J|A|C|S
COMMUNICATIONS

Published on Web 09/06/2003

Adding Amino Acids with Novel Reactivity to the Genetic Code of *Saccharomyces Cerevisiae*

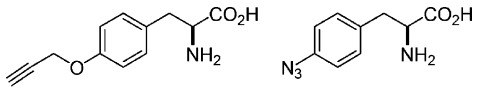
Alexander Deiters, T. Ashton Cropp, Mridul Mukherji, Jason W. Chin, J. Christopher Anderson, and Peter G. Schultz*

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

Received July 1, 2003; E-mail: schultz@scripps.edu

There is a considerable need for chemical reactions that modify proteins under physiological conditions in a highly selective fashion.[1] Most reactions currently used for the selective modification of proteins involve covalent bond formation between nucleophilic and electrophilic reaction partners: e.g., the reaction of α-halo ketones with histidine or cysteine side chains. Selectivity in these cases is determined by the number and accessibility of the nucleophilic residues in the protein. In the case of synthetic or semisynthetic proteins, other more selective reactions can be used, such as the reaction of an unnatural keto amino acid with hydrazides or aminooxy compounds.[2] Recently, we demonstrated that ketone-containing amino acids can be genetically encoded in bacteria and yeast using orthogonal tRNA−synthetase pairs with altered amino acid specificities.[3,4] This methodology has made possible the selective labeling of virtually any protein with a host of reagents, including fluorophores, cross-linking agents, and cytotoxic molecules.

We now report a second highly efficient method for the selective modification of proteins which involves the genetic incorporation of azide- or acetylene-containing unnatural amino acids into proteins in response to the amber nonsense codon, TAG. These amino acid side chains can then be modified by a Huisgen [3 + 2] cycloaddition reaction[5] with acetylene or azide derivatives, respectively. Because this method involves a cycloaddition rather than a nucleophilic substitution, proteins can be modified with extremely high selectivity.[6] This reaction can be carried out at room temperature under aqueous conditions with excellent regioselectivity (1,4 > 1,5) by the addition of catalytic amounts of Cu(I) salts to the reaction mixture.[7] Indeed, Finn and co-workers have shown that this azide−alkyne [3 + 2] cycloaddition can be conducted on the surface of an intact cowpea mosaic virus.[8,9] To selectively introduce either the acetylene or azide functional group into eukaryotic proteins at unique sites, we have evolved orthogonal TyrRS/tRNA$_{CUA}$ pairs in yeast that genetically encode the acetylene and azido amino acids **1** and **2**, respectively.



1 **2**

Previously we demonstrated that an *E. coli* tyrosyl tRNA−tRNA synthetase pair is orthogonal in yeast; i.e., neither the tRNA nor the synthetase cross-react with the endogenous yeast tRNA or synthetases.[10] This orthogonal tRNA−synthetase pair has been used to selectively and efficiently incorporate a number of unnatural amino acids in yeast in response to the TAG codon.[4c] To alter the amino acid specificity of the *E. coli* tyrosyl−tRNA synthetase to accept amino acid **1** or **2**, a libary of ∼10^7 mutants was generated by randomizing the codons for Tyr[37], Asn[126], Asp[182], Phe[183], and Leu[186]. These five residues were chosen on the basis of a crystal structure of the homologous synthetase from *B. stearothermophilus*. To obtain a synthetase for which the particular amino acid serves as a substrate, a selection scheme was used in which the codons for Thr[44] and Arg[110] of the gene for the transcriptional activator GAL4 were converted to amber nonsense codons (TAG).[10] Suppression of these amber codons in the MaV203:pGADGAL4-(2TAG) yeast strain leads to production of full-length GAL4,[11] which in turn drives expression of the *HIS3* and *URA3* reporter genes. The latter gene products complement histidine and uracil auxotrophy, allowing clones harboring active synthetase mutants to be selected in the presence of **1** or **2**. Synthetases that load endogenous amino acids are removed by growth on medium lacking **1** or **2** but containing 5-fluoroorotic acid, which is converted into a toxic product by URA3. By passing the library through three rounds of selection (positive, negative, positive), we identified synthetases selective for **1** (pPR-EcRS1-5) and for **2** (pAZ-EcRS1-6), as shown in Table 1 (see the Supporting Information).

