Bioorganic & Medicinal Chemistry Letters 18 (2008) 6004–6006

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins

Eric Brustad [a], Mark L. Bushey [a], Ansgar Brock [b], Johnathan Chittuluru [a], Peter G. Schultz [a,*]

[a] Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA
[b] Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121, USA

ARTICLE INFO

Article history:
Received 4 August 2008
Revised 10 September 2008
Accepted 10 September 2008
Available online 17 September 2008

Keywords:
Unnatural amino acids
Methionine homologs
Cysteine homologs
Alanine homologs

ABSTRACT

A mutant *Escherichia coli* leucyl-tRNA synthetase has been evolved for the selective incorporation of the methionine homolog **1** into proteins in yeast. This single aminoacyl-tRNA synthetase is capable of charging an amber suppressor *Ec*tRNA$^{Leu}_{CUA}$ with at least eight different amino acids including methionine and cysteine homologs, as well as straight chain aliphatic amino acids. In addition we show that incorporation yields for these amino acids can be increased substantially by mutations in the editing CP1 domain of the *E. coli* leucyl-tRNA synthetase.

Site-directed mutagenesis has proven to be a powerful tool in protein biochemistry, making it possible to systematically analyze protein structure–function relationships, as well as to generate proteins with new or enhanced properties. However, this method is limited to substitutions among the 20 canonical amino acids, severely restricting the nature of structural modifications that can be made. We have developed methods that significantly expand the number of amino acids that can be genetically encoded in bacteria, yeast, and mammalian cells.[1–3] This approach not only allows building blocks with novel properties and reactivities (e.g., photocrosslinkers,[4] metal ion chelators,[5] α-hydroxy acids,[6] keto containing amino acids,[7] fluorophores,[8,9] etc.) to be site-specifically introduced into proteins, it also allows one to precisely tailor the structure of a particular amino acid in a manner that cannot generally be achieved with classical mutagenesis methods. In an effort to provide new thiol and thioether containing amino acids, we now report the generation of an orthogonal, amber suppressor tRNA/aminoacyl-tRNA synthetase pair in yeast that incorporates a variety of methionine (**1** and **2**, Fig. 1) and cysteine homologs (**3**–**5**), as well as straight chain aliphatic amino acids (**6**–**8**). These amino acids should prove useful in exploring the effects of side-chain packing and disulfide bonds on protein folding and stability, and may provide novel ligands to metal ions in metalloproteins.

To incorporate methionine and cysteine homologs into proteins in yeast, we used an *Escherichia coli*-derived amber suppressor leucyl-tRNA/leucyl-tRNA synthetase (*Ec*tRNA$^{Leu}_{CUA}$/*Ec*LRS) pair that has

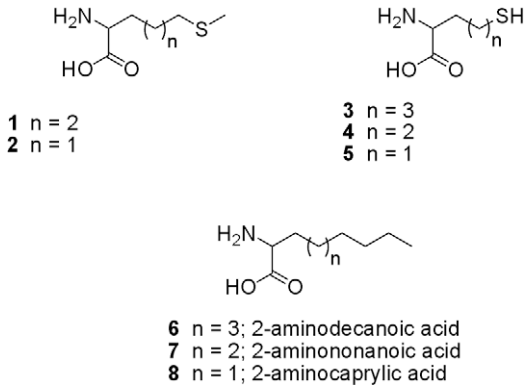

Figure 1. Structures of thiol, thioether, and aliphatic amino acids.

been previously shown to function efficiently in *Saccharomyces cerevisiae*, but does not cross react with any of the endogenous tRNAs and aminoacyl-tRNA synthetases in yeast (i.e., this pair is orthogonal in yeast). Moreover, this aminoacyl-tRNA synthetase has already been mutated to incorporate the unnatural amino acid 2-aminocaprylic acid (**8**, Fig. 1),[10] and was therefore expected to provide a good starting point for evolving active sites that bind amino acids **1**–**5**. To alter the specificity of *Ec*LRS, the active site residues Met40, Leu41, Tyr499, Tyr527, and His 537 were randomized by NNK saturation mutagenesis. Since no structure for the *Ec*LRS has been published to date, these sites were chosen on the basis of the crystal structure of the highly homologous *Thermus*

* Corresponding author.
E-mail address: schultz@scripps.edu (P.G. Schultz).

