J|A|C|S
COMMUNICATIONS

Published on Web 10/19/2004

A Genetically Encoded Photocaged Amino Acid

Ning Wu, Alexander Deiters, T. Ashton Cropp, David King, 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 19, 2004; E-mail: schultz@scripps.edu

Recently, over 30 unnatural amino acids have been genetically encoded in both prokaryotic and eukaryotic organisms in response to unique triplet and quadruplet codons.[1−3] These include glycosylated amino acids, amino acids with keto, azido, alkynyl, and iodo groups, and photoreactive and redox-active amino acids. To further increase the structural diversity of this "expanded" genetic code requires additional unique tRNA/aminoacyl tRNA-synthetase pairs. Here we describe the generation of a new orthogonal *Escherichia coli* $tRNA^{Leu}$/leucyl tRNA-synthetase (LRS) pair that has been used to selectively incorporate OMe-L-tyrosine, the C8 amino acid, α-aminocaprylic acid, and the photocaged amino acid, *o*-nitrobenzyl cysteine, into proteins in yeast in response to the amber nonsense codon, TAG. In addition, we show that the latter amino acid can be used to photoregulate the activity of the proapoptotic cysteine protease, caspase 3. The development of this and other orthogonal tRNA-synthetase pairs should further extend both the nature and number of amino acids that can be selectively introduced into proteins in bacteria and yeast.

On the basis of the crystal structure of the homologous *Thermus thermophilus* leucyl tRNA synthetase (ttLRS),[4] it was anticipated that the *E. coli* leucyl synthetase (ecLRS) active site could be evolved to accommodate a wide variety of amino acid side chains. The leucine binding site is a relatively large cavity comprised of side chains and no backbone elements, making possible a large number of mutant active-site configurations. In addition, previous experiments[5,6] suggest that an amber suppressor tRNA/synthetase pair derived from the *E. coli* leucyl tRNA and cognate synthetase will be orthogonal in yeast; i.e., this pair will not interact with any of the host tRNAs or synthetases. This condition ensures that only the unnatural amino acid will be incorporated into proteins at the site specified by the amber codon and at no other site.

The orthogonality of the *E. coli* leucyl suppressor $tRNA_{CUA}$ ($Leu5_{CUA}$) (Figure 1A) (which has U35 and A37 in the anticodon loop) and the corresponding ecLRS was examined in vivo in a selection strain of *Saccharomyces cerevisiae* [MaV203:pGAD-GAL4(2 TAG)].[2] This strain harbors the transcriptional activator GAL4 with amber codons at two permissive sites (Thr44 and Arg110). Efficient suppression at both sites leads to expression of three reporter genes: *lacZ*, *his3*, and *ura3*. In the presence of either $Leu5_{CUA}$ or ecLRS alone, no significant β-galactosidase activity was detected above the residual activity in lysates from transformed cells. However, coexpression of both $Leu5_{CUA}$ and ecLRS resulted in a more than a 10^4-fold increase in activity, similar to the activity of the previously reported *E. coli* $tRNA^{Tyr}_{CUA}$/YRS pair.[2] These results indicate that this leucyl pair is orthogonal and can efficiently deliver leucine in response to a TAG codon in yeast.

To evolve leucyl synthetases specific for unnatural amino acids, the residues Met40, Leu41, Tyr499, Tyr527, and His537 (the corresponding residues in ttLRS are Met40, Phe41, Tyr507, Tyr535, and His545) were randomized simultaneously by enzymatic inverse polymerase chain reaction mutagenesis to generate an active-site library containing ~10^7 mutants (Figure 1B).[7] These residues form a large hydrophobic pocket surrounding one γ-methyl group of the leucine side chain (Figure 2A) and are highly conserved in bacterial LRSs. Three unnatural amino acids with distinct electronic and steric properties were selected to probe the adaptability of the LRS active site: *O*-methyl tyrosine (OMeY), α-aminocaprylic acid (C8), and *o*-nitrobenzyl cysteine (nbC) (Figure 1C).

