Bioorganic & Medicinal Chemistry Letters xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Site-specific incorporation of a fluorescent terphenyl unnatural amino acid

Jessica S. Lampkowski, Diya M. Uthappa, Douglas D. Young *

Department of Chemistry, College of William & Mary, Williamsburg, VA 23187, USA

ARTICLE INFO

Article history:
Received 20 August 2015
Revised 16 September 2015
Accepted 21 September 2015
Available online xxxx

Keywords:
Unnatural amino acids
Green fluorescent protein
Fluorescent sensors
Protein expression
Aminoacyl-tRNA synthetase

ABSTRACT

The site-specific incorporation of unnatural amino acids into proteins has a wide range of biological implications. Of particular interest is the incorporation of fluorescent probes as a mechanism to track protein function, transport, and folding. Thus, the development of a novel system for the incorporation of new fluorescent unnatural amino acids has significant utility. Specifically, we have elucidated an aminoacyl-tRNA synthetase capable of recognizing a terphenyl UAA derivative, and charging a cognate tRNA with this amino acid for protein incorporation. Moreover, we have successfully incorporated this fluorescent UAA into GFP at several key residues, demonstrating a novel means to modulate fluorescence within the protein.

Modifying proteins with unnatural amino acids has allowed for expansion of the genetic code, facilitating the production of modified proteins with novel functionality.[1–3] The incorporation of unnatural amino acids (UAAs) into proteins provides a means to better study protein function and structure both in vivo and in vitro. To date, a wide range of chemical functionalities have been introduced into proteins via unnatural amino acids including, azides, alkynes, fluorophores, photosensitive moieties, and metal binders to name a few.[4–11] Specifically, modifying proteins with fluorescent probes provides the ability to examine structure and function of proteins as well as visualize their cellular location via fluorescence spectroscopy.[12] Numerous fluorescent probes have been utilized within a protein context to further advance our understanding of proteins. Different fluorophores possess unique spectral properties for different applications, and thus the generation of multiple fluorophore labels is advantageous, and allows for adaptation to the needs of the experiment.[13–15] Consequently, we aim to investigate the incorporation of a fluorescent terphenyl unnatural amino acid into *Escherichia coli* utilizing an orthogonal tRNA/aminoacyl-tRNA synthetase methodology.

The introduction of fluorescent probes into proteins can be accomplished by a variety of mechanisms. Perhaps the simplest means involves the chemical modification of a protein with a synthetic fluorophore post-translation. However, this method may be limited due to a low availability of reactive surface residues as well as non-specific labeling with limited control over the location of fluorophore modification or the number of residues modified with synthetic fluorophore.[16–20] Another methodology includes the use of chemically mis-acylated suppressor tRNAs to incorporate unnatural amino acids. This methodology, however, affords limited yields of protein and is limited to easily assessable positions on the protein.[21–23]

In order to alleviate these limitations, an approach utilizing an orthogonal tRNA and aminoacyl-tRNA synthetase, (aaRS) has previously been developed to incorporate a desired UAA in response to a TAG codon.[1] This methodology exploits the bacterial translational machinery to afford the site-specific incorporation of UAAs. Due to the high selectively of this approach, we hypothesized using an aaRS/tRNA system to successfully incorporate a fluorescent unnatural amino acid would prove to be advantageous to other methodologies.[2] Typically this approach requires a double-sieve selection of an aaRS mutant library to identify a mutant aaRS capable of both recognizing the desired UAA and charging the corresponding orthogonal tRNA.[3,24] This approach has already been employed to genetically incorporate several fluorescent UAAs, including L-(7-hydroxycoumarin-4-yl) ethyl glycine (CouA), 2-amino-3(5-dimethylamino)naphthalene-1-(sulfonamide)propanoic acid (DansA), and 6-propionyl-2-(*N*,*N*-dimethyl)-aminonaphthalene (Anap). These fluorescent UAAs have been employed in the cellular imaging of several proteins to identify various properties, such as cellular location and protein unfolding. However, expanding the fluorescent genetic code with new UAAs harboring different spectral properties may possess additional advantages. Specifically, we aimed to identify an aaRS capable of incorporating a terphenyl UAA into green

* Corresponding author.

http://dx.doi.org/10.1016/j.bmcl.2015.09.050
0960-894X/

fluorescent protein (GFP) in order to further enhance fluorescence activity and probe protein structure.

The terphenyl fluorophore represents an interesting π-conjugated molecule for use in biological applications. Characterization studies of terphenyl, as well as other conjugated systems, were performed to demonstrate their unique photophysical properties. Terphenyl moieties have high lifetimes ($\Phi_t = 0.49$; $\tau = 4.38$ ns) and novel emission spectra ($\lambda_{ex}\lambda_{em} = 280/342$ nm).[25] Moreover, these moieties have been found to be sensitive to environmental conditions including solvent.[26] Consequently, the terphenyl fluorophore has been utilized in a variety of applications including two-photon laser scanning microscopy,[27,28] femtosecond fluorescence spectroscopy,[29] and α-helical secondary structure investigations.[30] These studies have facilitated the expanded knowledge of charge transfer, isomerization and protein dynamics within biological systems.

