickable UAAs in proteins allows for high specificity in bioconjugation, providing well-defined conjugates. Bioconjugation with any of the canonical amino acids is imprecise in both location and number of conjugates formed, since there are multiple residues that can react.

![Figure 7: General mechanism of cysteine-maleimide conjugation](image)

**Photocaging biological systems**

Caging biomolecules, such as proteins, with a photolabile protecting group can inhibit the function of the protein until it is decaged by brief irradiation with UV light. Nitro groups readily undergo photochemical reactions by promoting radical formation. When 2-nitrobenzyl groups are subject to UV light, one of the oxygen atoms will migrate to the neighboring alkyl groups, leaving a nitroso group (Fig. 8). If this moiety was an R group of an amino acid chain, then the nitroso could further react with the amide nitrogen to form a 6-membered ring. After
rearrangement, hydrolysis releases the newly formed ring and the peptide is cleaved in two (Fig. 9).\(^{15}\)

**Figure 8: General decaging mechanism\(^{14}\)**

**Figure 9: Mechanism of photoinitiated peptide cleavage\(^{12}\)**
The inhibition can occur by the incorporation of a caged UAA, such as ONBY (Fig. 10), into the active site of a protein. This method has been performed on the RNA polymerase of bacteriophage T7 (T7RNAP).\textsuperscript{16} T7RNAP was mutated at Tyr639, a necessary component of the active site, with ONBY. Upon irradiation with UV light, the ortho-nitrobenzyl group was cleaved, activating the protein by revealing a tyrosine residue.\textsuperscript{16}

![Figure 10: (O-nitrobenzyl)tyrosine (ONBY)](image)

The goal of this study was to synthesize novel unnatural amino acids (UAAs) bearing substituents that afford unique chemical capabilities in biological systems upon incorporation into proteins. UAA 1 (Fig. 11) is a tyrosine derivative with two notable functional groups- an ortho-nitrobenzyl group and a para-propargyl group. The propargyl group provides a handle on which to perform “click” reactions with azide-containing substrates. The ortho-nitrobenzyl substituent affords the ability to cleave the peptide backbone, freeing the UAA and its conjugate upon UV irradiation.

![Figure 11: UAA 1](image)
Other UAAs were synthesized, most of which were propargyl-modified tyrosine derivatives like 1. UAA 10 (Fig. 12) was made possessing a propargyl group at the *para* position. Another UAA was attempted to be prepared with an ethylene linker between the phenolic oxygen of the tyrosine and a propargyl ether (Fig. 13). The latter UAA was intended to be one of several UAAs with linkers of varying length to either azides or alkynes in order to distance the reactive moiety from the core of the protein, but the products of early synthetic steps were not recovered.

![Figure 12: UAA 10](image)

![Figure 13: Proposed UAA with an ethylene-extended propargyl group](image)
Experimental

General. Solvents and reagents were obtained from either Sigma-Aldrich or Fisher Scientific and used without further purification, unless noted. Reactions were conducted under ambient atmosphere with un-distilled solvents, unless otherwise noted. Microwave reactions were conducted in a CEM Discover microwave reactor. NMR data was acquired on a Varian Gemini 400 MHz. Escherichia coli BL21(DE3) cell lines were obtained from Novagen. The various plasmids utilized during the transformations were obtained through either collaborators or Novagen. Microtiter plate data was acquired on a Wallac Victor2 1420 multilabel counter.

Synthesis of a Clickable Photoreactive Unnatural Amino Acid

(3-methoxy-4-(prop-2-yn-1-ylxy)benzaldehyde), 2: Vanillin (3 g, 20 mmol, 1 eq) was stirred with K₂CO₃ (8.29 g, 60 mmol, 3 eq) for 5 minutes in DMF (15 mL). The mixture was cooled to 0 °C in an ice bath, followed by slow addition of propargyl bromide (3.28 mL, 3.56 g, 30 mmol, 1.5 eq). The reaction was warmed to room temperature and stirred for 16 hours. The reaction was quenched by the addition of H₂O (30 mL) and extracted with EtOAc (3x40 mL) and back extracted with brine (3x80 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (5:1 Hexanes: EtOAc, then 3:1 Hexanes: EtOAc) yielding a white solid (2.88 g, 15 mmol, 77% yield).¹⁷