0960-894X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2008.09.050

thermophilus leucyl-tRNA synthetase (*Tt*LRS).[11] The construction of these libraries has been described elsewhere.[10] Selections were performed in *S. cerevisiae* strain MAV203:pGAD-Gal4(2TAG). This strain harbors a Gal4 transcriptional activator that has amber (TAG) codons at two permissive sites, Thr44 and Arg110. Suppression at these site controls the expression of three reporter genes *lacZ*, *his3*, and *ura3*. For positive selections, aminoacylation of *Ec*tRNA$^{Leu}_{CUA}$ in the presence of 1 mM unnatural amino results in growth on media deficient in uracil (−Ura). Aminoacyl-tRNA synthetases that charge *Ec*tRNA$^{Leu}_{CUA}$ with endogenous amino acids are then removed by selection in the absence of unnatural amino acid on 0.1% 5-fluoroorotic acid, which is converted to a toxic compound by the *ura3* gene product.[12]

Selections were initially performed using the methionine and cysteine homologs **1** and **3**, respectively, and 2-aminocaprylic acid (**8**) as a positive control. After selection on **1** (3 positive and 2 negative rounds) two unique clones were obtained that permitted growth on −Ura media in the presence of 1 mM of the unnatural amino acid. Similarly nine unique clones were obtained for the positive control (**8**). Unfortunately, no hits were obtained from selection on the long-chain cysteine analog, **3**. Table 1 summarizes the active site mutations obtained for each of the selected *Ec*LRSs. Analysis of these sequences reveals that residues 40, 41, 499, and 527 are enriched for hydrophobic residues (Ala, Val, Leu, Met, and Ile) which likely allow for packing around the hydrophobic alkyl chain of both **1** and **8**. His537 showed complete convergence to glycine for all selected synthetases. This same mutation was present in previously reported aminoacyl-tRNA synthetases selected against 2-aminocaprylic acid.[10] Analysis of the *Tt*LRS crystal structure shows that His537 (His545 in *Tt*LRS) lies just above the leucine side chain.[11] Mutation of this residue to glycine likely produces a larger cavity in the active site which allows longer amino acids such as **1** and **8** to bind while at the same time reducing the affinity of the enzyme for its cognate leucine amino acid.

The striking similarity between *Ec*LRS mutants isolated from selections with **1** and **8**, as well as the structural similarity of the two amino acids suggested that there could be some cross reactivity of the mutant enzymes to both amino acids. To examine this potential, *Ec*LRS mutant DHE6 (*Ec*LRS-DHE6) was cotransformed into SCY4 yeast cells with plasmid pC1hSOD33TAG that contains a 6× His tagged variant of the human superoxide dismutase (hSOD) gene with an amber codon in place of Trp33 and a constitutive TDH3 promoter.[13] Cells were grown in the presence of either 1 mM **1** or 1 mM **8**, or no amino acid for 48 h at 30 °C. Cells were then lysed and hSOD was purified by Ni–NTA chromatography. SDS–PAGE analysis confirmed that hSOD was produced when either **1** or **8** was added to the media; however, only a low background level of hSOD was observed in the absence of amino acid (Fig. 2), indicating little activity toward Leu, Ile, Met, or other hydrophobic amino acids. Incorporation of each amino acid was confirmed by MALDI-TOF mass spectrometry (**1** calcd 16,643 M + acetyl, obsd. 16643; **8** calcd 16,625 M + acetyl, obsd 16626; Fig. S1a and b). Yields for expression of **1** and **8** in this system were 0.5 and 0.3 mg/L, respectively.

