Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo*

David R. Liu*, Thomas J. Magliery*†, Miro Pastrnak*†, and Peter G. Schultz*‡

*Howard Hughes Medical Institute, Department of Chemistry, and †Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

ABSTRACT In an effort to expand the scope of protein mutagenesis, we have completed the first steps toward a general method to allow the site-specific incorporation of unnatural amino acids into proteins *in vivo*. Our approach involves the generation of an "orthogonal" suppressor tRNA that is uniquely acylated in *Escherichia coli* by an engineered aminoacyl-tRNA synthetase with the desired unnatural amino acid. To this end, eight mutations were introduced into tRNA$_2^{Gln}$ based on an analysis of the x-ray crystal structure of the glutaminyl-tRNA aminoacyl synthetase (GlnRS)–tRNA$_2^{Gln}$ complex and on previous biochemical data. The resulting tRNA satisfies the minimal requirements for the delivery of an unnatural amino acid: it is not acylated by any endogenous *E. coli* aminoacyl-tRNA synthetase including GlnRS, and it functions efficiently in protein translation. Repeated rounds of DNA shuffling and oligonucleotide-directed mutagenesis followed by genetic selection resulted in mutant GlnRS enzymes that efficiently acylate the engineered tRNA with glutamine *in vitro*. The mutant GlnRS and engineered tRNA also constitute a functional synthetase–tRNA pair *in vivo*. The nature of the GlnRS mutations, which occur both at the protein–tRNA interface and at sites further away, is discussed.

To expand our ability to probe the relationship between protein structure and function, as well as engineer new proteins with altered properties, we earlier developed an *in vitro* method that makes it possible to site-specifically incorporate into proteins amino acids not specified in the genetic code (1). Using this method, a large number of amino acids with novel structural, electronic, and spectroscopic properties have been introduced into proteins to probe protein stability, biomolecular recognition, and catalysis. The development of a general approach for the site-specific incorporation of unnatural amino acids into proteins *in vivo*, directly from the growth media, would greatly enhance the power of this technique. For example, the ability to synthesize large quantities of proteins containing heavy atoms would facilitate protein structure determination, and the ability to site-specifically substitute fluorophores or photocleavable groups into proteins in living cells would provide powerful tools for studying protein function *in vivo*. Alternatively, one might be able to enhance the properties of proteins by providing building blocks with new functional groups, such as a keto-containing amino acid. To this end, we have completed the first steps of a multistage approach for site-specifically incorporating unnatural amino acids into proteins *in vivo*.

Our strategy, similar to that used in the *in vitro* method, relies on suppression of amber codons by tRNAs aminoacylated with unnatural amino acids. Indeed, both natural and engineered suppressor tRNAs have been used to make mutations involving the common amino acids in *Escherichia coli* (2). The first re-

quirement of our strategy is a suppressor tRNA that is not a substrate for any endogenous aminoacyl-tRNA synthetase yet functions efficiently during translation. In addition, an aminoacyl-tRNA synthetase that uniquely recognizes and acylates the orthogonal suppressor tRNA, but not any endogenous tRNAs, is required. This orthogonal synthetase must then be engineered to uniquely acylate the tRNA with the desired unnatural amino acid, but not with any other amino acid. A final requirement is that the unnatural amino acid must be transported from the growth medium to the cytoplasm (or be biosynthesized *in vivo*) with reasonable efficiency. Our first target will be the construction of an engineered strain of *E. coli* that site-specifically inserts a keto-containing amino acid into proteins. The keto group, which is a central functional group in chemistry due to its novel reactivity, is notably absent from the genetically encoded amino acids. In addition, this group provides a unique chemical handle that may make it possible to site-specifically modify proteins *in vitro* or *in vivo* with fluorescent, spin-labeled, or biologically significant hydrazines or alkoxyamines (3). Ultimately we hope to generalize this approach to include a large number of amino acids, other cell types, and an expanded set of codons.

MATERIALS AND METHODS

Strains and Plasmids. *E. coli* strain *lacZ*$_{am1000}$ (4) was a gift of Hachiro Inokuchi (Kyoto University), strain *HAglns* (5) was a gift of Dieter Söll (Yale University), and strain Δ*lamB* was obtained from the *E. coli* Genetic Stock Center (Yale University). Plasmids for runoff transcription of suppressor tRNAs were derived from pYPhe2 (6), as described below. Plasmids for overexpression of tRNAs and glutaminyl-tRNA synthetases were derived from pACYCsupE and from pBRglnS^wt (7), respectively; both were generously provided by Dieter Söll. Plasmid pBAD/lamB was a gift of Phil Patten (Maxygen, Santa Clara, CA).

Construction and *in Vitro* Assays of Suppressor tRNAs. Genes encoding tRNAs for runoff transcription by T7 RNA polymerase were constructed from two overlapping synthetic oligonucleotides (Genosys, The Woodlands, TX) and inserted between the *Kpn*I and *Hin*dIII sites of pYPhe2 as previously described (6, 8). Genes encoding tRNAs for *in vivo* expression were similarly constructed from two overlapping synthetic oligonucleotides and inserted between the *Eco*RI and *Pst*I sites of pACYCsupE, placing transcription under control of the *lpp* promoter and the *rrn*C terminator. The sequences of the overlapping oligonucleotides used to construct the wild-type (wt) suppressor tRNA$_2^{Gln}$(A36) gene are as follows with the tRNA sequence underlined: 5′-GCGGGGTACCGCTCGAGTAATACGACT-CACTATAGGGGGTATCGCCAAGCGGTAAGGCACCG-GATTCTAATTCCGGC-3′ and 5′-GCGCGCAAGCTT-GGATGGATCACCTGGCTGGGGTACGAGGATTCGAA-CCTCGGAATGCCGGATTTAGAATCCGG-3′. Runoff transcription of the tRNAs was carried out as described (8). *In vitro* transcription and translation reactions were performed using 3 μg