(5-methoxy-2-nitro-4-(prop-2-yn-1-ylxy)phenyl)methanol, 3: 2 (0.244 g, 1.3 mmol, 1 eq) was wrapped in foil and placed in an ice bath at 0 °C. 70% HNO₃ (5mL, excess) was added slowly to the reaction, and it stirred for 10 minutes. The reaction was warmed to room
temperature and stirred an additional hour. The contents of the reaction were added to 0 °C H₂O (70 mL) and collected by vacuum filtration. Ice water (3x10 mL) was used to rinse the reaction vessel and collect the residual product. The product was extracted from the filter paper by rinsing with a small volume of DCM, which was removed by rotatory evaporation. The resulting bright yellow solid (0.220 g, 1.0 mmol, 77% yield) required no further purification.

¹H NMR (400MHz; CDCl₃): δ 10.45 (s, 1H), 7.79 (s, 1H), 7.43 (s, 1H), 4.90 (s, 2H), 4.02 (s, 3H), 2.62 (s, 1H)

¹³C NMR (400MHz; CDCl₃): δ 187.93, 153.94, 146.68, 140.22, 126.64, 110.48, 109.62, 78.00, 76.56, 57.45, 57.02

(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)methanol. 4: 3 (28 mg, 0.119 mmol, 1 eq) was dissolved in 1 mL EtOH, wrapped in foil, and placed in an ice bath. NaBH₄ (14 mg, 0.357 mmol, 3 eq) was dissolved in 1M NaOH (1 mL). The NaBH₄ solution was then added to the solution of 3 dropwise over a period of 2 minutes. Once all the reducing agent was added, it was removed from the ice bath and stirred at room temperature for 45 minutes. The reaction was quenched by neutralizing the reaction with 0.6M HCl and extracted in DCM (3x15 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting pale yellow solid (20 mg, 0.084 mmol, 71% yield) required no further purification.

¹H NMR (400MHz; CDCl₃): δ 7.89 (s, 1H), 7.22 (s, 1H), 4.98 (s, 2H), 4.84 (s, 2H), 4.01 (s, 3H), 2.58 (s, 1H), 1.56 (s, 1H)

¹³C NMR (400MHz; CDCl₃): δ 153.94, 146.68, 140.22, 126.64, 111.09, 78.00, 76.56, 63.02, 57.26, 29.94
1-(bromomethyl)-5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)benzene, 5: 4 (20 mg, 0.084 mmol, 1 eq) was dissolved in DCM (<1 mL) and wrapped in foil. PBr₃ (8 μL, 0.023 g, 0.084 mmol, 1 eq) was added to the reaction vessel and stirred for 2 hours. The reaction was quenched by the addition of H₂O (5 mL) and extracted in DCM (3x15 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (3:1 hexanes: EtOAc), yielding a yellow solid (20 mg, 0.067 mmol, 79% yield).

¹H NMR (400MHz; CDCl₃): δ 7.85 (s, 1H), 6.98 (s, 1H), 4.87 (s, 2H), 4.84 (s, 2H), 4.01 (s, 3H), 2.60 (s, 1H), 13C NMR (400MHz; CDCl₃): δ 153.94, 146.68, 140.22, 128.68, 114.29, 111.26, 57.25, 56.80, 29.94

Diethyl 2-acetamido-2-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)benzyl)malonate, 6:

Diethyl acetamidomalonate (5 mg, 0.025 mmol, 1.5 eq) and Cs₂CO₃ (8 mg, 0.025 mmol, 1.5 eq) were stirred in CH₃CN (1 mL) for 5 minutes. 5 (5 mg, 0.017 mmol, 1 eq) was dissolved in CH₃CN (2 mL) and added to the diethyl acetamidomalonate solution in a microwave vial. The reaction was subjected to microwave irradiation (130 °C, 0-300 W, 10 minutes) and quenched by the addition of H₂O (5 mL). The reaction was extracted in DCM (4x15 mL) and back extracted with brine (4x80 mL), dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (1:1 hexanes: EtOAc), yielding a peach-colored solid (9 mg, 0.020, 80% yield).