Table 1
Active site mutations for *Ec*LRSs selected on either dihomomethionine (**1**)[a] or 2-aminocaprylic acid (**8**)[b]

Synthetase	40	41	499	527	537
Wild-type	Met	Leu	Tyr	Tyr	His
DHA5[a]	Leu	Trp	Ile	Thr	Gly
DHE6[a]	Ile	Leu	Ile	Ala	Gly
CapB11[b]	Ile	Leu	His	Val	Gly
CapD7[b]	His	Leu	Ser	Ser	Gly
CapE9[b]	Ile	Met	Leu	Thr	Gly
CapE11[b]	Ala	Thr	Ser	Val	Gly
CapF9[b]	Val	Ser	Gln	Val	Gly
CapG5[b]	Ala	Val	Trp	Met	Gly
CapG9[b,d]	Ala	Val	Trp	Met	Gly
CapH9[b,d]	Thr	Pro	Ile	Met	Gly
Cap2X[b,e]	Ala	Lys	Ser	Leu	Gly
NW:1D7[c]	Ala	Ala	Pro	Val	Gly
NW:1G8[c]	Val	Met	Leu	Leu	Gly
NW:2F2[c]	His	Pro	Ala	Met	Gly
NW:2F5[c]	Val	Tyr	Leu	Leu	Gly

[a] Synthetase clones selected on 1 mM **1**.
[b] Synthetase clones selected on 1 mM 2-aminocaprylic acid (**8**).
[c] Synthetase clones selected on 1 mM 2-aminocaprylic acid previously published in Ref. 9.
[d] Mutants contain an additional point mutation of Glu410Gly that was not included in the original library.
[e] Mutant contains an additional point mutation of Glu20Lys that was not included in the original library.

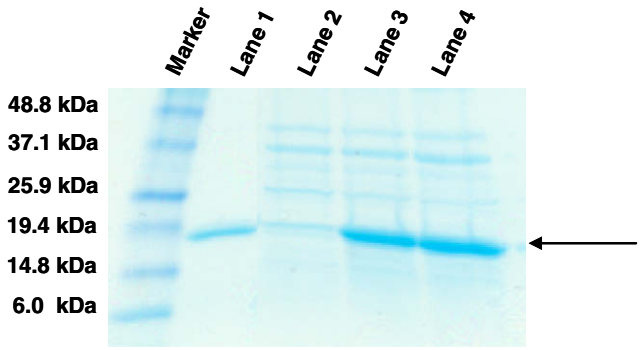

Figure 2. *Ec*LRS-DHE6-dependent expression of hSOD-Trp33TAG with either 1 mM **1** or **8**. Lane 1 contains wild-type hSOD as a reference; lane 2 contains protein expressed in the absence of unnatural amino acid; lane 3 contains protein expressed in the presence of 1 mM **1**; and lane 4 contains protein expressed in the presence of 1 mM **8**.

To further examine the substrate promiscuity of the selected aminoacyl-tRNA synthetase DHE6, MAV203:pGAD-Gal4(2TAG) cells containing *Ec*LRS-DHE6 were spotted on minimal −Ura agar plates that were individually supplemented with amino acids **1**–**8**, as well as *N*-methyllysine or *N*,*N*-dimethyllysine, which are also structurally similar. Colonies were observed after 48 h on plates supplemented with methionine analogs **1** and **2**, cysteine analogs **3**, **4**, and **5** and aliphatic amino acids **6**, **7**, and **8**. No growth was observed on either lysine derivative (Fig. 3). Optimal cell growth was observed on amino acids with side chains six atoms in length (**1**, **3**, and **8**). Cells grown in the presence of amino acids with longer and shorter chains also survived but grew less well in this assay. Amino acids **1**, **6**, **7**, and **8** were independently introduced into hSOD-Trp33TAG as described above and their incorporation was confirmed by MALDI-TOF (Figs. S1a and b, and S2a and b). Incorporation of cysteine analogs **3** and **4** into hSOD was confirmed by tryptic digest followed by LC–MS/MS (Fig. S3a and b). Database