of plasmid containing the *E. coli* chorismate mutase gene bearing an amber mutation at site Gln-88 and 10 μg of suppressor tRNA per 30 μl reaction at a final magnesium concentration of 7 mM (8). Reactions to detect acylation by endogenous *E. coli* amino-acyl-tRNA synthetases used full-length tRNA generated from the runoff transcription of *Bst*NI-digested template DNA described above; reactions to assay ribosomal acceptance used truncated ($-$ CA) tRNA generated from the runoff transcription of *Fok*I-digested DNA. This truncated tRNA was ligated using T4 RNA ligase to the dinucleotide pdCpA, which had been acylated with NVOC-protected valine, and then photodeprotected prior to addition to the transcription and translation reaction (9). The resulting [^{35}S]methionine-labeled crude reaction mixtures were subjected to SDS/PAGE and the amounts of truncated and full-length proteins were quantified using a Molecular Dynamics 445SI PhosphorImager. Suppression efficiencies were defined as the amount of full-length protein divided by the sum of the full-length and truncated products.

DNA Shuffling and Selection of Mutant Glutaminyl-tRNA Aminoacyl Synthetase (GlnRS) Enzymes. GlnRS genes isolated from selections were subjected to DNA shuffling as described (10). The desired 2.0-kb product was gel purified, digested with *Eco*RI and *Hin*dIII, and ligated with one equivalent of precut pBR322. Following the second round of DNA shuffling, the efficiency of the reassembly reaction and the yield of the final product began to decline significantly. These problems were remedied by using primers spaced roughly 300 bp further apart in the initial PCR reaction and by subjecting an 8 μl aliquot of the reassembled products to a second reassembly reaction in a total volume of 100 μl. The final PCR reaction was unchanged. The ligated pBRGlnRS libraries (typically 1–3 μg total DNA) were transformed by electroporation into 320 μl of electrocompetent BT235 cells harboring the pAC123 plasmid. Library sizes varied from 5×10^6 to 2×10^7 transformants. The transformed cells were washed twice with lactose minimal media, then plated on 15 cm lactose minimal media supplemented with 100 μg/ml ampicillin and 72 μg/ml (Rounds 1 and 2) or 20 μg/ml cysteine (Rounds 3–7) and incubated at 30°C. Lactose-utilizing colonies were harvested after 3 days, and the pBRGlnRS DNA was extracted by plasmid purification. This DNA was retransformed at least once into BT235/pAC123 cells and regrown on lactose minimal media for 2–3 days to remove false positives arising from reversion of the *lacZ* or tRNA mutations. Single colonies were picked from the resulting plates and pBRGlnRS genes were transformed into HAPPY101 (5) for purification and *in vitro* characterization.

Purification and *in Vitro* Assays of Mutant GlnRS Enzymes. Mutant enzymes were purified by anion exchange chromatography on a Mono Q HR 10/10 fast protein liquid chromatography column by adapting published protocols (5). To analyze each fraction for the presence of hemagglutinin (HA)-tagged wt GlnRS, aliquots of each column fraction were incubated in a 96-well plate and subjected to ELISA using 12CA5 mouse anti-HA IgG solution (Boehringer Mannheim, 1.0 μg/ml) and goat anti-mouse antibodies conjugated to alkaline phosphatase (Sigma, 2 μg/ml). Mutant GlnRS proteins were recovered in 30–96% purity and free of wt GlnRS. Impurities in the GlnRS proteins were assayed and possessed no glutaminylation activity. Purified enzymes were dialyzed into 50% glycerol, 20 mM Hepes, 50 mM KCl, 2 mM DTT (pH 7.2) prior to storage at -20°C.

Concentrations of GlnRS enzymes were determined by SDS/PAGE and staining with Coomassie blue R250 followed by comparison with known concentrations of BSA (Sigma). Specific activities were determined at concentrations corresponding to intracellular levels of 3 mM ATP (11), 150 μM glutamine (12, 36), and 2 μM tRNA (13). Assays (20 μl total volume) were carried out as described (14) at 30°C and contained 40 mM Hepes (pH 7.2), 10 mM MgCl$_2$, 50 mM KCl, 2 mM 2-mercaptoethanol, ATP, glutamine (1–100 mol% radiolabeled with ^{3}H or ^{14}C, DuPont/NEN), tRNA, and enzyme. Enzyme concentrations were ad-

justed such that less than 10% of limiting substrate (tRNA) was consumed. *E. coli* (wt) tRNA was used as whole tRNA from *E. coli* (Boehringer Mannheim); tRNA^{Gln} constitutes 1.8% of whole *E. coli* tRNA (15).

Construction of *lamB* Amber Mutants and Western Blot Analysis. The *Nsi*I–*Eco*RI fragment of pBAD/lamB, including the *lamB* gene and the adjacent *araC* and P$_{BAD}$ regulatory elements, was inserted between the *Fsp*I and *Eco*RV sites of pAC123 to yield pAC123LamB(wt). Amber mutants of *lamB* at Asn-72 and Asp-200 were generated from mutagenic oligonucleotides by PCR as described (16). The *Mlu*I–*Aat*II fragment encoding the wt *lamB* gene was replaced with the *lamB* amber mutants to yield pAC123LamB(N72Am) and pAC123LamB(D200Am). GS20 competent cells were transformed with pBR322 or with pBRGlnRS plasmids encoding mutant GlnRS proteins. Competent cells containing these plasmids were then transformed with pAC123LamB(wt), pAC123LamB(N72Am), or pAC123LamB(D200Am). Double transformants were grown in 5 ml cultures at 37°C to saturation, and 100 μl of each saturated culture was used to inoculate 5 ml of 2xYT containing antibiotics. The cells were grown to an OD$_{600}$ of 0.5 at 37°C. LamB expression was induced by the addition of 170 μl of 750 mM L-arabinose, followed by incubation at 30°C for 3 h. The cells were pelleted, resuspended in 200 μl of SDS/PAGE loading buffer, and boiled for 5 min. Aliquots of these crude lysates (10 μl) were analyzed on a SDS/10% polyacrylamide gel. The proteins were transferred to a nitrocellulose membrane. LamB protein was detected with a rabbit anti-LamB serum and visualized with an alkaline phosphatase-conjugated secondary antibody (Sigma).