¹H NMR (400MHz; CDCl₃): δ 7.69 (s, 1H), 6.76 (s, 1H), 6.52 (s, 1H), 4.80 (s, 2H), 4.33-4.13 (m, 4H), 4.10 (s, 3H), 3.90 (s, 2H), 2.59 (s, 1H), 1.94 (s, 3H), 1.29 (t, J = 7.2, 6H)
$^{13}$C NMR (400MHz; CDCl$_3$): $\delta$ 169.67, 167.83, 152.89, 145.75, 142.08, 126.11, 116.03, 115.95, 78.00, 76.56, 66.09, 63.05, 57.05, 56.52, 35.17, 14.02, 14.00

1-carboxy-2-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethan-1-aminium, 1: 7 (14 mg, 0.032 mmol, 1 eq) was stirred with 6M HCl (2 mL, excess) under microwave irradiation (90 °C, 10 minutes). The reaction was concentrated by rotatory evaporation, yielding a solid (9 mg, 0.030 mmol, 95% yield).

$^1$H NMR (400MHz; CDCl$_3$): $\delta$ 7.70 (s, 1H), 6.83 (s, 2H), 6.74 (s, 1H), 4.81 (s, 2H), 4.08 (s, 1H), 3.91 (s, 3H), 2.19 (s, 2H)

Cycloaddition of 6 with Sodium Azide

6 (0.010 g, 0.023 mmol, 1 eq), NaN$_3$ (0.002 g, 0.031 mmol, 1.3 eq), 1M CuSO$_4$ (7 $\mu$L), and 1M sodium ascorbate (17 $\mu$L) were wrapped in foil and stirred in DMSO (2 mL) and H$_2$O (2 mL) for 24 hours at room temperature. The reaction was extracted in DCM (3x10 mL), dried with Mg$_2$SO$_4$, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (1:1 hexanes: EtOAc), yielding a faintly yellow solid (0.010 g, 0.021 mmol, 67% yield).

Synthesis of Other Unnatural Amino Acids

Synthesis of p-propargyloxyphenylalanine (pPrF)

Methyl (S)-2-((tert-butoxycarbonyl)amino)-3-(4-(prop-2-yn-1-yloxy)phenyl)propanoate, 8: Methyl (tert-butoxycarbonyl)-L-tyrosine (0.500 g, 1.693 mmol, 1 eq) was stirred with Cs$_2$CO$_3$ (0.896 g, 2.539 mmol, 1.5 eq) and KI (catalytic amount) in DMF (5 mL) for 15 minutes.
Propargyl bromide (0.736 mL, 0.8056 g, 6.772 mmol, 4 eq) was added slowly. The reaction was allowed to proceed at 60 °C for 16 hours. The reaction was then quenched by the addition of H₂O (5 mL), extracted in DCM (4x15 mL) and back extracted with brine (4x120 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (3:1 hexanes: EtOAc), yielding a white solid (0.394 g, 1.182 mmol, 70% yield).¹⁸

(S)-2-((tert-butoxycarbonyl)amino)-3-(4-(prop-2-yn-1-yloxy)phenyl)propanoic acid, 9: 8 (0.394 g, 1.182 mmol, 1 eq) was placed in an ice bath and stirred with 1M NaOH and dioxane (1 mL ea) for five minutes. The reaction was warmed to room temperature and stirred for 1 hour. The reaction was concentrated by rotatory evaporation, and the remaining aqueous layer was acidified (pH 4) with 1M HCl and extracted with EtOAc (3x10 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation, yielding a white solid (0.117 g, 0.366 mmol, 31% yield).¹⁸

(S)-2-amino-3-(4-(prop-2-yn-1-yloxy)phenyl)propanoic acid, 10: 9 (0.117 g, 0.366 mmol, 1 eq) was stirred with 1% TFA (2 mL) on ice for 5 minutes. The reaction was warmed to room temperature and stirred for an additional 1 hour. The solution was concentrated by rotatory evaporation, yielding a white solid (0.080 g, 0.366 mmol, 100% yield).¹⁸

Synthesis of a Clickable Unnatural Amino Acid with an Extended Linker

Methyl (S)-3-(4-(2-bromoethoxy)phenyl)-2-((tert-butoxycarbonyl)amino)propanoate, 11:
Methyl (tert-butoxycarbonyl)-L-tyrosine (0.122 g, 0.413 mmol, 1 eq) was dissolved in DMF (3 mL) and stirred with Cs₂CO₃ (0.219 g, 0.620 mmol, 1.5 eq) for 15 minutes. Dibromoethane (71 μL, 0.154 g, 0.826 mmol, 2 eq) and KI (catalytic amount) were added to the solution and stirred for 24 hours at 60 °C. The reaction was quenched by the addition of H₂O (10 mL) and extracted in DCM (3x15 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation, yielding a white solid that was not the desired product, 11.

