Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 1521–1524
Bioorganic & Medicinal Chemistry Letters

In vivo incorporation of an alkyne into proteins in *Escherichia coli*

Alexander Deiters† and Peter G. Schultz*

Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

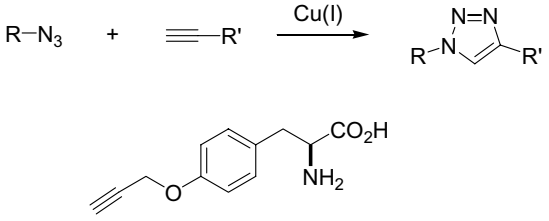
Received 29 July 2004; revised 18 November 2004; accepted 21 December 2004
Available online 23 January 2005

Abstract—Using a genetic selection we identified mutants of the *M. janaschii* tyrosyl-tRNA synthetase that selectively charge an amber suppressor tRNA with *para*-propargyloxyphenylalanine in *Escherichia coli*. These evolved tRNA-synthetase pairs were used to site-specifically incorporate an alkynyl group into a protein, which was subsequently conjugated with fluorescent dyes by a [3+2]-cycloaddition reaction under mild reaction conditions.

1. Introduction

The ability to site-specifically, chemically modify proteins with nonpeptidic molecules such as spectroscopic probes, catalytic auxiliaries, or polymers provides a powerful means to both investigate and manipulate the chemical and biological properties of proteins. A common approach involves the bioconjugation of nucleophilic surface residues on the protein, for example, the side chains of lysine, histidine, or cysteine, with electrophilic groups on an exogenous molecule, such as aldehydes, α-halo carboxamides, and *N*-hydroxy succinimides.[1] A general challenge is the modest selectivity of these reactions and the presence of multiple reactive amino acids in proteins, leading to the formation of heterogeneous mixtures of labeled proteins. One solution to this problem is the biosynthetic incorporation of unnatural amino acids with novel reactivity into proteins.[2,3] By evolving the specificity of orthogonal tRNA-synthetase pairs, we have selectively and efficiently incorporated over thirty unnatural amino acids into proteins in response to nonsense and frameshift codons in both prokaryotes and eukaryotes, including glycosylated amino acids, redox active amino acids, photo-crosslinking amino acids, and amino acids containing keto, azido, alkynyl, and iodo groups.[4] The chemistry of the alkynyl and the azido group is completely orthogonal to the chemistry of all endogenous functional groups present in proteins; an example of their unique reactivity is the irreversible formation of triazoles by a [3+2] cycloaddition (Scheme 1).[5] When this reaction is conducted in presence of copper(I) at room temperature in aqueous media (conditions compatible with most proteins) it proceeds in a completely regioselective fashion[6] and can be used to selectively modify proteins into which alkynyl and azido functional groups have been introduced.[7]

Previously, we showed that the alkynyl amino acid **1** could be selectively incorporated into proteins expressed in yeast using an orthogonal *E. coli* tRNA/tRNA-synthetase pair.[7a] In order to extend this chemistry to proteins expressed in bacteria, we now report the evolution of a *M. jannaschii* tyrosyl tRNA/tRNA-synthetase pair ($MjTyrRS/tRNA_{CUA}$) in *E. coli* that selectively accepts **1** (synthesized from commercially available *N*-Boc-tyrosine in three steps with an overall yield of 81%)[7] as a substrate, but does not utilize any of the common 20



1

Scheme 1. Triazole formation by [3+2] cycloaddition of an alkyne and an azide. Amino acid **1** containing an alkynyl group.

* Corresponding author. Tel.: +1 858 7849300; fax: +1 858 7849440; e-mail: schultz@scripps.edu
† Present address: Department of Chemistry, North Carolina State University, PO Box 8204, Raleigh, NC 27695-8204, USA.