RESULTS AND DISCUSSION

Choice of tRNA-Aminoacyl Synthetase System. The original strategy that was used to generate an orthogonal tRNA for our *in vitro* mutagenesis methodology involved the use of the yeast phenylalanyl suppressor tRNA, which was known not to be a substrate for any *E. coli* aminoacyl-tRNA synthetase (17). One could extend this strategy by introducing an orthogonal tRNA-aminoacyl synthetase pair from yeast into *E. coli*. However, the generation of synthetases that acylate tRNAs with unnatural amino acids will likely require multiple mutations both near and distant from the active site of each synthetase (18). Due to the complexity of this system (complete coverage of the protein sequence space of a typical tRNA-aminoacyl synthetase requires $\approx 10^{600}$ different mutants), it is likely that structural information will be required to guide the mutagenesis experiments. Consequently, we limited our initial choice of system to those aminoacyl-tRNA synthetases that had been structurally characterized as complexes with their cognate tRNAs. Among these, the *E. coli* GlnRS and the *Saccharomyces cerevisiae* aspartyl-tRNA synthetase stand out as promising candidates due to the availability of complete high-resolution crystal structures (19, 20), the wealth of biochemical data that has been accumulated on these systems, and the fact that these enzymes have been functionally expressed in *E. coli*. Our progress toward the generation of an orthogonal tRNA/aminoacyl-tRNA synthetase pair from the *E. coli* GlnRS system is described below.

Design and Characterization of an Orthogonal tRNA. The first requirement of a system to deliver unnatural amino acids site-specifically into proteins *in vivo* is the generation of an orthogonal tRNA, which is not aminoacylated by any endogenous aminoacyl-tRNA synthetase yet possesses the structural determinants needed for efficient translation. The three-dimensional x-ray crystal structure of *E. coli* GlnRS complexed with tRNA$_2^{Gln}$ (19) and of a ternary complex including a glutaminyl adenylate analog (L. Silvian and T. Steitz, personal communication) have both been determined to 2.5-Å resolution. Based on an analysis of these structures, three sites ("knobs") were identified in which tRNA$_2^{Gln}$ makes specific contacts with residues in GlnRS (Fig. 1): the exocyclic amine of base G3 makes a hydrogen bond with the
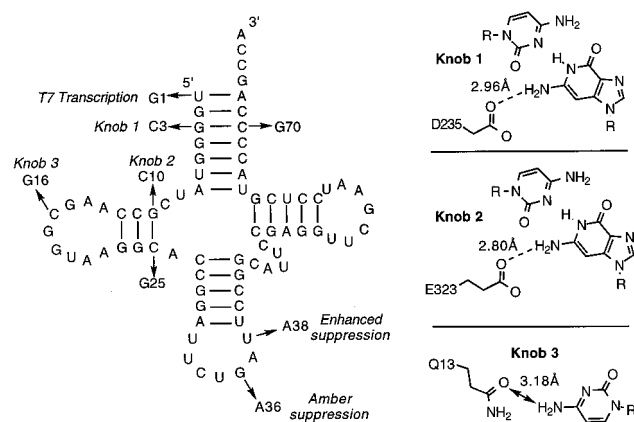
FIG. 1. Sequence of tRNA$_2^{Gln}$ and the base changes converting tRNA$_2^{Gln}$ to the orthogonal-tRNA (O-tRNA). The interactions between knob 1, knob 2, and knob 3 in the tRNA with residues Asp-235, Glu-323, and Gln-13 in GlnRS are also depicted (19).

carboxylate of Asp-235 (knob 1), the exocyclic amine of base G10 makes a hydrogen bond with the carboxylate of Glu-323 (knob 2), and the exocyclic amine of base C16 contacts the carboxamide side chain of Gln-13 (knob 3).

In an effort to eliminate recognition of tRNA$_2^{Gln}$ by GlnRS, while preserving translational competence, these three sites were mutated in the tRNA. Indeed, previous mutations made at these three sites in the tRNA resulted in 2- to 400-fold decreases in k_{cat}/K_m of GlnRS for the mutant tRNAs (21–23). In the case of knob 1, the base pair G3–C70 was changed to C3–G70; in the case of knob 2, the base pair G10–C25 was changed to C10–G25; and at knob 3, the pyrimidine C16 was changed to a purine (G16). In addition, the tRNA anticodon was changed from CUG to CUA to allow amber suppression, the first base of the tRNA was changed from U1 to G1 to allow efficient T7 RNA polymerase transcription and to provide the possibility of further discrimination by the wt GlnRS, and base U38 was mutated to A38 to enhance suppression efficiency (24).

Mutant suppressor tRNA genes containing all possible combinations of knobs 1, 2, and 3 were generated by runoff transcription. The ability of these mutant tRNAs to be recognized by endogenous *E. coli* tRNA-aminoacyl synthetases was initially assayed using *in vitro* transcription/translation reactions containing DNA bearing an amber codon at position 88 of *E. coli* chorismate mutase (25) and one of the mutant tRNAs. Since these *in vitro* suppression reactions are driven by *E. coli* cell extract containing all soluble cellular proteins including tRNA-aminoacyl synthetases, the amount of full-length chorismate mutase generated in these reactions reflects the ability of the suppressor tRNA to be recognized by any endogenous aminoacyl synthetase. The two tRNAs bearing mutations at knobs 2 and 3, or at knobs 1, 2, and 3, were unable to suppress the nonsense mutation at position 88: suppression efficiencies were 6% for both tRNAs, compared with 7% readthrough product for the reaction lacking tRNA (40).