3-(2-bromoethoxy)prop-1-yne, 13: Propargyl alcohol (23 μL, 0.022g, 0.0387 mmol, 1 eq) was stirred with t-BuOK (0.192 g, 1.16 mmol, 3 eq) for 15 minutes in DMF (5 mL). 1,2-dibromoethane (100 μL, 0.218 g, 1.16 mmol, 3 eq) was added to the reaction and stirred for 24 hours at room temperature. The reaction was quenched by the addition of H₂O (10 mL) and extracted in DCM (3x15 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation, yielding a white solid that was not the desired product, 13.

Synthesis of a Malonate-Derived Unnatural Amino Acid

Methyl (S)-2-((tert-butoxycarbonyl)amino)-3-(4-iodophenyl)propanoate, 14: (S)-2-((tert-butoxycarbonyl)amino)-3-(4-iodophenyl)propanoic acid (0.100g, 0.256 mmol, 1 eq) was stirred with NaHCO₃ (0.060 g, 0.767 mmol, 3 eq) in DMF (2 mL) for 10 minutes. CH₃I (14 μL, 0.0319 g, 0.282 mmol, 1.1 eq) was added to the solution and stirred at 60 °C for 48 hours. The reaction was quenched by the addition of H₂O (10 mL) and extracted in EtOAc (4x10 mL) and back extracted with brine (4x100 mL). The organic layer was dried with Mg₂SO₄, filtered, and concentrated by rotatory evaporation. The resulting oil was purified by flash chromatography (1:1 hexanes:EtOAc), yielding a white solid (0.070 g, 0.172 mmol, 67% yield).
Diethyl (S)-2-(4-((tert-butoxycarbonyl)amino)-3-methoxy-3-oxopropyl)phenyl)malonate, 15: 14 (0.010 g, 2.468 mmol, 1 eq) was stirred with diethyl malonate (5 μL, 0.005 g, 2.96 mmol, 1.2 eq), CuI (0.001 g, 0.617 mmol, 0.25 eq), L-proline (0.001 g, 0.987 mmol, 0.4 eq), and Cs₂CO₃ (0.032 g, 9.871 mmol, 4 eq) in DMSO (1.5 mL) under argon (1 atm) for 16 hours at 40 ºC. The reaction was quenched by the addition of H₂O (5 mL), extracted in EtOAc (3x5 mL) and back extracted with brine (3x50 mL). The organic layer was dried with MgSO₄, filtered, and concentrated by rotatory evaporation; however, the desired product, 15, was not recovered.

Assessing the Expression of GFP by the Insertion of UAA1

A pET-GFPTAG151 plasmid and one of seven plasmids (pEVOL-ONBY, pEVOL-pPrF, pEVOL-Ac32, pEVOL-CN3265, pEVOL-pAzoF, pEVOL-Bipy, pEVOL-Ambrx; 0.5 μL each) were each transformed to competent BL21(DE3) cells (50 μL; Novagen). The cells were transformed by electroporation (1.8 V, 1s), then recovered in 2xYT media (200 μL) at 37 ºC for 1 h. After plating 75 μL on LB agar containing ampicillin (50 μg/mL) and chloramphenicol (50 μg/mL), the cells were incubated overnight. A single colony from each plate was then used to inoculate 2xYT containing ampicillin (50 μg/mL) and chloramphenicol (50 μg/mL) (4 mL). After overnight incubation at 37 ºC, the starter cultures (500 μL) were used to inoculate 2xYT media containing ampicillin and chloramphenicol (10 mL ea.). Each starter culture was used to inoculate two expression cultures. These expression cultures were incubated with shaking at 37 ºC until the OD₆₀₀ reached 0.78. At this point, the expression cultures were induced with 1M IPTG (10 μL) and 20% arabinose (10 μL). One of each of the pairs of expression cultures was supplemented with 1 (100 μL, 100 mM). The cultures were then grown for 16 hours at 30°C.
Both expression cultures were pelleted at 5000 rpm for 10 minutes, and then frozen at -80 °C for 24 hours. The cell pellets were resuspended in Bug Buster (500 µL; Novagen) supplemented with lysozyme and incubated at room temperature for 20 minutes to lyse the cells. The cell lysate (100 µL) was assessed for fluorescence on a microtiter plate (Table 2).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>ONBY +</td>
<td>ONBY -</td>
<td>pPrF +</td>
<td>pPrF -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Ac32 +</td>
<td>Ac32 -</td>
<td>Bipy +</td>
<td>Bipy -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>pAzoF +</td>
<td>pAzoF -</td>
<td>Ambrx +</td>
<td>Ambrx -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>CN3265 +</td>
<td>CN3265 -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Well coordinates of expression products of different aaRSs with and without UAA supplementation

