2015

Synthesis and Incorporation of Unnatural Amino Acids To Probe and Optimize Protein Bioconjugations

Johnathan C. Maza  
*College of William and Mary*

Jaclyn R. McKenna  
*College of William and Mary*

Benjamin K. Raliski  
*College of William and Mary*

Matthew T. Freedman  
*College of William and Mary*

Douglas D. Young  
*College of William and Mary*

Follow this and additional works at: [https://scholarworks.wm.edu/aspubs](https://scholarworks.wm.edu/aspubs)

Recommended Citation  

This Article is brought to you for free and open access by the Arts and Sciences at W&M ScholarWorks. It has been accepted for inclusion in Arts & Sciences Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Synthesis and Incorporation of Unnatural Amino Acids To Probe and Optimize Protein Bioconjugations

Johnathan C. Maza, Jaclyn R. McKenna, Benjamin K. Raliski, Matthew T. Freedman, and Douglas D. Young*

Department of Chemistry, College of William & Mary P.O. Box 8795, Williamsburg, Virginia 23187, United States

Supporting Information

ABSTRACT: The utilization of unnatural amino acids (UAAs) in bioconjugations is ideal due to their ability to confer a degree of bioorthogonality and specificity. In order to elucidate optimal conditions for the preparation of bioconjugates with UAAs, we synthesized 9 UAAs with variable methylene tethers (2−4) and either an azide, alkyne, or halide functional group. All 9 UAAs were then incorporated into green fluorescent protein (GFP) using a promiscuous aminoacyl-tRNA synthetase. The different bioconjugations were then analyzed for optimal tether length via reaction with either a fluorophore or a derivatized resin. Interestingly, the optimal tether length was found to be dependent on the type of reaction. Overall, these findings provide a better understanding of various parameters that can be optimized for the efficient preparation of bioconjugates.

A growing area of interest in chemical biology is the ability to perform chemistry on biological systems, particularly conjugation reactions between bioactive molecules and either molecular probes or surfaces. These bioconjugates can be employed as therapeutic agents or diagnostic tools, and toward fundamental research.1−6 One of the most useful strategies for bioconjugation involves the formation of strong covalent linkages between the bioactive molecule and its partner (as opposed to processes such as adsorption or encapsulation).7,8 Due to the breadth of their applicability, there is a constant impetus for the optimization and development of methods for their preparation.

In order for such chemistry to be feasible, certain conditions must be met by the chemical reaction employed. Most importantly, the chemistry must be compatible with a biological setting. That is, 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 chemical functionalities encountered in biological systems to provide a degree of chemoselectivity.9 Should the conditions above be met, the chemistry is deemed bioorthogonal. While a range of bioorthogonal reactions have been elucidated, several issues preclude their widespread use.10 These bioorthogonal reactions include 1,3-dipolar cycloadditions, Sonogashira and Suzuki couplings, oxime ligations, photo-cross-linking reactions, and, most recently, Glaser-Hay couplings.11−18 One primary obstacle is the introduction of novel functionality into the biomolecule/protein; however, this can rapidly be accomplished via the site-specific incorporation of unnatural amino acids (UAAs).19 This novel chemical functionality can serve as a reactive handle to facilitate the generation of a stable bioconjugate in a well-defined fashion.

Multiple techniques exist for the introduction of UAAs that range from completely synthetic to exploitation of endogenous translational machinery.20−22 The evolution of orthogonal aminoacyl-tRNA synthetase/tRNA pairs has found a wide degree of success in bioconjugations and various other applications.22 Specifically, several examples of the use of unnatural amino acids to generate well-defined bioconjugates include the reaction of an azide containing calmodulin onto carbon nanotubes and nanoparticles via the Staudinger-Bertozzi ligation,23,24 the immobilization of other proteins to M-20 Dynabeads in a cell-free system with a click reaction,25,26 and the labeling of proteins with an agarose resin and azido UAA for secretome MS analysis (however, in this case the UAA was not site-specifically incorporated into the pool of proteins).27 More recently, we demonstrated the optimization of protein immobilization utilizing an incorporated azido UAA, conferring increased stability to organic solvents.28 The combination of these studies demonstrates the utility of site-specific bioconjugations, and the stabilizing benefits of the solid support. Other studies have used similar techniques to label proteins with molecular probes or generate novel therapeutic bioconjugates.22,29 Given the power of this technology, we aimed to expand the utility of UAA bioconjugations through the alteration of the reactive UAAs to optimize reaction parameters.