To verify that the absence of full-length protein was not a result of the failure of the mutant tRNA to be accepted by elongation factor-Tu or the ribosome, the four engineered tRNAs bearing combinations of two or three of the knob mutations were generated by runoff transcription as the truncated tRNAs lacking the 3′ CA dinucleotide. These four truncated tRNAs were then enzymatically ligated to the chemically aminoacylated dinucleotide pdCpA-valine to afford full-length suppressor tRNAs acylated with valine (9). *In vitro* suppression assays using these acylated tRNAs revealed that the valine-acylated tRNA bearing mutations at knobs 2 and 3 again did not suppress the nonsense mutation (unpublished work) (40), indicating a lack of acceptance by elongation factor-Tu or the ribosome and thus a lack of suitably as an orthogonal tRNA. Surprisingly, the tRNA bearing

all three knob mutations suppressed the amber codon in chorismate mutase with 39% efficiency, indicating that the three sets of mutations interact in a complex and nonadditive fashion (40). Equal or greater amounts of full-length protein were produced with this tRNA than with the efficient wt *supE* suppressor tRNA$_2^{Gln}$(A36). This tRNA therefore satisfies the main requirements of an O-tRNA—the inability to serve as a substrate for an endogenous aminoacyl-tRNA synthetase and the ability to function in translation.

Two additional experiments confirmed the suitability of the tRNA bearing all three knob mutations to serve as the orthogonal tRNA. *E. coli* strain BT235 (4) harbors genomic amber mutations at a critical glutamine codon in β-galactosidase and in a cysteine biosynthesis gene, rendering it unable to grow on lactose minimal media or without cysteine supplementation. When doubly transformed with a plasmid expressing wt GlnRS and a plasmid expressing the *supE* suppressor, BT235 survived on lactose minimal media supplemented with cysteine at a rate approaching 100%. When doubly transformed with a plasmid (pBRglnS^{wt}) expressing wt GlnRS and a plasmid (PAC123) expressing the suppressor tRNA bearing mutations at knobs 1, 2, and 3, however, this strain survived on lactose minimal media at a rate of only 1 in 100,000. This tRNA was also assayed *in vitro* as a substrate for wt *E. coli* GlnRS using [³H]glutamine. At physiological concentrations of ATP (3.0 mM), tRNA (2.0 μM), and glutamine (150 μM), purified wt *E. coli* GlnRS acylated the knob 1/2/3 tRNA 13,000-fold slower than wt tRNA^{Gln} [the kinetic properties of tRNA$_1^{Gln}$ acylation and tRNA$_2^{Gln}$ acylation are similar (26)] (Table 1). Taken together, both the *in vitro* and *in vivo* experiments suggest that the suppressor tRNA$_2^{Gln}$(G1C3C10G16G25A36A38G70) is a promising candidate as the O-tRNA for the delivery of an unnatural amino acid *in vivo*.

Evolution of a Synthetase That Aminoacylates the O-tRNA. The second step in our strategy involves the generation of an aminoacyl-tRNA synthetase that charges the orthogonal tRNA with a natural amino acid. We used multiple rounds of "DNA shuffling" (27) (random point mutagenesis and *in vitro* homologous recombination) together with oligonucleotide-directed mutagenesis to generate libraries of mutant GlnRS enzymes bearing a low level of mutations throughout the protein as well as a high mutation rate at targeted residues. These mutant enzymes were selected for their ability to acylate the O-tRNA *in vivo* using strain BT235. Suppression of the genomic amber codon in *lacZ* with a glutamine-inserting suppressor tRNA restores the ability of BT235 cells to metabolize lactose. BT235 doubly transformed with a mutant GlnRS library in pBR322 and with a pACYC184 derivative (pAC123) expressing the O-tRNA will therefore grow on lactose only if a mutant GlnRS enzyme charges the orthogonal tRNA with glutamine.

The x-ray crystal structure of the GlnRS-tRNA$_2^{Gln}$ complex was used to direct mutagenesis to key sites in GlnRS during the shuffling process. In an effort to increase recognition of the O-tRNA, synthetic oligonucleotides containing three randomized bases at positions encoding knob 1 (Asp-235), knob 2 (Glu-323), knob 3 (Gln-13), and Arg-402 (Arg-402 contacts G36 of the wt anticodon, which corresponds to A36 in the O-tRNA) were added to the reassembly reaction in the fourth round (see Fig. 3). Mutations in Asp-235, Glu-323, and Arg-402 have been previously found to result in GlnRS enzymes with altered aminoacylation activities for wt and noncognate tRNAs (28, 29). To decrease recognition of the wt tRNA^{Gln}, oligonucleotides bearing the specific mutations Leu-136-Phe, Asp-235-Asn, and Arg-402-Ala were introduced into one library in the fifth round, and the Leu-136-Phe mutation was further introduced into one library in the seventh round (Fig. 2). Mutant GlnRS enzymes with these substitutions have previously been shown to exhibit decreased wt tRNA^{Gln} recognition while allowing the recognition of various noncognate amber suppressor tRNAs (5, 28).