**Expression of GFP by the Insertion of UAA1**

A pEVOL-Bipy plasmid and one of four plasmids (pET-GFPTAG3, pET-GFPTAG133, pET-GFPTAG151, pET-GFPWALDOTAG151; 0.5 µL each) were added to competent BL21(DE3) cells (50 µL; Novagen) and transformed, inoculated, and induced as previously described. The cell lysate was centrifuged at 13,000rpm for 10 minutes. The supernatant was then added to the QIAGEN® Ni-NTA Spin Column, and the resulting proteins were purified according to the manufacturer’s protocols. Native PAGE (10%) was performed on each of the native GFP mutants. The gel was subject to UV irradiation to visualize the fluorescent bands and stained with Coomassie Blue to visualize all protein bands.

**GFP Click Reaction**

The GFP mutants (30 µL), N$_3$-PEG (1 µL), CuSO$_4$ (5 µL, 1 mM), TCEP (4 µL, 2mM), and TBTA (1 µL, 5 mM) were wrapped in foil and agitated at 190 rpm for 48 hours. The reaction was assessed by native PAGE.
Results and Discussion

Synthesis of Unnatural Amino Acids

The synthesis of 1 proceeded according to the planned route (Fig. 14) with minor alterations to conditions for most steps, but the syntheses of 4 and 6 required more substantial changes. After optimization of the synthesis twice with small amounts of vanillin (0.75 g), a large scale synthesis was attempted using 3 g of vanillin starting material to obtain significant quantities of the desired unnatural amino acid.

Figure 14: Synthetic scheme for UAA 1

The initial reaction of the synthesis involved propargylation of vanillin (2). This nucleophilic substitution reaction generated a nucleophilic phenoxide which reacted with the
electrophilic propargyl bromide. Overall the reaction proceeded to near completion and required little optimization. Following an extraction and purification by column chromatography, 2 was obtained in 77% yield. The reaction was performed in DMF, which was not dried at the time of utilization. The yield could likely be improved by the use of anhydrous solvent.

The propargylated vanillin (2) was next reacted with nitric acid to install a nitro group, effectively generating the photoreactive moiety. The reaction proceeds by electrophilic aromatic substitution (Fig. 15). The electrophilic NO$_2^+$ cation is generated by acid-catalyzed loss of H$_2$O. The nucleophilic aromatic ring of 2 reacts with the electrophile, and a phenyl proton is abstracted by water to reestablish aromaticity. This step of the synthesis proceeded to completion at small scales, but proved to be somewhat problematic upon scaling up the reaction conditions. Primarily, the lack of complete nitration was especially challenging. This incompleteness was not realized until the following step where the aldehyde was reduced to an alcohol when an NMR spectroscopic analysis on the crude product of 4 was performed. This was remedied by resubmitting the non-nitrated alcohol to the same nitration conditions, which not only installed the nitro group to the correct position, but also oxidized the alcohol to yield 3 (Fig. 16).

![Figure 15: Mechanism of nitration](image-url)
Relatively significant difficulties arose with the reduction of 3 to 4. This reduction occurs by the attack of a hydride ion at the electrophilic aldehyde carbon, generating a tetrahedral alkoxide intermediate. Following acid workup, extraction, and column chromatography, 4 is recovered. In early attempts, the reduction was performed by adding NaBH₄ to the solution of 3 in EtOH at neutral pH. The yields were not optimal (3, 0.220 g, 1.5 eq NaBH₄ → 4, 0.083 g, 37%), but there was enough 4 retrieved to continue with the synthesis. The residual starting material recovered and subjected to the reaction conditions again in similar yields. It was thought that the NaBH₄ was rendered less potent as a result of reacting with atmospheric water, so it was used in greater excess with longer reaction times, and different sources of reducing agent were utilized in attempts to drive the reaction to completion. It was ultimately realized that the hydride ions were most likely reacting with the protons on ethanol. Consequently, the reaction was prepared in 1M NaOH and the reduction was much more successful, nearly doubling the yield from 37% to 71%.