Received: July 28, 2015
Revised: August 18, 2015
Published: August 19, 2015

DOI: 10.1021/acs.bioconjchem.5b00424
Bioconjugate Chem. 2015, 26, 1884−1889
Initial investigations attempted to further utilize UAAs to optimize bioconjugations through several different bioorthogonal reactions. Our previous studies immobilizing GFP demonstrated that the UAA residue location on the protein dictated its immobilization efficiency.28 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. Moreover, removing the reactive partner from conjugation with the aromatic ring may alter its chemical reactivity. Thus, it becomes important to ascertain 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 can be derived from a modified tyrosine (Scheme 1). These functionalities afford access to the Cu(I)-catalyzed 1,3-dipolar cycloaddition, the Glaser-Hay coupling, and a nucleophilic substitution, respectively.

![Scheme 1. Synthesis of Tethered UAAs](image)

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 a S$_{N}$2 reaction (Scheme 1).

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. These bromo-tyrosines (1−3) would then have alkyl chains of varying lengths, and a good leaving group in future reactions. Then, a secondary S$_{N}$2 reaction with either an azide (4−6) or alkyne nucleophile would install the desired novel reactivity to the tyrosine derivatives. However, synthesis of the alkynyl-tyrosine series (7−9) proved more difficult than initially anticipated, and the series was prepared directly via S$_{N}$2 reactions with commercially available bromo-alkynes (see Supporting Information). Deprotection with LiOH followed by 2% TFA, afforded the 9 UAAs in moderate overall yields (see Supporting Information).

Typically with the new synthesis of UAAs a unique aaRS must be evolved to recognize and incorporate the UAA; however, based on the previous demonstration of the promiscuity of the pCNF-aaRS in *E. coli*, it was our hope that the structural similarity of the 9 prepared UAAs could facilitate incorporation without the need of a double sieve aaRS selection. With respect to the bromo-derivatives, a previous report used a similar approach to their incorporation in mammalian cells. Based on our previous studies we employed GFP as a model protein for incorporation due to its well-documented fluorescent properties facilitating rapid visualization and quantitation. Our previous findings indicated that residue 151 conferred the best conjugation efficiencies, and was thus selected for this study.28 BL21(DE3) *E. coli* were cotransformed with pEVOL-pCNF-aaRS and pET-GFP-151-TAG plasmids to introduce both the GFP protein and the translational machinery necessary for TAG suppression with UAAs. Following protein expression (30 °C, 16 h) in the presence or absence of the UAA, cells were lysed and GFP was purified via Ni-NTA chromatography. Gratifyingly, the application of the polyspecific aaRS was successful, as the 9 UAAs were able to be incorporated into GFP, albeit with differential expression levels (Table 1). The incorporation of the UAA was visualized both by fluorescence and SDS-PAGE analysis (see Supporting Information). As the pCNF synthetase was evolved to recognize pCNF, it was hypothesized that the p2yneY-, p3yneY-, and p4yneY-UAAs would be most easily incorporated due to the structural similarity between the cyano group and alkyne group in size and polarizability.

This result was observed, with the p3yneY (n = 2) producing a larger quantity of mutant-GFP, and the p4yneY producing the lowest quantity of mutant-GFP within the alkyne series. These qualitative observations make sense based on the evolution of the synthetase. We hypothesize that the increased tether distance may have more difficulty fitting into the binding pocket of the tRNA synthetase relative to smaller derivatives, preventing the UAA from being incorporated into the polypeptide chain. Thus, a trade-off 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. This trend was observed for all of the UAAs: as the tether length increased, the expression yield decreased.

### Table 1. Protein Expression Yields for the Tethered UAAs

<table>
<thead>
<tr>
<th>n</th>
<th>Br</th>
<th>N$_3$</th>
<th>C≡CH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.704 mg/mL</td>
<td>0.347 mg/mL</td>
<td>0.702 mg/mL</td>
</tr>
<tr>
<td>2</td>
<td>0.291 mg/mL</td>
<td>0.685 mg/mL</td>
<td>0.767 mg/mL</td>
</tr>
<tr>
<td>3</td>
<td>0.277 mg/mL</td>
<td>0.246 mg/mL</td>
<td>0.447 mg/mL</td>
</tr>
</tbody>
</table>
However, all 9 UAAs were capable of being incorporated using a single aaRS, not only obviating the necessity of a tedious selection, but also increasing the versatility of this approach. Given the ability to incorporate three different functionalities with variable tether lengths into GFP, we next sought to employ these new UAAs in bioorthogonal conjugation reactions. Studies were initiated with the azido-GFP series, as the cycloaddition reaction is the most common and well-understood bioorthogonal reaction. Moreover, the various tethers could be directly compared to the commonly utilized p-azidophenylalanine (pAzF) that has no methylene tether and contains the azido group in direct conjugation with the aromatic ring. We hypothesized that in addition to providing distance between the protein and the conjugation partner, the removal of the azido conjugation may also alter its reactivity. The various azido-containing GFP mutants were conjugated with a fluorescent AlexaFluor 488 alkyne dye using a CuSO4/TBTA/TCEP catalyst system based on previously reported conditions. Gratifyingly, all tethered derivatives could be conjugated to the dye, as determined by both fluorimetry and SDS-PAGE analysis (Figure 1). Control reactions were also performed in the absence of catalyst system to ensure that protein labeling was not due to nonspecific reactions or noncovalent association by the dye. These reactions yielded no fluorescent product following protein denaturation, and thus indicated that any fluorescence is indeed a result of the coupling to form the triazole ring. Interestingly, the p3AzY handle appears to have the best conjugation efficiency, compared to the more sterically hindered p2AzY and the more flexible p4AzY handles.