Table 1. *In vitro* aminoacylation activities of mutant GlnRS enzymes

| Clone | Activity* | | Activity change[†] | | Discrimination ratio[‡] |
	wt-tRNA ($\times 10^3$)	O-tRNA	wt-tRNA	O-tRNA	
wt	230 ± 32	17 ± 2.1	1	1	13,000 ± 2,500
2/1	140 ± 18	9.5 ± 1.8	1.7 ± 0.32	0.57 ± 0.13	14,000 ± 3,300
3/1	250 ± 41	15 ± 1.6	0.92 ± 0.20	0.89 ± 0.14	17,000 ± 3,200
4/1	97 ± 13	80 ± 0.80	2.3 ± 0.46	4.8 ± 0.59	1,200 ± 160
4/2	99 ± 5.1	280 ± 8.5	2.3 ± 0.34	17 ± 2.1	350 ± 21
5/1	84 ± 13	500 ± 81	2.7 ± 0.56	30 ± 6.0	170 ± 38
5/2	120 ± 16	1,500 ± 290	1.9 ± 0.37	91 ± 20	78 ± 18
5/3	40 ± 2.3	3,000 ± 220	5.7 ± 0.87	180 ± 26	13 ± 1.2
5s/1	75 ± 4.8	180 ± 17	3.0 ± 0.47	11 ± 1.6	420 ± 48
5s/2	42 ± 9.6	1,400 ± 220	5.4 ± 1.4	84 ± 17	30 ± 8.3
6/1	53 ± 1.4	230 ± 30	4.3 ± 0.61	13 ± 2.4	240 ± 32
6/2	32 ± 0.84	180 ± 8.2	7.1 ± 1.0	11 ± 1.4	170 ± 9.1
6/3	11 ± 2.8	140 ± 32	20 ± 5.9	8.6 ± 2.2	77 ± 26
6/4	24 ± 0.31	870 ± 160	9.3 ± 1.3	52 ± 11	28 ± 5.2
7d/1	14 ± 0.96	450 ± 26	16 ± 2.5	27 ± 3.6	31 ± 2.8
7d/2	7 ± 0.17	290 ± 41	32 ± 4.6	17 ± 3.2	24 ± 3.6
7/1[§]	61 ± 7.0	120 ± 22	3.7 ± 0.67	7.3 ± 1.6	500 ± 110
7/2	61 ± 2.3	1,800 ± 86	3.7 ± 0.54	110 ± 14	34 ± 2.1
7/3	8.6 ± 0.73	300 ± 14	26 ± 4.3	18 ± 2.3	29 ± 2.8
7/4	19 ± 2.4	810 ± 120	12 ± 2.2	48 ± 9.2	24 ± 4.6
7/5	39 ± 3.2	1,900 ± 64	5.7 ± 0.93	120 ± 15	20 ± 1.8
7/6	8.1 ± 0.094	920 ± 42	28 ± 3.9	55 ± 7.1	8.8 ± 0.42

*All assays were carried out two to six times. Activities are listed in picomoles of acylated tRNA produced per milligram of enzyme per minute.
[†]Changes in activity are given as ratios corresponding to decreases in wt tRNA^{Gln} acylation and increases in O-tRNA acylation.
[‡]Discrimination ratios are defined as the wt tRNA^{Gln} acylation activity divided by O-tRNA acylation activity.
[§]The GlnRS from clone 7/1 had a 5- to 10-fold higher expression level than the other Round 7 enzymes.

Efforts to increase the stringency of the selection in the fifth and sixth rounds by decreasing cysteine supplementation and thereby forcing suppression of both the *lacZ* and *cys* amber mutants failed, presumably due to the inability of the mutant GlnRS enzymes or the O-tRNA to provide high enough suppression efficiencies to complement both defects (4). All colonies surviving each round of selection were carried forward for the first six rounds to allow evolution along diverse pathways. The five mutants with the lowest ratios of wt tRNA to O-tRNA recognition after the first six rounds were used to generate a new library (round seven) in an effort to increase the performance of each member of the library (Fig. 2).

***In Vitro* Characterization of Mutant Synthetases.** A purification of GlnRS mutants free of endogenous wt GlnRS has recently been developed using a strain of *E. coli* (HAPPY101), in which a hemagglutinin epitope has been attached to the genomic copy of the glutaminyl-tRNA aminoacyl synthetase
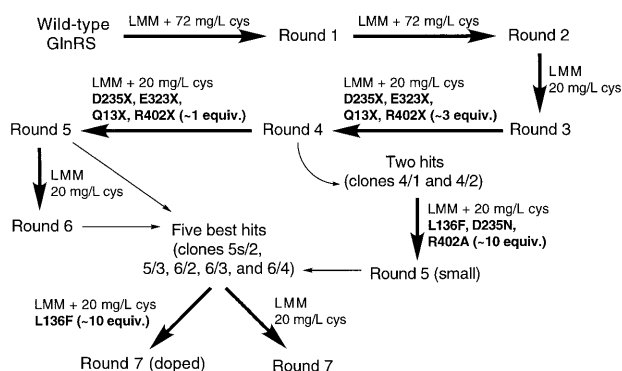

FIG. 2. Flow chart describing the history of each GlnRS library. Thick arrows denote a round of DNA shuffling and selection on the media indicated next to each arrow. LMM, lactose minimal media. Thin arrows indicate that a small number of characterized GlnRS genes were used to begin a new round of DNA shuffling and selection. Mutagenic oligonucleotides, when added to a reassembly reaction during DNA shuffling, are listed in boldface type.

gene (5). Using this method, several mutant synthetases surviving each round of selection were purified and assayed *in vitro* for their ability to acylate wt tRNA^{Gln} (as a mixture of all *E. coli* tRNAs including tRNA$_1^{Gln}$ and tRNA$_2^{Gln}$) and the O-tRNA. With the caveat that only a few mutant enzymes were purified and assayed after each round of selection, the ratio of wt tRNA^{Gln} acylation to O-tRNA acylation decreased significantly upon multiple rounds of selection. Whereas wt GlnRS acylates the wt tRNA^{Gln} 13,000 times as efficiently as the O-tRNA, the best mutant assayed after seven rounds of DNA shuffling and selection acylated the O-tRNA at one-ninth the rate of wt tRNA^{Gln} (Table 1). This 1,400-fold change in discrimination is dominated by increased recognition of the O-tRNA (as much as 180-fold) accompanied by a significant decrease in recognition of the tRNA^{Gln} (as much as 32-fold) (Table 1). Twelve mutant *glnS* genes were fully sequenced and contained a total of 138 silent and 177 nonsilent mutations in the translated sequence; the only frameshift mutations observed occurred near the 3′ end of clones 6/2 and 7/4 and resulted in the loss of the terminal three residues in both cases.