The bromination of 4 to 5 occurs by an activated nucleophilic substitution reaction. First, the nucleophilic alcohol reacts with PBr₃, displacing a Br⁻ ion (Fig. 17). The resulting phosphorus ester is a good leaving group, and the Br⁻ ion reacts via an S_N2 manner to form a C-Br bond. The phosphonic dibromide is a better leaving group than Br⁻, so equilibrium favors the brominated product.
The reaction of 5 to 6 required the modification of conditions. This step of the synthesis tended to be performed and was followed by periods of inactivity, so it was unclear whether the product was degrading over time or if the reaction was simply not working. The reaction conditions that were used most often began by stirring NaH (0.003 g, 0.137 mmol, 2 eq) with diethyl acetamidomalonate (0.015 g, 0.103 mmol, 1.5 eq) for 5 minutes in DMF (1 mL) in a flame-dried vial. 5 (0.023 g, 0.069 mmol, 1 eq) was dissolved in DMF (1 mL) and added to the reaction and the reaction vessel was wrapped in foil (16 h, 60 °C). These conditions were attempted four times, but 6 was never produced. Under these conditions, the reaction should occur by nucleophilic attack of the malonate derivative on the electrophilic bromine-bearing carbon. A carbanion is readily formed by the abstraction of a proton from diethyl acetamidomalonate by the hydride ion, but its nucleophilicity is mitigated by steric hindrance. The electrophile is a primary carbon, but the presence of the bulky phenyl ring at the alpha position further impedes the progress of the reaction. Consequently, based on literature precedence\textsuperscript{20}, the reaction was attempted under microwave irradiation. Instead of using NaH to abstract a proton from the malonic ester, Cs\textsubscript{2}CO\textsubscript{3} was used to address safety issues that could arise from heating NaH to such high temperatures. Acetonitrile was used as the solvent instead
of DMF, since it can tolerate large amounts of microwave energy. This allowed the reaction to be heated to 130 °C, well above the conditions that were previously attempted. This reaction was irradiated at this temperature for 10 minutes, providing a yield of 80%. The success of these conditions is likely due to the increased input of energy that allowed the substitution reaction to overcome steric hindrance and the high-energy transition state.

The conversion of 6 to 1 occurred under microwave irradiation in 6M HCl. Compound 6 had limited solubility in HCl, but as the deprotection of the malonic ester progressed, the more water soluble compound, 1 was produced. This led to the solvation of some of the product, revealing the interior of the insoluble granules of starting material. 1 was still not particularly soluble, but the reaction was determined to have run to completion by NMR spectroscopic analysis. 1 could not be purified by column chromatography since the amino acid backbone would bind quite strongly to the polar silica gel. Another notable result is the lack of stereoselectivity in the deprotection; the reaction presumably results in a racemic mixture of the (D) and (L) diastereomers. Only (L) amino acids can be translated by ribosomes, which implies that only about half of the quantity of 1 produced can be incorporated into proteins.

**Incorporation of the Synthesized Unnatural Amino Acid**

With the synthesized amino acid in hand, the next step involved the incorporation of the UAA into a protein. GFP was used in this study as a model system to determine whether or not the UAA was incorporated into the protein. A successful translation involving the UAA would lead to a functional fluorescent protein product, whereas in the absence of the UAA, GFP would be expressed at much lower levels or not at all. This is due to the fact that the UAA is introduced in relatively high concentrations, and the pEVOL vector that is used to transform the bacteria
contains genes that encode the tRNA with the UAG anticodon as well as two copies of the aaRS that charge the UAA to the exogenous tRNA. In the absence of the UAA, bacterial release factors will bind to the stop codon on a more frequent basis, terminating translation of GFP.