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.12.065

amino acids. The orthogonality[8] of this pair ensures that neither the tRNA nor the synthetase cross reacts with endogenous *E. coli* tRNAs or synthetases so that the unnatural amino acid gets delivered only in response to an amber nonsense codon, TAG. A library of $\sim 10^7$ different *M. jannaschii* tyrosyl tRNA-synthetases was generated by randomizing five active site residues (Tyr34, Glu107, Asp158, Ile159, and Leu162), based on a crystal structure of the homologous tyrosyl tRNA-synthetase from *Bacillus stearothermophilus*. This pool of synthetases was then passed through positive and negative rounds of selection. The positive selection is based on suppression of an amber stop codon at a permissive site (Asp112) in the chloramphenicol acetyltransferase (CAT) gene. When cells harboring the MjTyrRS library and the mutated CAT gene were grown on minimal media in the presence of **1** (1 mM) and chloramphenicol (80 $\mu g mL^{-1}$) they only survive if a mutant synthetase aminoacylates the $tRNA_{CUA}$ with either an endogenous amino acid or **1**. The synthetase genes were then transformed into cells containing a mutated gene encoding the toxic protein barnase, which has three amber mutations at permissive sites (Gln2, Asp44, Gly65). Growth of these cells in the absence of **1** selected against synthetases capable of accepting endogenous amino acids as a substrate. After just three rounds of selection, 96 clones were screened for a growth rate dependence on the presence or absence of **1**, and eight candidate clones were identified and sequenced (Table 1). Although no convergence was obtained, a preponderance of the following amino acids was found in the binding pocket of most of the clones: Ala32, Pro107/Gln107, Ala158, Ile159, and Ala162/Pro162. The mutations Tyr32→Ala32 and Asp158→Ala158 may result in the loss of hydrogen bonds between Tyr32, Asp158, and the natural substrate tyrosine, thus disfavoring its binding. The occurrence of small and mostly hydrophobic side chains might be expected to facilitate binding of **1**.

Synthetase pPR-MjRS-1 was further characterized. This synthetase confers chloramphenicol resistance on *E. coli* with IC_{50} values of 110 and 5 $\mu g mL^{-1}$ in the presence and absence of **1**, respectively. The large difference between the chloramphenicol resistance with and without **1** suggests a substantial in vivo specificity of pPR-MjYRS-1 for **1**. The $tRNA_{CUA}$/MjYRS pair was then used to selectively incorporate **1** into sperm whale myoglobin, a monomeric 153-residue heme protein that has been the focus of a number of structural, mechanistic, and protein folding studies.[9] To produce alkyne modified myoglobin, the fourth codon (Ser4) was mutated to TAG and a C-terminal 6 × His tag was added. In the presence of the mutant MjTyrRS, $tRNA_{CUA}$, and **1** (1 mM in liquid minimal media), full-length myoglobin was produced with a yield of 2 mg/L after purification by Ni-affinity chromatography (>90% homogeneous by SDS-PAGE and Gelcode Blue staining). In the absence of **1** no myoglobin was visible after staining or Western blot (using an anti-His6 antibody), indicating a high selectivity of the evolved synthetase (Fig. 1). To further confirm the identity of the amino acid incorporated, a tryptic digest of a mutant myoglobin with **1** at position 74 was subjected to liquid chromatography/tandem mass spectrometry. The precursor ions corresponding to the singly and doubly charged ions of the peptide HGVTVLTALGY*ILK containing the unnatural amino acid (denoted Y*) were separated and fragmented with an ion trap mass spectrometer. The fragment ion masses could be unambiguously assigned, confirming the site-specific incorporation of **1**.[10] The LC MS/MS runs did not suggest incorporation of any natural amino acid at this position, confirming the high selectivity of the evolved synthetase.

The Ser4→**1** mutant myoglobin was then derivatized with the azido functionalized dyes **2** and **3**, containing the dansyl and fluoresceine fluorophore, respectively (Scheme 2).[7] A [3+2] cycloaddition reaction was carried out with purified protein at a concentration of 0.5 mg/mL in 0.1 M phosphate buffer (pH 8) containing $CuSO_4$ (1 mM) and Cu-wire (1 mg) as a reducing agent (Scheme 2). The use of tris(carboxyethyl)phosphine (2 mM) as an alternative reducing agent generally led to a similar labeling efficiency. In contrast to previous observations,[7b] the presence or absence of the ligand tris(1-benzyl-1*H*-[1,2,3]triazol-4-ylmethyl)amine did not have a substantial influence on the outcome of these reactions. The fluorescently labeled proteins[11] were then analyzed by SDS-PAGE and imaged (Scheme 2).[12] The labeling

Table 1. Active site sequences of pPR-MjRSs selected for incorporation of **1**

Synthetase	32	107	158	159	162
Wild type	Tyr	Glu	Asp	Ile	Leu
pPR-MjRS-1[a]	Ala	Pro	Ala	Ile	Ala
pPR-MjRS-2	Ala	Lys	Ala	Ile	Ala
pPR-MjRS-3	Ala	Arg	Ala	Ile	Pro
pPR-MjRS-4	His	Ala	Ala	Ile	Pro
pPR-MjRS-5	Ser	Gln	Ala	Ile	Ala
pPR-MjRS-6	Thr	Ser	Leu	His	Pro
pPR-MjRS-7	Ala	Gln	Pro	Gly	Thr
pPR-MjRS-8	Ala	Pro	Ser	Leu	His

[a] Additional Leu110→Phe110 mutation.

