J|A|C|S
COMMUNICATIONS

Published on Web 02/06/2002

Adding L-3-(2-Naphthyl)alanine to the Genetic Code of *E. coli*

Lei Wang,[†] Ansgar Brock,[‡] and Peter G. Schultz*[,‡,§]

Department of Chemistry, University of California at Berkeley, Berkeley, California 94720, Genomics Institute of the Novartis Research Foundation, 3115 Merryfield Row, San Diego, California 92121, and Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received October 8, 2001

The ability of chemists to rationally manipulate the structure and function of proteins is quite limited in comparison to that for smaller organic molecules, despite the central role of these macromolecules in nature. The development of methods that make it possible to increase the number of genetically encoded amino acids in living organisms beyond the common 20 would provide a tremendous opportunity to better understand and possibly enhance protein (and perhaps organismal) function. To this end, we have developed a general strategy in which additional components including a novel tRNA−codon pair, an aminoacyl-tRNA synthetase, and an amino acid are added to the translational machinery of the cell.[1] This new set of components functions orthogonally to the counterparts of the common 20 amino acids; that is, the orthogonal synthetase (and only this synthetase) aminoacylates the orthogonal tRNA (and only this tRNA) with the unnatural amino acid only, and the resulting acylated tRNA inserts the unnatural amino acid in response to only the unique codon. Using this strategy, we have shown recently that it is possible to augment the protein biosynthetic machinery of *Escherichia coli* to accommodate the additional genetically encoded amino acid *O*-methyl-L-tyrosine with fidelity close to that of the common amino acids. We now report the site-specific incorporation of a second unnatural amino acid, L-3-(2-naphthyl)alanine into proteins in *E. coli*, suggesting that this overall scheme may be applicable to a host of amino acids. Preliminary results suggest that alkyl-, aryl-, acyl-, and azido-substituted amino acids can also be selectively incorporated.[2]

An amber stop codon and its corresponding orthogonal amber suppressor tRNA, mutRNA$^{Tyr}_{CUA}$, were selected to encode the unnatural amino acid.[3] The *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS) was used as the starting point for the generation of an orthogonal synthetase with unnatural amino acid specificity. This TyrRS does not aminoacylate any endogenous *E. coli* tRNAs[4] but aminoacylates the mutRNA$^{Tyr}_{CUA}$ with tyrosine.[5] L-3-(2-Naphthyl)alanine was chosen for this study since it represents a significant structural perturbation from tyrosine and may have novel packing properties. To change the amino acid specificity of the TyrRS so that it charges the mutRNA$^{Tyr}_{CUA}$ with L-3-(2-naphthyl)-alanine and not any common 20 amino acids, a library of *M. jannaschii* TyrRS mutants was generated and screened. On the basis of an analysis of the crystal structure of the homologous TyrRS from *Bacillus stearothermophilus*,[6] five residues (Tyr[32], Asp[158], Ile[159], Leu[162], and Ala[167]) in the active site of *M. jannaschii* TyrRS that are within 7 Å of the para position of the aryl ring of tyrosine were mutated (Figure 1).[7] To reduce the wild-type synthetase contamination in the following selection, these residues (except Ala[167]) were first mutated to alanine. The resulting inactive Ala$_5$

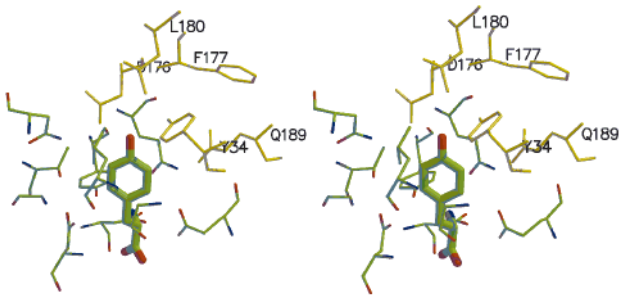

Figure 1. The active site of TyrRS. Residues from *B. stearothermophilus* TyrRS are shown. Corresponding residues from *M. jannaschii* TyrRS are Tyr[32] (Tyr[34]), Asp[158] (Asp[176]), Ile[159] (Phe[177]), Leu[162] (Leu[180]), and Ala[167] (Gln[189]) with *B. stearothermophilus* TyrRS residues in parentheses. Randomly mutated resides are shown in yellow.

TyrRS gene was used as a template for polymerase chain reaction (PCR) random mutagenesis with oligonucleotides bearing random mutations at the corresponding sites.