Typically, for each unnatural amino acid incorporated, an aaRS must be evolved that recognizes the UAA and charges the corresponding tRNA. However, due to the tedious nature of aaRS selections, we first attempted to exploit the inherent promiscuity of previously evolved aaRSs. Identifying an aaRS that could aminoacylate 1 to the orthogonal tRNA suppressor required the analysis of several candidates, since each aaRS in the laboratory was evolved to incorporate a few structurally similar UAAs. The UAA 1 contains several different substituents that taken singularly have a previously evolved aaRS, such as pPrF, so an array of aaRSs was assessed. Other aaRSs were selected for other properties, such as their incorporations of large aryl substituents like BipyA (2,2'-bipyridylalanine). BipyA, Ac32, and pPrF showed modest increases in fluorescence in samples supplemented with 1 (Fig. 18). Other aaRSs had stronger fluorescence in wells lacking 1, indicating that the UAA is not incorporated with those synthetases, and that it may be toxic. The toxicity might kill the cells at a rate such that GFP is more likely to be expressed by the incidental incorporation of natural amino acids to the UAG site in the control cells. The wells containing 1 that had a stronger fluorescent signal than the respective -UAA control indicate that the UAA is being successfully incorporated. Ultimately, BipyA showed to be the most efficacious aaRS based on results from PAGE analysis.
Figure 18: Table and graphical depiction of fluorescent intensities of different aaRSs
Different GFP genes were used that replaced wild type codons with nonsense codons at residue sites 3, 133, and 151 to find the optimal site for conjugation. The UAG codon in GFP3 is near the C terminus residue, 133 is in one of the loops that extends out from the barrel, and 151 is on the barrel itself. These sites are used because they are not in the fluorophore, but are important in the maintenance or creation of it. GFP 3 is a useful model, because no part of the protein is expressed in the absence of UAA. GFP 133 and 151 both contain the nonsense codon downstream of the fluorophore, but if translation is terminated at those sites, the fluorophore cannot be formed. The formation of the fluorophore is necessarily kept to the interior of the protein away from the aqueous environment. Each of the three mutant sites is surface exposed, though the local environment differs between the three. Testing the efficacy of the photoinitiated cycloaddition by the ortho-substituted nitro group can be assessed in the 133 and 151 variants, but not GFP3. This is simply due to the fact that the reaction cleaves the peptide in two, and one of the products of this process in GFP3 would be a PAGE-undetectable tripeptide. Incorporation of 1 to each of these sites should theoretically expose the alkyne to click reaction conditions.

**Optimization of Click Conjugation**

Initial work involved the click between 6 and NaN₃, outside of the context of a protein in the presence of CuSO₄ and sodium ascorbate. Based on the success of this reaction (determined by NMR spectroscopy), the conditions were translated to a protein system with the various GFP mutants containing UAA 1. Unfortunately, these results were somewhat inconclusive. Instead of clicking with NaN₃, the protein (30 μL) was clicked with N₃PEG (1 μL, 5kD), a polymer of known weight that can be analyzed via PAGE. Other changes in conditions included the use of TCEP (4 μL, 2mM) and TBTA (1 μL, 5 mM). Proteins have more steric hindrance that may
lead to difficulty in setting up the geometry needed for the click reaction to proceed. Furthermore, Cu(I) catalyzes the formation of reactive oxygen species, which are reactive to proteins and may cause complications. Moreover, the protein was clicked with a 5kD PEG-azide, resulting in a difference in size that is within the range of error of low percentage PAGE when a slightly uneven interface is present between the stacking and resolving layers of the gel. Consequently, more analyses must be performed to determine the success of this reaction in a protein context.

**Photocleavage of the Unnatural Amino Acid**

The capacity of the ortho-nitrobenzyl group of UAA 1 to undergo photoinitiated cleavage has been observed at various points throughout the synthesis. The reactive nitro group was installed in the second reaction (3), and from that point onward, the photoreaction product was readily observed on TLC plates as a brown-yellow spot after irradiation with 365 nm light for 30 seconds. Unfortunately, the reaction has not been observed in GFP thus far. UV light not only interacts with the nitro group of the UAA, but it can also cause other undesired ionizing effects on the protein and lead to numerous undesired side products. When the protein containing 1 was irradiated with 365 nm light for 10 minutes and analyzed by PAGE, there were no distinct bands that indicated the peptide had successfully undergone this reaction. We would expect for two distinct lighter bands to be present, each representing one of the two portions of the cleaved protein.
Synthesis of Other Unnatural Amino Acids

Synthesis of pPrF

The propargylation of protected tyrosine by propargyl bromide to yield 8 proceeded by the same mechanism as the reaction with vanillin. The yield was slightly lower for this reaction than the propargylation of vanillin (70% vs 77%). The slight difference could be a result of using a different source of propargyl bromide as well as wet DMF.

The hydrolysis of the methyl ester (8) to a carboxylic acid (9) by 1M NaOH and dioxanes resulted in a low yield, likely because of the middling polarity of the product and its subsequent loss to the aqueous layer in the extraction. This could be remedied by performing multiple extractions of smaller volume each time and using saturated brine.