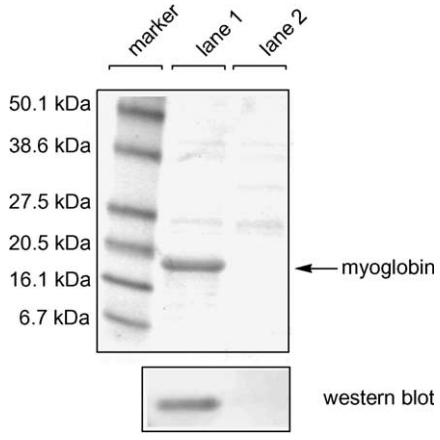
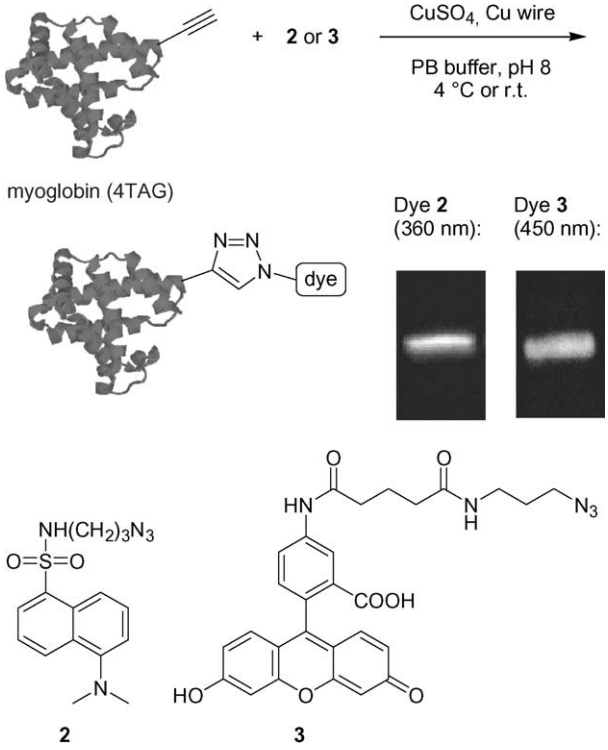


Figure 1. Gelcode Blue stained SDS-PAGE gel of purified Ser4→**1** myoglobin. Lane 1 contains protein expressed in minimal media in the presence of **1**; lane 2 contains a sample expressed in the absence of **1**. An anti-His6 antibody was used to detect the hexahistidine tag at the C-terminus of myoglobin.



Scheme 2. [3+2] Cycloaddition of mutant myoglobin and the dyes **2** and **3**. The labeling was visualized by gel imaging using UV irradiation.

efficiency was ~75% as determined by comparison of the A_{280}/A_{495} values for myoglobin labeled with **3**.[6] The selectivity of this bioconjugation was verified by the fact that no reaction between wild type myoglobin and **2** or **3** was observed.

In summary, we have demonstrated that an alkyne amino acid can be genetically encoded in *E. coli* and used to selectively modify a protein by a [3+2] cycloaddition reaction with an exogeneous small molecule azide. We are currently applying this technology to the production of PEGylated proteins and are investigating the structural basis for the high selectivity of this and other evolved synthetases by structural studies.

2. Experimental

To express protein, plasmid pBAD/JYAMB-4TAG (which encodes the mutant sperm whale myoglobin gene with an arabinose promoter and a rrnB terminator; the tyrosyl $tRNA_{CUA}$ on a lpp promoter and a rrnC terminator; and a tetracycline resistance marker) was co-transformed with a pBK vector (encoding the mutant synthetase and a kanamycin resistance gene) into DH10B *E. coli*. Cells were amplified in Luria–Bertani media (5 mL) supplemented with tetracycline (25 mg/L) and kanamycin (30 mg/L), washed with phosphate buffer, and used to innoculate 500 mL of liquid glycerol minimal media (supplemented with 0.3 mM leucine) containing the appropriate antibiotics, **2** (1 mM), and arabinose (0.002%). Cells were grown to saturation and then harvested by centrifugation. The protein was purified using Ni-affinity chromatography, yielding ~1 mg of mutant myoglobin.