In some libraries, all of the "parent" GlnRS enzymes entering the shuffling process and some of the "daughter" enzymes surviving selection were characterized. As a representative example of the mutagenesis, recombination, and selection process, two mutant *glnS* genes from clones 4/1 and 4/2 were shuffled together with 10 equivalents each of the Leu-136-Phe, Asp-235-Asn, and Arg-402-Ala mutagenic oligonucleotides. Each parental *glnS* gene contained 10 different nonsilent mutations resulting in wt tRNA^{Gln} to O-tRNA discrimination factors of 1,200 and 350 (Table 1). One daughter GlnRS (clone 5s/2) that emerged from the shuffling and oligonucleotide-directed mutagenesis of these two genes inherited three mutations from one parent, two mutations from the other parent, acquired two mutations encoded by the synthetic oligonucleotides, and developed five new nonsilent mutations. This daughter GlnRS possessed a discrimination factor of 30, which resulted from an 5.4-fold decrease in wt tRNA^{Gln} recognition and an 84-fold increase in O-tRNA recognition (Table 1).

The substantial decrease in the ability of many mutant GlnRS enzymes to acylate wt tRNA^{Gln} is especially noteworthy given that no negative selection against wt tRNA recognition was employed thus far. This result implies that some mutations beneficial to O-tRNA recognition may eliminate favorable interactions with the wt tRNA and further suggests that this selection strategy may eventually lead to mutant synthetases with reversed specificities toward the wt- and O-tRNAs. Because even low levels of misacylation of tRNAs other than tRNA$_1^{Gln}$ and tRNA$_2^{Gln}$ should result in a lethal phenotype due to widespread missense mutations throughout the cell (30), it is unlikely that any of the mutant synthetases surviving selection acylate other endogenous tRNAs.

***In Vivo* Suppression to Generate a Cell Surface Protein.** To further demonstrate the ability of the O-tRNA and mutant GlnRS to serve as a translationally active system *in vivo*, as well as establish a screen for the incorporation of a keto amino acid by an orthogonal aminoacyl-tRNA synthetase, two amber mutants of the gene encoding the *E. coli* maltoporin LamB were constructed by PCR mutagenesis. The sites of these amber mutations, positions Asn-72 and Asp-200, both lie within loops of LamB, which are known from an x-ray crystal structure (31) and from biochemical studies (32) to exist in the extracellular space. In addition, insertion of the C3 epitope from poliovirus at these sites results in the display of this epitope on the surface of the cell in a manner accessible to anti-C3 polyclonal and monoclonal antibodies (33). Consequently, unnatural amino acids inserted into these sites should be detectable based on their unique reactivity (in the case of a keto amino acid) or using an antibody specific for a peptide containing the desired unnatural amino acid (34).

The LamB nonsense mutants, under control of the *araC* repressor and promoter, were inserted into the pAC123 plasmid containing the O-tRNA gene to afford pAC123LamB(N72Am) and pAC123LamB(D200Am). The pAC123LamB plasmids along with a representative set of three mutant GlnRS genes contained in pBR322 were transformed into *E. coli* strain GS20, which has a genomic deletion of *lamB*. Arabinose induction of mutant LamB expression led to suppression of the amber codon in LamB with the O-tRNA acylated by the mutant GlnRS. A Western blot using rabbit anti-LamB serum showed that efficient suppression at sites 72 and 200 results in full-length LamB production (Fig. 3, lanes 8, 10, 12, 14, 16, 18). The absence of the
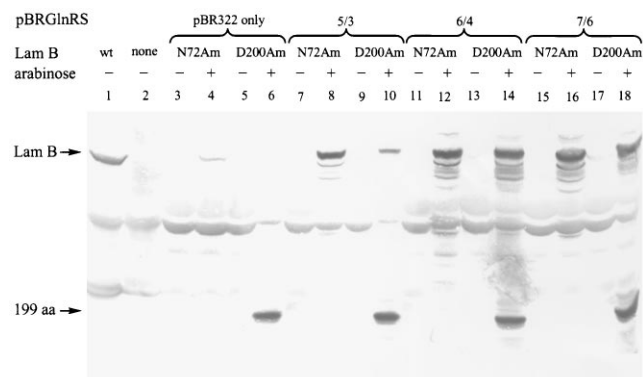

FIG. 3. *In vivo* suppression of *lamB* amber mutants by the O-tRNA and mutant GlnRS enzymes. Crude cell lysates were subjected to SDS/PAGE and the resulting proteins were visualized by Western blot using anti-LamB polyclonal antibodies. Lanes: 1, XL-1 blue cells (a *lamB* positive strain); 2, GS-20 cells (a strain lacking the *lamB* gene); 3–18, GS-20 cells transformed with pAC123LamB harboring either the N72Am or the D200Am *lamB* mutant gene as indicated; 3–6, pBR322 transformants; 7–10, pBRGlnRS clone 5/3 transformants; 11–14, pBRGlnRS clone 6/4 transformants; 15–18, pBRGlnRS clone 7/6 transformants. Induction of mutant LamB expression with arabinose is indicated above each lane.

mutant GlnRS gene resulted in the production of little or no full-length LamB (Fig. 3, lanes 4 and 6).