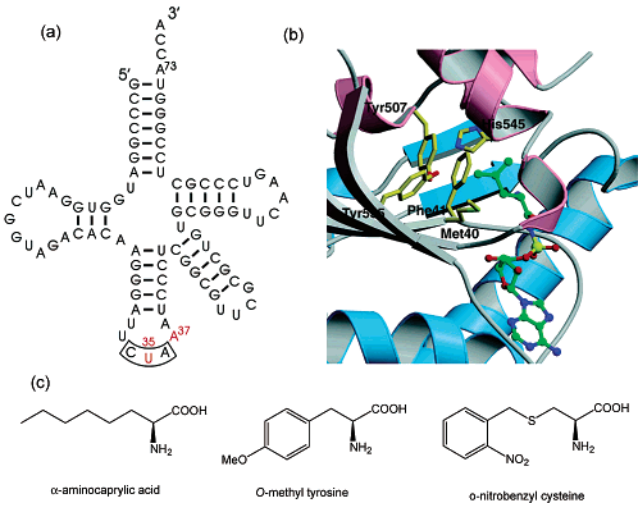

Figure 1. (A) Primary sequence of the leucyl suppressor tRNA, $Leu5_{CUA}$. (B) The ttLRS active site is shown with a bound leucyl sulfamoyl adenylate inhibitor (green, PDB entry 1H3N). The residues randomized in generating the synthetase library are in yellow. The catalytic domain of LRS is composed of two discontinuous stretches of primary sequences, with the N-terminal half in blue and the C-terminal half in pink. (C) Structrues of *O*-methyl tyrosine, α-aminocaprylic acid, and *o*-nitrobenzyl cysteine.

Mutant synthetases specific for each unnatural amino acid were selected on the basis of suppression of amber codons at positions 44 and 110 in the *gal4* gene in yeast. Specifically, charging of $Leu5_{CUA}$ by a mutant synthetase with the unnatural amino acid (added to the media at 1 mM concentration) or an endogenous amino acid results in growth either on media lacking histidine (−His) but containing 20 mM 3-aminotriazole (3-AT, a competitive inhibitor of the HIS3 protein), or on media lacking uracil (−Ura). In the absence of the unnatural amino acid, cells harboring synthetases that utilize endogenous amino acids show slowed growth on media containing 0.1% 5-fluoroorotic acid (5-FOA), which is converted into a toxic compound by URA3.[7] Three rounds of alternating positive and negative selections were carried out for each unnatural amino acid. After the final round, clones were chosen for their strict growth dependence on −Ura and −His/20 mM 3-AT plates supplemented with 1 mM unnatural amino acid. These clones were then tested for suppression of an amber codon at the permissive site 33 in human superoxide dismutase containing a C-terminal hexahistidine tag (hSOD-33TAG-His6) in the presence and in the absence of the unnatural amino acids in yeast. The best

10.1021/ja040175z CCC: $27.50

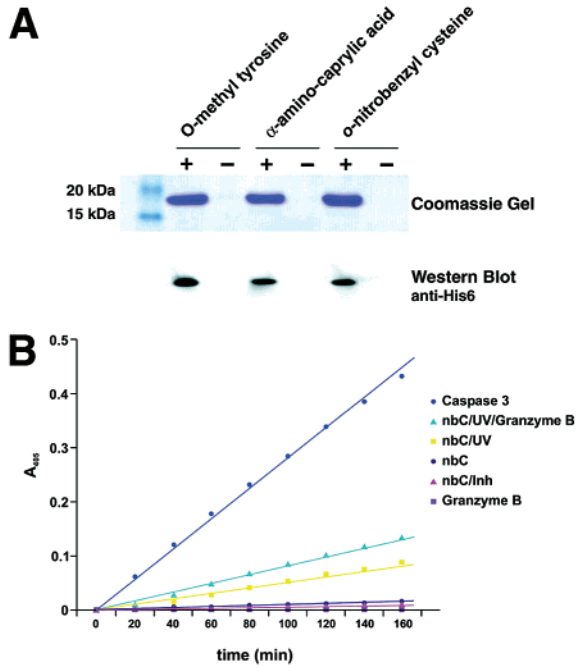

Figure 2. (A) Expression of hSOD-33TAG-His6 in the presence (+ lanes) and absence (− lanes) of 1 mM unnatural amino acids, detected with both Coomassie blue and anti-His6 antibody after Ni-NTA purification. (B) A 7-amino-4-trifluoromethyl coumarin substrate (absorption at 405 nm) was used to measure caspase 3 activity in an untreated cell lysate (nbC), after irradiation (nbC/UV), after irradiation in the presence of granzyme B (nbC/UV/granzyme B), and in the presence of a caspase 3 inhibitor (nbC/Inh). Commercial recombinant caspase 3 was used as a positive control and granzyme B as a negative control.

synthetase clones (designated C8RS, OMeYRS, and nbCRS) produced protein only in the presence of their corresponding unnatural amino acid (Figure 2A) (Table 1, Supporting Information). In each case, the yield of purified protein is comparable to that produced by the wild-type orthogonal suppressor tRNA/synthetase pair (~0.6 mg/L). In addition, electrospray ionization ion-trap mass spectrometry revealed that the total mass of the intact hSOD protein is consistent with site-specific incorporation of the unnatural amino acid in response to the amber codon (Table 2, Supporting Information). These results indicate that all three amino acids are incorporated into proteins with high efficiency and high fidelity.