For the cycloaddition reaction, 1 μL of $CuSO_4$ (50 mM in H_2O), 2 μL of dye **2** or **3** (50 mM in EtOH), 2 μL of tris(1-benzyl-1*H*-[1,2,3]triazol-4-ylmethyl)amine (50 mM in DMSO), and 1 mg Cu wire or 1 μL tris(carboxyethyl)phosphine (100 mM in H_2O) were added to 45 μL of mutant myoglobin (~0.5 mg/mL) in phosphate buffer (pH = 8). After 8 h at room temperature or overnight at 4 °C, 450 μL H_2O were added and the mixture was centrifuged through a dialysis membrane (10 kDa cut off). After washing the supernatant with 2×500 μL phosphate buffer by centrifugation, the solution was brought to a volume of 50 μL. A sample of 20 μL was analyzed by SDS-PAGE and imaged.

Acknowledgements

A.D. gratefully acknowledges a postdoctoral fellowship by the Deutsche Forschungsgemeinschaft. We thank Mridul Mukherji for the mass spectrometry analysis and the Sharpless group for providing us with tris(1-benzyl-1*H*-[1,2,3]triazol-4-ylmethyl)amine. This work is supported by funding from the National Institutes of Health (GM62159) and is manuscript 16697-CH of The Scripps Research Institute.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2004.12.065.

References and notes

1. Lemineux, G. A.; Bertozzi, C. R. *TIBTECH* **1996**, *16*, 506.
2. Wang, L.; Schultz, P. G. *Chem. Commun.* **2002**, *1*, 1.
3. van Maarseveen, J. H.; Back, J. W. *Angew. Chem.* **2003**, *115*, 6106.
4. (a) Anderson, J. C.; Wu, N.; Santoro, S. W.; Lakshman, V.; King, D. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7566; (b) Alfonta, L.; Zhang, Z.; Uryu, S.; Loo, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 14662; (c) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56; (d) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964; (e) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020; (f) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.
5. (a) Padwa, A. In *Comprehensive Organic Synthesis*; Trost, B. M., Ed.; Pergamon: Oxford, 1991; Vol. 4, p 1069; (b) Huisgen, R. In *1,3-Dipolar Cycloaddition Chemistry*; Padwa, A., Ed.; Wiley: New York, 1984; p 1.
6. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.
7. (a) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**,

125, 11782; (b) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192; (c) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164.

8. Steer, B. A.; Schimmel, P. *J. Biol. Chem.* **1999**, *274*, 35601.
9. (a) Reedy, C. J.; Gibney, B. R. *Chem. Rev.* **2004**, *104*, 617, and references cited therein; (b) Uzawa, T.; Akiyama, S.; Kimura, T.; Takahashi, S.; Ishimori, K.; Morishima, I.; Fujisawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1171, and references cited therein; (c) Wright, P. E.; Baldwin, R. L. In *Frontiers in Molecular Biology: Mechanisms of Protein Folding*; Pain, R., Ed.; Oxford University Press: London, 2000; p 309.
10. The myoglobin-74TAG mutant was used due to improved properties for LC MS/MS analysis (see Supporting information).
11. (a) Blake, R. A. *Curr. Opin. Pharmacol.* **2001**, *1*, 533; (b) Wouters, F. S.; Verveer, P. J.; Bastiaens, P. I. H. *Trends Cell Biol.* **2001**, *11*, 203; (c) Zacharias, D. A.; Baird, G. S.; Tsien, R. Y. *Curr. Opin. Neurobiol.* **2000**, *10*, 416.
12. Protein modified with the dansyl dye **2** (λ_{ex} = 337 nm, λ_{em} = 506 nm) was in-gel imaged at 360 ± 30 nm using an Eagle Eye densitometer (Stratagene). Attachment of the fluoresceine dye **3** (λ_{ex} = 495 nm, λ_{em} = 516 nm) was visualized at 450 ± 30 nm with a Storm Phosphorimager (Molecular Dynamics).