All synthetases show a strong sequence similarity, including a conserved Asn[126], suggesting an important functional role for this residue. Surprisingly, the synthetases pPR-EcRS-2 and pAZ-EcRS-6, evolved to bind **1** and **2**, respectively, converged to the same sequence (Tyr[37] → Thr[37], Asn[126] → Asn[126], Asp[182] → Ser[182], and Phe[183] → Ala[183], Leu[186] → Leu[186]). Both hydrogen bonds between the phenolic hydroxy group of bound tyrosine and Tyr[37] and Asp[182] are disrupted by mutations to Thr and Ser, respectively. Phe[183] is converted to Ala, possibly providing more space to accommodate the unnatural amino acid. To confirm the ability of this synthetase (and the other synthetases) to accept either amino acid as a substrate, selection strains harboring the synthetase plasmids were grown on media lacking uracil (the same results were obtained for media lacking histidine) but supplemented with either **1** or **2**. Growth results revealed that four of the five alkyne synthetases were able to load both unnatural amino acids onto its tRNA. The azido synthetases seem to be more selective, since only pAZ-EcRS-6 (which is identical with pPR-EcRS-2) was able to amino acylate its tRNA with both **1** and **2**. The fact that no growth was detected in the absence of **1** or **2** suggests that the synthetases do not accept any of the 20 common amino acids as a substrate (see the Supporting Information).

For all further experiments pPR-EcRS-2 (pAZ-EcRS-6) was used, allowing one to control which unnatural amino acid is incorporated simply by adding either **1** or **2** to media containing the expression strain. For protein production the codon for the permissive residue Trp[33] of human superoxide dismutase-1 (SOD) fused to a C-terminal 6×His tag was mutated to TAG. Protein was expressed in the presence or absence of 1 mM **1** or **2** and purified by Ni-NTA chromatography. Analysis by SDS-PAGE and Western blot revealed unnatural amino acid dependent protein expression with a fidelity

 10.1021/ja0370037 CCC: $25.00

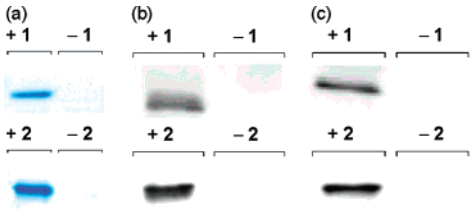


Figure 1. SOD expression in the presence or absence of **1** and **2**: (a) Gelcode Blue stain; (b) Western blot with anti-SOD antibody; (c) Western blot with anti-6×His antibody.

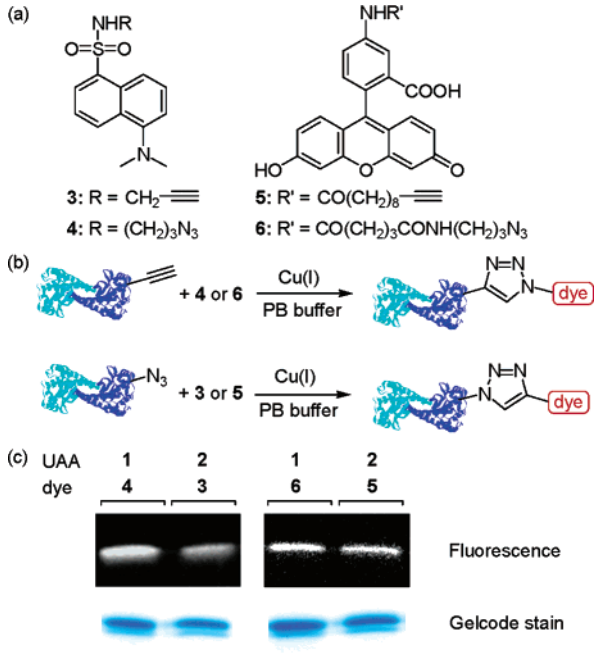


Figure 2. Protein labeling by [3 + 2] cycloaddition: (a) synthesized dye labels 3−6; (b) reaction between SOD and dye; (c) in-gel fluorescence scanning and Gelcode Blue staining.

of >99%, as judged by densitometry comparisons to protein expression in the absence of **1** or **2** (Figure 1). To further confirm the identity of the amino acid incorporated, a tryptic digest was subjected to liquid chromatography and tandem mass spectrometry. The precursor ions corresponding to the singly and doubly charged precursor ions of the peptide VY*GSIK 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 each unnatural amino acid. LC MS/MS did not indicate incorporation of any natural amino acid at this position. The signal-to-noise ratios of the peptide for all mutants were >1000, suggesting fidelity of incorporation better than 99.8% (see the Supporting Information).