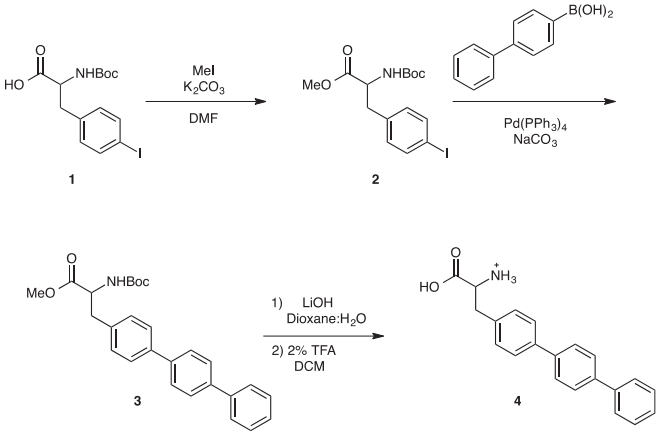
Incorporating a terphenyl UAA as a fluorescent probe into protein is beneficial in that the terphenyl is of relatively small size and allows for conformational mobility within a protein structure. Recently, various terphenyl derivatives have been incorporated into dihydrofolate reductase (DHFR) from *Escherichia coli* for FRET fluorescence studies.[31,32] These studies utilized a chemically acylated tRNA methodology, requiring extra synthetic maniupulations and lowered protein yields. Excitingly, the incorporation of a terphenyl UAA provided a means to make measurements of changes in protein conformations, and further knowledge of protein dynamics. When chemically incorporated, these derivatives did not result in any loss of functional catalytic activity of DHFR.[31] Moreover, they also afforded rotational flexibility within protein folding patterns due to the biphenyl bonding associated with their structures. These studies truly demonstrate the utility of this novel amino acid, and consequently we aim to facilitate the increased utilization of this UAA via its complete genetic incorporation.

In order to assess the feasibility of the approach, the fully deprotected 4-biphenyl-L-phenylalanine (**4**) was synthesized from conditions adapted from the literature (Scheme 1).[32] The synthesis was initiated with the esterification of N-Boc-4-iodo-L-phenylalanine (**1**) to generate the protected methyl ester **2** in good yield (90%). A Suzuki coupling was then employed with 4-biphenylboronic acid to yield the protected terphenyl unnatural amino acid, **3**. Finally, the protection groups were removed in a 2-step process to afford the fully deprotected 4-biphenyl-L-phenylalanine (**4**) in quantitative yield.

Following the synthesis of **4**, site-specific incorporation into a protein was attempted. Unlike previous approaches involving the chemical aminoacylation of tRNA, we attempted to identify an aminoacyl t-RNA synthetase (aaRS) capable of both recognizing **4**, and charging it onto the appropriate tRNA. Based on previous findings that some already evolved aaRSs demonstrate a degree of polyspecificity, an aaRS screen was undertaken with known polyspecific synthetases to alleviate the potential necessity of an aaRS selection.[33] Several synthetases were selected due to either known polyspecificity or due to their incorporation of structurally similar UAAs. Plasmids encoding the aaRS and tRNA were co-transformed into BL21(DE3) *E. coli* with a pET-GFP-TAG-151 plasmid, harboring the TAG codon at position 151. Following protein expression, GFP fluorescence was measured using a BioTek multiplate reader and expression cultures grown in the presence and absence of **4** were compared (Fig. 1). Gratifyingly, a differential was observed for cultures containing either the well documented polyspecific *p*CNF aaRS, or the NapA aaRS (which encodes the fluorescent naphthylalanine UAA).[11,34] To confirm the actual incorporation of **4**, expressions were repeated and the GFP mutant was purified using a Ni–NTA resin and analyzed by gel electrophoresis and mass spectrometry (Fig. 2). Expressions employing the *p*CNF aaRS occasionally yielded an unknown higher molecular weight band in conjunction with the desired GFP. Consequently, the NapA aaRS was used for further protein expressions to alleviate the need for additional protein purification, despite a slightly decreased protein yield.

The fluorescent UAA **4** was next incorporated into GFP at multiple residues to probe for environmental alterations in fluorescence and overall effect on the protein. Residues, 3, 66, 133, and 151 were selected in the study due to their varied structural components (Fig. 3).

Residue 3 is at the N-terminus of the protein in a non-structured loop, residue 66 comprises a key member of the normal GFP fluorophore, residue 133 is on a more structured loop opposite of residue 3, and residue 151 is at the terminus of a β-sheet that comprises the β-barrel of the protein and is most rigid. When examining the fluorescent mutants containing **4** by excitation at 395 nm, the UAA has a dramatic effect on the overall spectra based on its position (Fig. 4). Relative to the wild-type protein containing no UAA, an approximate 3 nm red-shift is observed for all spectra. Moreover, the ~512 nm emission corresponding to the deprotonated tyrosine residue 66 in the core fluorophore is impacted by the presence of **4**, especially in relation to the protonated ~455 nm emission. When located at the rigid 151 position, little change in ratio of the two states is detected relative to the wild-type protein. However, at the more flexible positions, the presence of **4** tends to favor the protonated state relative to the deprotonated fluorophore signifying a structural change in the overall protein that alters the pK_a of tyrosine 66. Moreover, the placement of **4** at that key residue eliminates the 512 nm emission as an acidic proton is no longer present within the fluorophore.



Scheme 1. Synthesis of 4-biphenyl-L-phenylalanine.

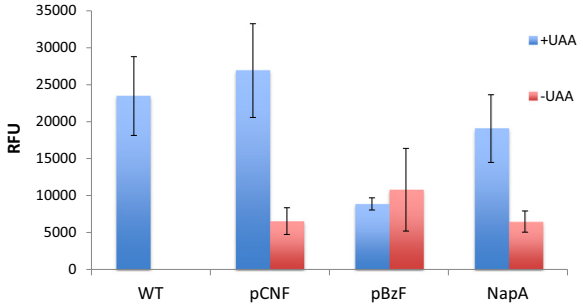


Figure 1. Initial aaRS screen to identify previously evolved aminoacyl-tRNA synthetases capable of recognizing **4**. Culture containing different synthetases and a GFP-151TAG were grown in the presence and absence of **4**. Fluorescence indicates functional GFP protein, and thus incorporation of **4**.