Sequence Analysis of Mutant GlnRS Clones. Analysis of the DNA sequences encoding 12 mutant GlnRS enzymes, which recognize the O-tRNA to varying degrees, revealed 4 distinct classes of mutations. In 9 of the 12 cases examined, mutations in GlnRS occurred at one of the residues corresponding to knobs 1 (Asp-235), knob 2 (Glu-323), knob 3 (Gln-13), or Arg-402 contacting the anticodon base G36. Among the 12 sequences, position 235 encoded Asp, Asn, Thr, Ile, or Val; position 323 encoded Glu, Ala, or Leu; position 13 encoded Gln, Lys, Ser, or His; and position 402 encoded Arg, Ala, or Cys. With the exception of position 13, which is the only solvent-exposed amino acid among these four residues, most of the mutated residues at these positions have lost their ability to hydrogen bond with the tRNA. The increase in O-tRNA recognition resulting from these mutations, therefore, may arise from subtle and complex changes in side chain and tRNA packing arising from hydrogen bond removal, which are impossible to predict *a priori*. In the case of residue 13, the Gln to Lys mutation found in clones 4/1, 5s/2, 7d/1, 7d/2, and 7/6 may result in a new a hydrogen bond to the 5′ phosphate of tRNA base C15, whereas the Gln to His mutation found in clone 6/4 and 7/5 may result in a hydrogen bond with base G16 (knob 3) of the O-tRNA. Although site-directed mutagenesis and structural studies are needed to define the precise role of each mutation at these sites in determining tRNA specificity, the large number of studies confirming the importance of Asp-235, Glu-323, and Arg-402 suggest that the multiple mutations found at these positions contribute significantly to the large differences in substrate specificity between the wt and mutant GlnRS enzymes. Given that the Gln-13-Lys and Arg-402-Cys mutations each occurred in 60% of the best enzymes characterized from Round 7 despite the low frequency (20%) of these mutations in the parental DNA entering the seventh round, it is tempting to speculate that these two mutations may represent a consensus emerging in response to mutations at knob 1 and knob 3 in the O-tRNA.

The fact that a mutant GlnRS (clone 5/3) with the greatest degree of O-tRNA recognition, as well as one of the lowest ratios of wt tRNA^{Gln} to O-tRNA recognition, has no mutations at positions 13, 235, 323, or 402 indicates that functional recognition of a noncognate tRNA can also be achieved by residues distant from the knob sites. Indeed, a second group of mutations was mapped to residues other than Asp-235, Glu-323, Gln-13, or Arg-402, which in the GlnRS crystal structure lie close to the tRNA. These include Pro-126-Leu (clone 4/1), Ile-129-Val (clone 6/4), Arg-186-Gly (clone 7d/2), and Pro-188-Gln (clone 4/2), which occur at residues near bases 72, 73, or 74 in the tRNA; Phe-182-Cys (clones 6/2, 6/3, 7/4, and 7/6), Arg-238-Cys (clone 5s/2) and Asn-320-Ser (clones 4/1, 5s/2, 5/3, 7d/1, 7d/2, and 7/6), which are close to bases 3 (knob 1), 4 and 11, respectively, in the tRNA; and Asn-336-Gly (clone 6/2), His-368-Tyr (clones 6/4 and 7/4), Lys-401-Glu (clone 6/4), Asp-546-Asn (clone 5/3) and Asp-546-Gly (clone 6/3), which lie near the anticodon loop. The latter mutations may allow the mutant synthetases to better accommodate the A36 substitution in the O-tRNA, which is known to cause a 1,700-fold decrease in k_{cat}/K_m for the tRNA (22). Interestingly, the mutation at Ile-129 (clone 6/4) occurs at the same position as the mutation in the previously isolated *glnS15* allele encoding a GlnRS capable of aminoacylating the noncognate *supF* tRNA^{Tyr} amber suppressor (29).

Three mutations were found in residues that lie near or in the active site of GlnRS. While the 3′ end of the O-tRNA is identical with that of wt tRNA^{Gln}, a number of studies have demonstrated that mutations in either the tRNA or the GlnRS far from the active site often lead to large changes in the k_{cat} for aminoacylation or in the K_m for glutamine (28, 35–37). This "functional communication" between the active site and distal regions of GlnRS is consistent with the observation that tRNA

binding is a prerequisite for glutaminyl-adenylate formation (38). The mutations Ser-46-Phe (clone 4/2) and Ile-47-Val (clones 6/4, 7d/2, and 7/6) lie in close proximity to the bound glutaminyl-adenylate analog in the crystal structure (L. Silvian and T. Steitz, personal communication) and these mutations may compensate for mutations in the O-tRNA, which are communicated to the active site. The mutation of Lys-72, a residue located near the ribose of the 3′ terminus of the tRNA, to Glu-72 (clones 5/3, 7d/2, and 7/6) may play a similar role.

A fourth class of mutations map to the surface of GlnRS. Whereas many of these residues likely do not affect the specificity or activity of GlnRS, it is possible that subtle rearrangements arising from changes made at or near the enzyme's surface in some cases contribute to the ability of these mutants to acylate the O-tRNA. Indeed, many mutations mapping to the surface of GlnRS were found to occur in more than one mutant synthetase and may represent consensus sequences, which are responsible in part for altered substrate specificity. In summary, the evolution of GlnRS enzymes that recognize the O-tRNA involve both mutations at knobs 1, 2, 3, and the anticodon binding region, as well as mutations removed from these sites. The latter mutations reinforce previous observations (18) that mutations throughout proteins can play an important role in reconfiguring binding and catalytic sites.

Generating a Completely Orthogonal Synthetase/tRNA Pair. Any aminoacyl-tRNA synthetase that uses an unnatural amino acid must not acylate any wt tRNA, since even modest misacylation of a wt tRNA with a noncognate amino acid will likely result in a lethal phenotype (30). While the selections described above have afforded mutant GlnRS proteins with a 1,400-fold change in their discrimination toward the O-tRNA versus the wt tRNA^{Gln}, and while it is possible to modulate the degree of mutant GlnRS acylation of wt tRNA^{Gln} by controlling the expression levels of the O-tRNA and mutant GlnRS (39), a synthetase that does not acylate the wt tRNA^{Gln} would be ideal and has yet to be isolated. Although successive rounds of selection generally provided mutant GlnRS enzymes with higher ratios of O-tRNA to wt tRNA^{Gln} activities, this improvement clearly decreased between successive rounds (Table 1). This suggests that O-tRNA aminoacylation among our most active GlnRS mutants may no longer be limiting the growth of the cell. Efforts are currently underway to evolve a GlnRS that no longer recognizes the wt tRNA by continuing the current selection scheme under more stringent conditions (for example, by decreasing O-tRNA or GlnRS expression or by increasing wt tRNA expression) and by direct screen or selection against wt tRNA charging, using both natural and the keto amino acids. Alternatively, we are also replacing the genomic copies of *E. coli* GlnRS, tRNA$_1^{Gln}$, and tRNA$_2^{Gln}$ with their counterparts from yeast.