To demonstrate that small organic molecules can be conjugated to proteins by an azide−alkyne [3 + 2] cycloaddition reaction, the dyes **3**−**6**, which contain either an acetylenic or an azido group and bear a dansyl or fluoresceine fluorophore, were synthesized (see the Supporting Information). The cycloaddition itself was carried out with 0.01 mM protein in phosphate buffer, pH 8, in the presence of 2 mM **3**−**6**, 1 mM $CuSO_4$, and ∼1 mg Cu wire for 4 h at 37 °C (see Figure 2 and the Supporting Information). After dialysis the labeled proteins were analyzed by SDS-PAGE and in-gel imaged using a densitometer in the case of the dansyl dyes **3** and **4** (λ_{ex} = 337 nm, λ_{em} = 506 nm) or a phosphorimager in the case of the fluoresceine dyes **5** and **6** (λ_{ex} = 483 nm, λ_{em} = 516 nm).[12] The labeled proteins were characterized by LC MS/MS analysis of tryptic digests showing site-specific attachment of the fluorophores (data not shown), and the conversion was 75% on average.[13] The selectivity of this bioconjugation was verified by the fact that there was no observable reaction between **3** and alkyne protein or between **4** and azido protein (data not shown).

In conclusion, we have demonstrated a site-specific, fast, reliable, and irreversible method of bioconjugation to proteins based on a [3 + 2] cycloaddition. We are currently investigating the possibility of carrying out the cycloaddition in vivo as well as using this approach to selectively modify proteins with a variety of other moieties, including poly(ethylene glycol) and cross-linking agents.

Acknowledgment. A.D. is supported by the Deutsche Forschungsgemeinschaft. J.W.C. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (Grant No. DRG1707-02). J.C.A. is a NSF predoctoral fellow. This work was supported by the National Institutes of Health (Grant No. GM62159). We thank the Sharpless group for providing us with the tris-(triazolyl)amine ligand. This is manuscript #15910-CH of The Scripps Research Institute.

Supporting Information Available: A table giving sequences of all pPR-EcRS and pAZ-EcRS mutants obtained, text giving details of the syntheses of **1** and **3**−**6**, the protein expression experiment, and the cycloaddition experiment, and figures giving the mass spectra of alkyne- and azido-SOD. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Review: Lemineux, G. A.; Bertozzi, C. R. *TIBTECH* **1996**, *16*, 506−513.
(2) (a) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150−8151. (b) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125−1128.
(3) (a) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498−500. (b) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026−9027. (c) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020−11024.
(4) (a) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56−61. (b) Zhang, Z.; Smith, B. A.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G. *Biochemistry* **2003**, *42*, 6735−6746. (c) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science*, in press.
(5) Reviews: (a) Padwa, A. In *Comprehensive Organic Synthesis*; Trost, B. M., Ed.; Pergamon: Oxford, U.K., 1991; Vol. 4, pp 1069−1109. (b) Huisgen, R. In *1,3-Dipolar Cycloaddition Chemistry*, Padwa, A., Ed.; Wiley: New York, 1984; pp 1−176.
(6) Another method is the ligand exchange on a bis-arsenic compound with a tetracysteine motif: Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269−272.
(7) (a) Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057−3064. (b) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596−2599.
(8) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192−3193.
(9) For another recent example of the electrophilic introduction of an azido group into a protein and a subsequent [3 + 2] cycloaddition see: Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686−4687.
(10) Chin, J. W.; Cropp, T. A.; Chu, S.; Meggers, E.; Schultz, P. G. *Chem. Biol.* **2003**, *10*, 511−519.
(11) (a) Keegan, L.; Gill, G.; Ptashne, M. *Science* **1986**, *231*, 699−704. (b) Ptashne, M. *Nature* **1988**, *335*, 683−689.
(12) (a) Blake, R. A. *Curr. Opin. Pharmacol.* **2001**, *1*, 533−539. (b) Wouters, F. S.; Verveer, P. J.; Bastiaens, P. I. H. *Trends Cell Biol.* **2001**, *11*, 203-211. (c) Zacharias, D. A.; Baird, G. S.; Tsien, R. Y. *Curr. Opin. Neurobiol.* **2000**, *10*, 416−421.
(13) As determined by comparison of A_{280}/A_{495} values for SOD labeled with **5** or **6**.

JA0370037