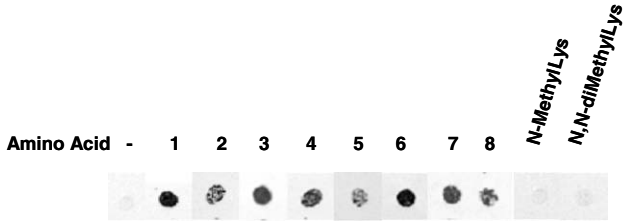

Figure 3. Promiscuous activity of EcLRS-DHE6. Cells harboring pGAD-Gal4(2TAG) and EcLRS-DHE6 were grown on −Ura minimal media in the presence of the indicated amino acids. Pictures of respective colonies are shown above. Numbers correspond to amino acids listed in Figure 1.

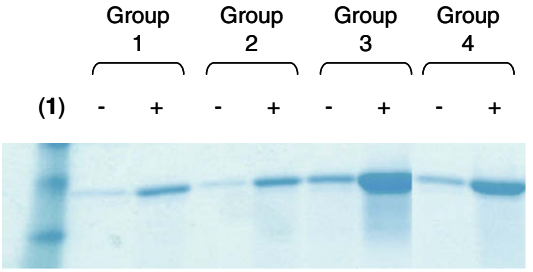

Figure 4. Expression of hSOD-Trp33TAG by editing domain mutants of the *Ec*LRS-DHE6 in the absence (−) or presence (+) of 1 mM **1**. Group 1, unmodified *Ec*LRS-DHE6; group 2, *Ec*LRS + 3 tRNAs; group 3, *Ec*LRS-DHE6-Thr247Val/Thr248Val; group 4, *Ec*LRS-DHE6-Thr252Phe.

searches of other tryptic fragments revealed an increased background incorporation of Leu, Gln, Thr, and Lys. While DHE6 shows activity for amino acids **1**–**8**, further evolution will be required to achieve complete selectivity. Protein yields for each amino acid were comparable to that of **1** and **8** (data not shown).

It is known that the wild-type *E. coli* leucyl-tRNA synthetase itself aminoacylates endogenous amino acids other than its cognate leucine. In vitro studies have shown that the enzyme can charge *E. coli* leucyl-tRNA with similar amino acids, such as isoleucine and methionine.[13,14] To overcome this promiscuity, the enzyme has evolved an aminoacyl editing domain (the CP1 domain) that removes misacylated amino acids, thereby maintaining high fidelity for leucine incorporation.[13] One potential consequence of the editing domain with our selection methodology is that the enzyme may remove an unnatural amino acid of interest after it is attached to the tRNA, decreasing the yield of incorporation. Mutation of Thr252Ala in the CP1 domain is known to decrease the ability of the editing domain to discriminate leucine from other amino acids while mutation of Thr252 to Phe has been shown to decrease editing domain activity presumably by obstructing the editing domain binding site.[15] Two other mutations, Thr247Val and Thr248Val, have been shown to decrease editing domain activity by preventing proper orientation of the aminoacylated tRNA into the CP1-binding site.[16] Finally, we have previously shown that Thr252, which lies at the back of the editing domain, can be mutated to alanine to increase the fidelity of unnatural amino acid incorporation.[9]

To test whether mutations to the editing domain would increase the fidelity or yield of unnatural amino acid incorporation by *Ec*LRS-DHE6, two mutants (Thr247Val/Thr248Val and Thr252Phe) were constructed using a modified quick change mutagenesis procedure.[17] These mutants were cotransformed with pC1hSOD33TAG, and hSOD was expressed as described above in the presence or absence of **1**. In addition, a second plasmid containing the unmutated *Ec*LRS-DHE6 plus an additional two copies of $Ec\text{tRNA}^{\text{Leu}}_{\text{CUA}}$ was also created to examine the effects of increased tRNA levels on expression. We have previously shown that increased tRNA levels can have beneficial effects on the suppression yields of certain unnatural amino acids. As shown in Figure 4, increasing tRNA copy number from 1 to 3 produced only a slight increase in hSOD levels in the presence of **1**. Mutations to the editing domain, however, showed significant increases in hSOD production (>10-fold) versus that of the wild-type editing domain. As expected, the background incorporation in the absence of unnatural amino acid is also increased somewhat due to the inability of the synthetase to remove misacylated endogenous amino acids.