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.34 Based on previous UAA immobilization results, an epoxyc-activated sepharose resin was derivatized with propargyl alcohol in the presence of trimethylamine (TEA) by stirring overnight at room temperature.28 This afforded a resin with an alkyne functionality that could be employed as a reaction partner with the azido-GFP series.

It was expected that by extending the tethering distance of the functional handle of the UAA, and thus the site of bioorthogonal reactivity, the steric hindrance due to the protein’s bulk could be decreased and immobilization efficiency could be improved. As such, the p2AzY, p3AzY-, and p4AzY-mutated GFPs were exposed to copper click conditions (50 mM CuSO4, 50 mM TCEP, and 5 mM TBTA) at fixed concentration of GFP (roughly 0.250 mg/mL). As a positive control, the pAzF GFP, which has no tether between the aromatic ring and the azide functional group, was also subjected to the same copper click conditions. We have previously immobilized the p-azidophenylalanine mutated GFP on the alkylamine-functionalized sepharose resin using similar conditions.28 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 nonspecific interactions as the reason for immobilization rather than the selective copper click reaction. Results indicated that that copper click-based immobilization strategy was effective for all of the mutants in the series (Figure 2).

Figure 1. Expression and conjugation of GFP-151 mutants with an AlexaFluor488 dye. The Coomassie stained gel (top) indicates the successful incorporation of all three azides in the series as an appropriate band is observed for GFP at ∼26 kDa. The fluorescence of the same gels (bottom) indicates that 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, as no fluorescence is detected in the absence of catalyst.

Figure 2. Click immobilization of GFP-151 mutants on a sepharose resin with either an azide tethered tyrosine UAAs or a p-azidophenylalanine. (A) Immobilization of GFP at a fixed concentration of 0.246 mg/mL of each mutant protein. Fluorescence signal at 528 nm on a plate reader indicates that the n = 2 GFP-151 azide tethered variant performed best for immobilization (immobilizations conducted in triplicate, average standard deviation ±296). (B) Image of the immobilized resins as irradiated at 365 nm on a transilluminator. Overall, the tethered variants appeared to have a greater immobilization efficiency than the traditional pAzF.

Interestingly, the tether distance did factor into the efficiency of the immobilization reaction but not as expected, as increased distance from the protein did not necessarily lead to improved immobilization as measured by relative fluorescence intensity on a microplate reader (Figure 2). However, confirming the results observed when coupling to the Alexafluor488 dye, the p3AzY (n = 2) derivative again had the highest coupling efficiency, indicating that the interplay of structural rigidity and distance is important when optimizing the bioconjugation reaction. This 3-carbon tether seems to exploit the best immobilization efficiency of each factor, conferring a high coupling efficiency.

Based on these results, we next investigated the role of the methylene tether on alternate reactions. Recently, we reported...
the first Glaser-Hay bioconjugation, linking two terminal alkynyl functionalities to yield a conjugated diyne.\(^6\) Consequently, this reaction was selected to explore the reactivity of the alkynyl-GFP series. The alkynyl-containing GFPs were first conjugated to the same AlexaFluor 488 alkynyl dye previously employed in the click reactions. All mutant proteins were subjected to conjugation conditions in the presence of CuI (50 mM) and TMEDA (50 mM) and a control in the presence of the alkynyl fluorophore but in the absence of catalyst system. After running for 4 h at 4 °C, the proteins were denatured at 98 °C for 15 min and then analyzed via SDS-PAGE (Figure 3). These

![Figure 3](image)

**Figure 3.** Glaser-Hay reactions of alkynyl-containing GFP mutants. (A) Bioconjugation of mutant GFPs with AlexaFluor488 under Glaser-Hay conditions. The Coomassie stained gel (top) indicates the incorporation of all tethered mutants into GFP-151. Moreover, the successful Glaser-Hay coupling with the fluorophore was confirmed via fluorescent imaging of the same gel (bottom), as AlexaFluor fluorescence is only observed in reactions containing the fluorophore and the CuI/TMEDA system (immobilizations conducted in triplicate, average standard deviation ±774). (B) Immobilization of alkynyl-GFP mutants on a sepharose resin under Glaser-Hay conditions. Measuring resin fluorescence on a plate-reader (Ex 395/Em 512), all GFPs were able to be immobilized. (C) Immobilization was also confirmed by GFP fluorescent imaging on a transilluminator.

couplings all were successful, as demonstrated by a fluorescent band at ~26 kDa; however, the different tethered derivatives reacted with different efficiencies. 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.