1. Mendel, D., Cornish, V. W. & Schultz, P. G. (1995) *Annu. Rev. Biophys. Biomol. Struct.* **24**, 435–462.
2. Cupples, C. G. & Miller, J. H. (1988) *Genetics* **120**, 637–644.
3. Cornish, V. W., Hahn, K. M. & Schultz, P. G. (1996) *J. Am. Chem. Soc.* **118**, 8150–8151.
4. Yamao, F., Inokuchi, H. & Ozeki, H. (1988) *Jpn. J. Genet.* **63**, 237–249.
5. Thomann, H.-U., Ibba, M., Hong, K.-W. & Söll, D. (1996) *Bio/Technology* **14**, 50–55.
6. Noren, C., Anthony-Cahill, S. J., Suich, D. J., Noren, K. A., Griffith, M. C. & Schultz, P. G. (1990) *Nucleic Acids Res.* **18**, 83–88.
7. Weygand-Durasevic, I., Schwob, E. & Söll, D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2010–2014.
8. Cload, S. T., Liu, D. R., Froland, W. A. & Schultz, P. G. (1996) *Chem. Biol.* **3**, 1033–1038.
9. Thorson, J. S., Cornish, V. W., Barrett, J. E., Cload, S. T., Yano, T. & Schultz, P. G. (1997) *Methods in Molecular Biology—Protein Synthesis: Methods and Protocols,* in press.
10. Crameri, A., Whitehorn, E., Tate, E. & Stemmer, W. (1996) *Nat. Biotechnol.* **14**, 315–319.
11. Neidhardt, F. & Umbarger, H. E. (1996) in *Escherichia coli and Salmonella*, ed. Neidhardt, F. (Am. Soc. Microbiol., Washington, DC), Vol. 1, pp. 13–16.
12. Raunio, R. & Rosenqvist, H. (1970) *Acta Chim. Scand.* **24**, 2737–2744.
13. Ikemura, T. (1985) *Mol. Biol. Evol.* **2**, 13–34.
14. Hoben, P. & Söll, D. (1985) *Methods Enzymol.* **113**, 55–59.
15. Lloyd, A. J., Thomann, H.-U., Ibba, M. & Söll, D. (1995) *Nucleic Acids Res.* **23**, 2886–2892.
16. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1995) *Current Protocols in Molecular Biology*, (Wiley, New York), Vol. 1, pp. 8.5.7–8.5.7.
17. Kwok, Y. & Wong, J. T. (1980) *Can. J. Biochem.* **58**, 213–218.
18. Patten, P. A., Gray, N. S., Yang, P. L., Marks, C. B., Wedemayer, G. J., Boniface, J. J., Stevens, R. C. & Schultz, P. G. (1996) *Science* **271**, 1086–1091.
19. Rould, M. A., Perona, J. J. & Steitz, T. A. (1991) *Nature (London)* **352**, 213–218.
20. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C. & Moras, D. (1991) *Science* **252**, 1682–1689.
21. Hayase, Y., Jahn, M., Rogers, M. J., Sylvers, L. A., Koizumi, M., Inoue, H., Ohtsuka, E. & Söll, D. (1992) *EMBO J.* **11**, 4159–4165.
22. Jahn, M., Rogers, M. J. & Söll, D. (1991) *Nature (London)* **352**, 258–260.
23. Englisch-Peters, S., Conley, J., Plumbridge, J., Leptak, C., Söll, D. & Rogers, M. J. (1991) *Biochimie* **73**, 1501–1508.
24. Bradley, D., Park, J. V. & Soll, L. (1981) *J. Bacteriol.* **145**, 704–712.
25. Stewart, J., Wilson, D. B. & Ganem, B. (1990) *J. Am. Chem. Soc.* **112**, 4582–4584.
26. Seno, T., Agris, P. F. & Söll, D. (1974) *Biochim. Biophys. Acta* **349**, 328–338.
27. Stemmer, W. P. C. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10747–10751.
28. Sherman, J. M., Thomann, H. & Söll, D. (1996) *J. Mol. Biol.* **256**, 818–828.
29. Perona, J. J., Swanson, R. N., Rould, M. A., Steitz, T. A. & Söll, D. (1989) *Science* **246**, 1152–1154.
30. Yarus, M. (1979) *Prog. Nucleic Acids Res. Mol. Biol.* **23**, 195–225.
31. Schrimer, T., Keller, T. A., Wang, Y.-F. & Rosenbusch, J. P. (1995) *Science* **267**, 512–514.
32. Francis, G., Brennan, L. & Ferenci, T. (1991) *Biochim. Biophys. Acta* **1067**, 89–96.
33. Newton, S. M. C., Klebba, P. E., Michel, V., Hofnung, M. & Charbit, A. (1996) *J. Bacteriol.* **178**, 3447–3456.
34. Kalo, M. S. (1996) *Biochemistry* **35**, 999–1009.
35. Rogers, M. J., Adachi, T., Inokuchi, H. & Söll, D. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 291–295.
36. Ibba, M., Hong, K.-W., Sherman, J. M., Sever, S. & Söll, D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 6953–6958.
37. Weygand-Durasevic, I., Rogers, M. J. & Söll, D. (1994) *J. Mol. Biol.* **240**, 111–118.
38. Schimmel, P. R. & Söll, D. (1979) *Annu. Rev. Biochem.* **48**, 601–648.
39. Swanson, R., Hoben, P., Summer-Smith, M., Uemura, H., Watson, L. & Söll, D. (1988) *Science* **242**, 1548–1551.
40. Liu, D. R., Magliery, T. J. & Schultz, P. G. (1997) *Chem. Biol.*, in press.