The efficient and selective biosynthetic incorporation of *o*-nitrobenzyl cysteine into proteins should allow both temporal and spatial photoregulation of proteins with essential cysteine residues.[8] UV irradiation of the nbC-33-hSOD mutant with a UV lamp for 20 min resulted in cleavage of the benzylic C−S bond and subsequent release of the free thiol group and *o*-nitrosobenzaldehyde. A 70% conversion to the decaged hSOD was detected by RPLC and full-length mass spectral analysis (observed mass 16 584.2 Da; calculated mass 16 585.2 Da); nbC-33-hSOD was the only other observed protein. This is in agreement with previous studies of the photodeprotection of Boc-nbC.[9] To further illustrate the utility of *o*-nitrobenzyl cysteine, the active-site cysteine of the human proapoptotic protein caspase 3 was substituted with *o*-nitrobenzyl cysteine. Like other caspases involved in apoptosis, this enzyme exists in the cytosol as an inactive zymogen. A two-chain mature enzyme is generated after precise cleavage at internal aspartate residues, either by activated caspase 3 or by other proteases such as the serine protease granzyme B.[10] To photocage caspase 3 activity, the codon for the catalytic Cys163 residue was mutated to TAG, and the protein was expressed under an inducible galactose promoter in the presence of 1 mM *o*-nitrobenzyl cysteine, nbCRS, and $Leu5_{CUA}$. Caspase 3 activity was measured in cell lysates after 10 min of irradiation with or without further incubation with granzyme B. Only after photolysis was there any detectable protease activity in the cell lysate (Figure 2B), both in the presence and in the absence of granzyme B. Under these conditions, approximately 40% of the caged caspase was converted to the active enzyme. Moreover, expression of wild-yype caspase 3 is toxic to yeast, whereas expression of the nbC163 caspase mutant had no effect on growth rates.

In conclusion, a new orthogonal pair has been generated in yeast, and three structurally distinct unnatural amino acids have been efficiently incorporated into proteins using this new orthogonal pair, including a photocaged cysteine. It should be possible to evolve LRS mutants to incorporate additional unnatural amino acids, including caged amino acids that can be deprotected at longer wavelengths in vivo (e.g., nitroveratyrl Cys, Tyr or Ser), as well as fluorophores and spin-labeled amino acids.

Acknowledgment. N.W. is supported by a fellowship from the Canadian Institutes of Health Research. A.D. gratefully acknowledges a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. T.A.C. is supported by a National Research Service Award from the National Institutes of Health. This work was supported by funding from the National Institutes of Health (GM62159). This is manuscript CH-16544 of The Scripps Research Institute.

Supporting Information Available: Experimental protocol for library construction and selection; synthesis of *o*-nitrobenzyl cysteine; caspase 3 enzymatic assays; sequences of all evolved leucyl synthetases; and protein expression experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020−4. (b) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026−7. (c) Alfonta, L.; Zhang, Z.; Uryu, S.; Loo, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 14662−3. (d) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56−61. (e) Zhang, Z.; Gildersleeve, J.; Yang, Y. Y.; Xu, R.; Loo, J. A.; Uryu, S.; Wong, C. H.; Schultz, P. G. *Science* **2004**, *303*, 371−3. (f) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. *Nature Biotech*. **2004**, in press. (g) Wang, L.; Schultz, P. G. *Angew. Chem*. **2004**, in press.

(2) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964−7.

(3) 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*, 7566−71.

(4) Cusack, S.; Yaremchuk, A.; Tukalo, M. *EMBO J.* **2000**, *19*, 2351−61.

(5) Soma, A.; Himeno, H. *Nucleic Acids Res.* **1998**, *26*, 4374−81.

(6) (a) Soma, A.; Kumagai, R.; Nishikawa, K.; Himeno, H. *J. Mol. Biol.* **1996**, *263*, 707−14. (b) Asahara, H.; Himeno, H.; Tamura, K.; Hasegawa, T.; Watanabe, K.; Shimizu, M. *J. Mol. Biol.* **1993**, *231*, 219−29.

(7) Chin, J. W.; Cropp, T. A.; Chu, S.; Meggers, E.; Schultz, P. G. *Chem. Biol.* **2003**, *10*, 511−9.

(8) (a) Self, C. H.; Thompson, S. *Nat. Med.* **1996**, *2*, 817−20. (b) Philipson, K. D.; Gallivan, J. P.; Brandt, G. S.; Dougherty, D. A.; Lester, H. A. *Am. J. Physiol. Cell Physiol*. **2001**, *281*, C195−206. (c) Pollitt, S. K.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl*. **1998**, *37*, 2104−7.

(9) Smith, A. B. I.; Savinov, S. N.; Manjappara, U. V.; Chaiken, I. M. *Org. Lett.* **2002**, *4*, 4041−4.

(10) Trapani, J. A. *Genome Biol*. **2001**, *2* (12), REVIEWS3014.1−3014.7.

JA040175Z