The alkynyl series was also reacted with the alkynyl-derivatized sepharose resin to assess coupling efficiencies. All of the alkynyl-UAAs containing GFP mutants were able to be immobilized, as indicated by the presence of fluorescence on the resin after reaction with CuI/TMEDA (Figure 3B). However, it appears that there is no relative benefit to increasing the tether length, in contrast to the trends observed with the azide tethers. Interestingly, the p-propargylphenylalanine, which has an alkynyl moiety closest to the phenyl ring of phenylalanine, performed the best. Based on the proposed mechanism for the Glaser-Hay reaction, we hypothesize that the enhanced rigidity of the pPrF helps properly orient the protein–resin interactions as the copper binds in a \(\eta^1\) fashion. Moreover, the enhanced flexibility afforded by the novel alkynyl-UAAs may diminish their propensity to react, as the steric bulk of the protein may be more predisposed to interact with the resin. An odd/even effect is observed as the p3yneY-containing GFP performed the best among the UAA-mutated GFPs in the immobilization experiment (Figure 3C). This suggests that there may be some benefit, although slight, to increasing the distance of the protein from the site of reactivity. This increase drops once the alkyl chain gets too flexible, as evidenced by the drop in the p4ynY-containing GFP. Interestingly, within the series, the conjugation partner also plays a role in efficiency as the Glaser-Hay reaction with the small fluorophore was most efficient with the longer tether (Figure 3A), whereas reaction with the bulky resin required decreased flexibility.

Finally, the bromo-GFP series was investigated utilizing a \(S_N^2\) reaction. The incorporation of this series with a different synthetase, and their reaction with thiols, has previously been reported in a mammalian system;\(^{32,33}\) however, we were interested to see if these synthetic intermediates could also be incorporated with the pCNF aaRS. Due to the presence of many nucleophilic functional groups within biological settings, this reaction is not truly bioorthogonal, as cross-reactivity is likely. Nevertheless, we reasoned that successful nucleophilic substitution with an amine resin, thereby immobilizing the GFP onto the resin, was an ideal means to demonstrate the successful incorporation of the bromo-UAAs and the generation of the p2BrY-, p3BrY-, and p4BrY-GFPs. An amino-functionalized sepharose resin was generated using ethylenediamine. This resin was then subjected to the bromo-GFP series in the presence of a carbonate buffer (pH = 9.5). As a control, wild-type GFP was subjected to the same conditions. Gratifyingly, all of the bromo-containing GFPs were successfully immobilized, with p4BrY performing the best (Figure 4). Compared to the other functionalities and reactions, in which the three carbon alkyl chain variant of the
UAA consistently performed the best, this was the only case where we found a four-carbon alkyl chain to increase the propensity to immobilize. This is to be expected, as this bromo-tethered variant is the least sterically hindered by the protein’s bulk and will therefore perform the best in a nucleophilic substitution reaction.

Furthermore, the wild-type GFP was not immobilized, as indicated by minimal fluorescence on a microplate reader, verifying that the presence of the bromine functionality was responsible for the immobilization, and thereby demonstrating that we had successfully synthesized and incorporated bromotyrosine UAA.

In conclusion, nine novel UAAs with various functional handles and alkyl chains were synthesized. Furthermore, all UAA were successfully incorporated into GFP using a single promiscuous aaRS. The novel functional handles present in the UAA allowed access to a range of 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. In particular, we have elaborated upon a previously reported immobilization technique using an azido-containing GFP. We improved the immobilization efficiency by increasing the distance between the protein and the site of the CuAAC reaction. In addition, we have also demonstrated the utility of the Glaser-Hay reaction for protein immobilization.

Based on these results, we have illustrated the potential of UAA-mutagenesis to generate unique, site-specific functional handles in proteins for the preparation of various bioconjugates. This technology has downstream molecular imaging applications for the study of protein dynamics in vivo, as well as applications involving immobilized enzymes and the synthesis of well-defined therapeutic agents.

**ASSOCIATED CONTENT**

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.bioconjchem.5b00424.

Experimental protocols, control bioconjugation results, fluorescence imaging (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

*E-mail: dyoung01@wm.edu.*

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors would like to thank the Jeffress Memorial Trust and the College of William & Mary for financial support. Also, we thank Dr. Peter G. Schultz for providing plasmids necessary for incorporation of unnatural amino acids, and Dr. JC Potsma and Kathy Hyunsh for assistance with mass spectrometry.

**REFERENCES**

nanobeads containing calmodulin site-specifically immobilized via an azido group. J. Biochem. 154, 159–65.