The mutant TyrRS library was first passed through a positive selection based on suppression of an amber stop codon at a nonessential position (Asp112) in the chloramphenicol acetyltransferase (CAT) gene. Cells transformed with the mutant TyrRS library and the mutRNA$^{Tyr}_{CUA}$ gene were grown in minimal media containing 1 mM L-3-(2-naphthyl)alanine and 80 μg/mL chloramphenicol. Cells can survive only if a mutant TyrRS aminoacylates the mutRN A$^{Tyr}_{CUA}$ with either natural amino acids or L-3-(2-naphthyl)alanine. The surviving cells were then grown in the presence of chloramphenicol and the absence of the unnatural amino acid. Those cells that did not survive must encode a mutant TyrRS that charges the mutRNA$^{Tyr}_{CUA}$ with L-3-(2-naphthyl)alanine, and were picked from a replica plate supplied with the unnatural amino acid. After three rounds of positive selection followed by a negative screen, four TyrRSs were characterized using an in vivo assay based on the suppression of the Asp112TAG codon in the CAT gene (Table 1). In the absence of L-3-(2-naphthyl)alanine, cells expressing the selected TyrRS and the mutRNA$^{Tyr}_{CUA}$ survived in 25−35 μg/mL chloramphenicol on minimal media plates containing 1% glycerol and 0.3 mM leucine (GMML plate); in the presence of L-3-(2-naphthyl)alanine, cells survived in 100−120 μg/mL chloramphenicol on GMML plates. Compared to the IC$_{50}$ value in the absence of any TyrRS (4 μg/mL chloramphenicol), these results indicate that the selected TyrRSs accept L-3-(2-naphthyl)alanine, but also still charge natural amino acids to some degree.

To further reduce the activity of the mutant TyrRS toward natural amino acids, one round of DNA shuffling was carried out using the above four mutant genes as templates. The resulting mutant TyrRS library was passed through two additional rounds of positive selections and negative screens. One mutant TyrRS (SS12-TyrRS) was evolved, whose activity for natural amino acids was greatly reduced (IC$_{50}$ = 9 μg/mL chloramphenicol) while its activity toward

* To whom correspondence should be addressed. E-mail: schultz@scripps.edu.
† University of California at Berkeley.
‡ Genomics Institute of the Novartis Research Foundation.
§ The Scripps Research Institute.

10.1021/ja012307j CCC: $22.00

Table 1. In Vivo Chloramphenicol Acetyltransferase Assay of Mutant TyrRS[a]

mutant TyrRS	IC_{50} (μg/mL of Chloramphenicol)	
	no L-3-(2-naphthyl)Ala	add L-3-(2-naphthyl)Ala
no TyrRS	4	4
wt TyrRS	240	240
	After Selection	
S1-TyrRS	30	120
S2-TyrRS	30	120
S3-TyrRS	25	110
S4-TyrRS	35	100
	After DNA Shuffling	
SS12-TyrRS	9	150

[a] A pYC-J17 plasmid was used to express the mutRNA$^{Tyr}_{CUA}$ gene and the chloramphenicol acetyltransferase gene with an amber stop codon at Asp112. A pBK plasmid was used to express TyrRS, and was cotransformed with pYC-J17 into *E. coli* DH10B. Cell survival on GMML plates was titrated in the presence of different concentrations of chloramphenicol.

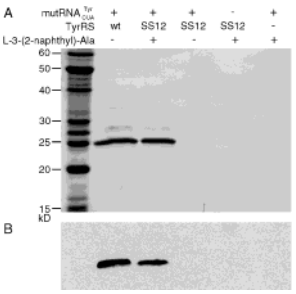

Figure 2. Accumulation of mouse DHFR protein under different conditions. (A) Silver-stained SDS-PAGE gel of purified mouse DHFR. Expression conditions are notated at the top of each lane. The very left lane is molecular weight marker. (B) Western blot of gel in (A).

L-3-(2-naphthyl)alanine was enhanced (IC_{50} = 150 μg/mL chloramphenicol) (Table 1).

An L-3-(2-naphthyl)alanine mutant of mouse dihydrofolate reductase (DHFR) was generated and characterized to confirm the ability of the mutRNA$^{Tyr}_{CUA}$/SS12-TyrRS pair to site-specifically incorporate L-3-(2-naphthyl)alanine in response to an amber stop codon. The Tyr163 codon of the mouse DHFR gene was mutated to TAG, and a His6 tag was added to the COOH-terminus of DHFR to facilitate protein purification using Ni^{2+} affinity chromatography. As a positive control, wild-type *M. jannaschii* TyrRS was coexpressed with the mutRNA$^{Tyr}_{CUA}$, resulting in efficient suppression of the TAG codon with tyrosine (Figure 2). When SS12-TyrRS was coexpressed with the mutRNA$^{Tyr}_{CUA}$ in the presence of 1 mM L-3-(2-naphthyl)alanine, full-length mouse DHFR was also generated (with yield of 2.2 mg/L in liquid GMML minimal medium). In the absence of either L-3-(2-naphthyl)alanine, mutRNA$^{Tyr}_{CUA}$, or SS12-TyrRS, no full-length DHFR was produced. A penta-His antibody was used to detect the His6 tag at the COOH-terminus of DHFR in a Western blot. No DHFR could be detected in the absence of each of the above three components.