Synthesis of a Malonate-Derived Unnatural Amino Acid

The commercially available starting material for this UAA (BocNH(Phe4-I)) is not protected at the carboxylic acid, so it was esterified to prevent side reactions in future steps. This was achieved by deprotonating the acid with NaHCO₃, a weak base. The carboxylate reacted with CH₃I via a nucleophilic substitution reaction. The reaction took two days of stirring at 60 °C to run to completion, possibly a result of dilute conditions. This substitution reaction is concentration-dependent, since the nucleophile and the electrophile must directly interact with each other, which could contribute to the slow rate at which the reaction proceeds. The protected phenylalanine derivative is then reacted with diethyl malonate to displace the para-iodide through nucleophilic aromatic substitution. Cs₂CO₃ abstracts the acidic hydrogen of diethyl malonate, generating a resonance-stabilized carbanion. This nucleophilic carbanion reacts with the carbon bonded to the iodide. This is not an S_N2 reaction, because the nucleophile would have
to add from inside the benzene ring to achieve the correct geometry to generate a Meisenheimer complex. Instead, the malonic ester adds to the iodide carbon, forming a negatively charged species. The iodide ion is a good leaving group, so it dissociates from the ring, leaving behind the desired product. The product of this reaction was not recovered. The transition state of this reaction is quite high, as it is difficult for the sterically hindered nucleophile to add to the electron-rich benzene ring. Furthermore, the equilibrium of this reaction is probably somewhat equivalent between the dissociation of iodide and malonate, making it difficult to drive to completion.
Conclusion

Several syntheses of UAAs were attempted, two of which were completed. The syntheses that were unable to be completed will require future efforts and modifications to the reaction conditions. The formation of the malonate-derived UAA proceeds by nucleophilic aromatic substitution with a hindered nucleophile, so the use of microwave irradiation is a potential candidate to help facilitate the reaction. The use of microwave irradiation in the reaction of 5 to 6 proved to be instrumental, which had analogous issues under thermal conditions. The synthesis of UAAs with propyl and butyl linkers to propargyl groups have been shown to be successful (Maza, McKenna, unpublished), but the reaction with dibromoethane may be yielding the elimination product instead of the substitution product. Changes to the conditions should divert the reaction from the elimination route.

The UAA pPrF has been used as a known substrate for the aaRS by the same name as a control against other UAAs that are candidates for aminoacylation by that synthetase. It has also been incorporated into proteins by other members of the lab and subjected to click reactions and Glaser-Hay coupling.

UAA 1 has been shown to be incorporated into proteins, but successful click reactions and photocleavage reactions have not been observed. Recent developments in the laboratory have indicated that effective click conditions involving proteins have been determined, however, so there is promise that this reaction can be observed soon. The photoreaction was performed by irradiating the mutant GFP for 10 minutes with UV light, which may have been too harsh of conditions. Decreased exposure time to irradiation will be used to determine the optimal duration of radiation that will lead to the peptidyl cleavage.
**Future Directions**

One of the overarching goals to the development of UAA technologies is the ability to treat diseases. The incorporation of UAA 1 into an antibody and conjugation of a cytotoxic agent through a click reaction (Fig. 19) could be useful in the treatment of human diseases, including cancers. Antibodies are immune proteins that have two variable regions that are highly discriminatory and bind tightly to a specific substrate (Fig. 20).\(^1\) Cancer cells express unique surface markers for which antibodies can be developed against. Conjugation to canonical amino acids on an antibody is not ideal since there is little control over how many residues will react and the location of those residues.\(^{13}\) If those reactive side chains are in the variable region, the function of the antibody is negated and no binding can occur. The site-specific incorporation of a UAA to the non-variable region of an antibody provides a way to overcome these two obstacles. With a single side chain alkyne in this region, one azide-containing drug can be conjugated without disrupting antibody function and the immunoconjugate can selectively bind cancer cells when introduced to the patient. The presence of the ortho-nitrobenzyl group provides temporal control for the release of the UAA and delivery of the drug once the antibody has been endocytosed by cells (Fig. 19).
Figure 19: General scheme for the incorporation of $I$ into a protein, conjugation, and UV-controlled release

Figure 20: Illustration of an antibody

Works Cited


Figures

5. http://www.chm.bris.ac.uk/motm/GFP/GFP-3structure.gif

Tables