In summary, we have demonstrated that a mutant leucyl-tRNA synthetase that was evolved to incorporate the long-chain methionine homolog (**1**) into proteins can also accommodate various long-chain methionine, cysteine, and alkyl analogs. Such promiscuous protein activity is not uncommon in nature. In fact enzymes that display low levels of promiscuous (“moonlighting”) activity have been implicated in the evolution of novel enzymatic function through gene duplication and subsequent random mutagenesis.[18,19] In this regard, we are currently applying error-prone PCR and DNA shuffling to increase the activity and specificity of *Ec*LRS-DHE6 towards each of the aforementioned amino acids as well as other amino acids of similar structure including methylated and acetylated lysine analogs. In addition, it is likely that other evolved aminoacyl-tRNA synthetases may display promiscuous activity profiles towards various unnatural amino acids and that the simple screening of existing mutant aminoacyl-tRNA synthetases may yield novel hits without the need for library design and selection.

Acknowledgments

We thank Emily Remba for her assistance during the preparation of this manuscript and The Scripps Research Institute Center for Mass Spectrometry for their assistance in gathering the MALDI-TOF data. This work is supported by the US Department of Energy, Division of Materials Sciences, under Award No. DE-FG03-00ER46051 and the Skaggs Institute for Chemical Biology.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2008.09.050.

References and notes

1. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.
2. Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964.
3. Liu, W.; Brock, A.; Chen, S.; Chen, S.; Schultz, P. G. *Nat. Methods* **2007**, *4*, 239.
4. Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020.
5. Xie, J. M.; Liu, W. S.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2007**, *46*, 9239.
6. Guo, J. T.; Wang, J. Y.; Anderson, J. C.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 722.
7. Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56.
8. Wang, J. Y.; Xie, J. M.; Schultz, P. G. J. *Am. Chem. Soc.* **2006**, *128*, 8738.
9. Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J. W.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9785.
10. Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. J. *Am. Chem. Soc.* **2004**, *126*, 14306.
11. Tukalo, M.; Yaremchuk, A.; Fukunaga, R.; Yokoyama, S.; Cusack, S. *Nat. Struct. Mol. Biol.* **2005**, *12*, 923.
12. Boeke, J. D.; Lacroute, F.; Fink, G. R. *Mol. Gen. Genet.* **1984**, *197*, 345.
13. Chen, J. F.; Guo, N. N.; Li, T.; Wang, E. D.; Wang, Y. L. *Biochemistry* **2000**, *39*, 6726.
14. Liu, Y. Q.; Liao, J.; Zhu, B.; Wang, E. D.; Ding, J. P. *Biochem. J.* **2006**, *394*, 399.
15. Mursinna, R. S.; Lee, K. W.; Briggs, J. M.; Martinis, S. A. *Biochemistry* **2004**, *43*, 155.
16. Zhai, Y. X.; Martinis, S. A. *Biochemistry* **2005**, *44*, 15437.
17. Zheng, L.; Baumann, U.; Reymond, J. L. *Nucleic Acids Res.* **2004**, *32*, 5.
18. Aharoni, A.; Gaidukov, L.; Khersonsky, O.; Gould, S. M.; Roodveldt, C.; Tawfik, D. S. *Nat. Genet.* **2005**, *37*, 73.
19. Jensen, R. A. *Annu. Rev. Microbiol.* **1976**, *30*, 409.