Tryptic digests of the L-3-(2-naphthyl)alanine mutant of mouse DHFR were analyzed by MALDI FT-ICR and liquid chromatography tandem mass spectrometry to confirm unambiguously the incorporation of L-3-(2-naphthyl)alanine. The peptide map of the internally calibrated digest shows a major peak at 1867.962, which is within 3.5 ppm of the theoretical mass of the tryptic peptide LLPEX*TGVLSEVQEEK where X* represents the L-3-(2-naphthyl)alanine residue (Pro164 was mutated to Thr to improve the amber suppression efficiency). Further, the interpreted tandem mass spectrum of precursor ion at *m*/*z* 934.5, which corresponds to the doubly charged ion of the peptide of interest is shown in Figure 3. The sequence information gleaned from the spectrum clearly demonstrates the site-specific incorporation of L-3-(2-naphthyl)-

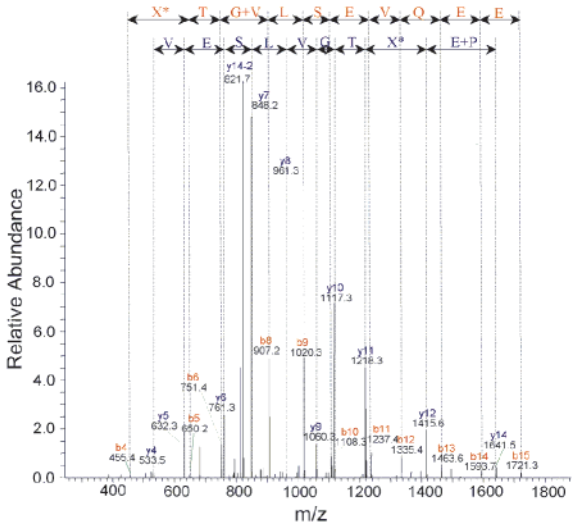

Figure 3. Tandem mass spectrum of the tryptic peptide LLPEX*TGVLSEVQEEK (X* = L-3-(2-naphthyl)alanine). The sequence can be read from the annotated b (red) or y (blue) ion series; even so, b7 and y13 are not observed. The base peak 821.7 (100%) assigned to the doubly charged y14 ion is truncated for clarity.

alanine into the protein. Neither peptide maps nor LC MS/MS runs produced any indication of mutants in which the L-3-(2-naphthyl)-alanine residue is substituted by other amino acids. The signal-to-noise ratio of more than 1500 observed in the peptide maps suggests a fidelity in the incorporation of L-3-(2-naphthyl)alanine of better than 99.8%.

The evolved SS12-TyrRS has the following mutations: Tyr^{32}→Leu^{32}, Asp^{158}→Pro^{158}, Ile^{159}→Ala^{159}, Leu^{162}→Gln^{162}, and Ala^{167}→Val^{167} (Figure 1). Based on the crystal structure of the homologous *B. stearothermophilus* TyrRS, we speculate that the mutations of Tyr^{32}→Leu^{32} and Asp^{158}→Pro^{158} result in the loss of hydrogen bonds between Tyr^{32}, Asp^{158}, and the native substrate tyrosine, thus disfavoring the binding of tyrosine to SS12-TyrRS. Most residues are mutated to amino acids with hydrophobic side chains, which are expected to favor binding of L-3-(2-naphthyl)-alanine. Efforts to solve the crystal structure of the wild-type *M. jannaschii* TyrRS and the evolved SS12-TyrRS are under way.

In summary, the cell growth, protein expression, and mass spectrometry experiments demonstrate that the mutRNA$^{Tyr}_{CUA}$/SS12-TyrRS pair is capable of selectively inserting L-3-(2-naphthyl)alanine into proteins in response to the amber codon with fidelity rivaling that of the natural amino acids. This result, which involves an amino acid that is structurally distinct from tyrosine, suggests that the above methodology should be applicable to a variety of unnatural amino acids. We are currently evolving synthetases that have specificities for unnatural amino acids with novel chemical and physical properties, as well as expanding the methodology to eukaryotic cells and four-base codon−anticodon pairs.

Acknowledgment. We acknowledge support of this work by NIH (GM62159).

References

(1) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498−500. Wang, L.; Schultz, P. G. *Chem. Commun.* **2002**. In press.
(2) Schultz, P. G. et al. Unpublished results.
(3) Wang, L.; Schultz, P. G. *Chem. Biol.* **2001**, *8*, 883−890.
(4) Steer, B. A.; Schimmel, P. *J. Biol. Chem.* **1999**, *274*, 35601−35606.
(5) Wang, L.; Magliery, T. J.; Liu, D. R.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 5010−5011.
(6) Brick, P.; Bhat, T. N.; Blow, D. M. *J. Mol. Biol.* **1989**, *208*, 83−98.
(7) No synthetases specific for L-3-(2-naphthyl)alanine were selected from the mutant TyrRS library reported in ref 1a.

JA012307J