Chapter 4

A High-Throughput, Flow Cytometry-Based Method for the Identification of Novel Aminoacyl-tRNA Synthetase Activity

Abstract

The incorporation of non-canonical amino acids into recombinant proteins in *E. coli* is often facilitated by the introduction of new aminoacyl-tRNA (aaRS) activities to the expression host. We describe here a screening procedure for the identification of such new activities based on the cell surface display of non-canonical amino acids. Screening of a saturation mutagenesis library of the *E. coli* methionyl-tRNA synthetase (MetRS) led to the discovery of three MetRS mutants capable of incorporating the long-chain amino acid azidonorleucine into recombinant proteins with modest efficiency. The leucine 13 to glycine (L13G) mutation is found in each of the three MetRS mutants, and MetRS containing this single mutation is highly efficient in producing recombinant proteins containing azidonorleucine.

Introduction

The aminoacyl-tRNA synthetases (aaRS) are responsible for much of the fidelity of protein synthesis in cells by virtue of catalyzing the ligation of an amino acid to its cognate tRNA. Despite the specificity of these enzymes for their natural substrates, protein engineers have exploited the promiscuity of aaRS to incorporate a wide array of non-canonical amino acids into recombinant proteins in organisms such as *E. coli.*[1,2] Elevation of the aaRS activity of the expression host is often critical for efficient *in vivo* production of proteins in which one of the natural amino acids is globally replaced by a non-canonical amino acid.[3-6] Alternatively, mutant aaRS activity may be introduced into an expression host to permit the incorporation of non-canonical amino acids that are translationally silent regarding the wild-type protein synthesis machinery.[7-14] The discovery of new mutant aaRS activities for global reassignment of a natural amino acid to a non-canonical amino acid can be greatly accelerated by the efficient screening of random libraries of mutant aaRS. Such an approach has proved fruitful in generating novel aaRS activity for the related problem of site-specific incorporation of a non-canonical amino acid using a heterologous aaRS/suppressor tRNA pair.[15,16] Herein we describe a rapid flow cytometry-based screening protocol to examine libraries of mutant aaRS for their ability to incorporate reactive amino acids into proteins and demonstrate its application on the *E. coli* methionyl-tRNA synthetase (MetRS).

Our screening protocol (Figure 1) relies on our previous observations that the introduction of a non-canonical amino acid of interest into the *E. coli* cell surface protein OmpC (outer membrane protein C) and the subsequent covalent tagging of this amino acid provide a very sensitive mode of detection of the translational capability of the non-

canonical amino acid.[17,18] *E. coli* cells displaying recombinant OmpC expressed in media supplemented with azidohomoalanine (AHA, **1**, Figure 2) were covalently biotinylated via Cu-catalyzed azide-alkyne ligation[19] and subsequently stained with fluorescent avidin. These cells were readily differentiable from unlabeled cells in flow cytometric analyses. One potential difficulty in using Cu-catalyzed azide-alkyne ligation in our screening protocol is the toxicity of the copper catalyst toward *E. coli*, thus we investigated the use of biotin-PEO-cyclooctyne as an alternative way of covalently labeling azide-bearing cell surfaces.

As a first test of the screening protocol, we designed and built a saturation mutagenesis library of methionyl-tRNA synthetases (MetRS). High-resolution crystal structures of the *E. coli* MetRS are available both without[20] and with[21] methionine bound. Additional structures are available for MetRS bound with several analogs of methionine and methionyl-AMP.[22] This wealth of structural data allows for informed selection of residues for saturation mutagenesis in the binding pocket of MetRS. Furthermore, MetRS lacks the proofreading or editing activity found in related aaRS such as valyl-,[23] isoleucyl-,[24] and leucyl-tRNA synthetase,[25] so only engineering of the synthetic site of the synthetase need be considered. Screening of the MetRS library led to the discovery of three different mutants that are able to efficiently incorporate the long chain amino acid azidonorleucine (**2**) into recombinant proteins with modest protein yields. Lastly, we report that a single amino acid mutation, leucine 13 to glycine (L13G) that occurs in each of the three active mutants is sufficient to recapitulate the activity toward azidonorleucine. In fact, the L13G MetRS mutant is able to incorporate azidonorleucine into recombinant proteins more efficiently than any of the mutants obtained in the screen.

Materials and Methods

Azido Amino Acids and Tagging Reagents. Azidohomoalanine (**1,** AHA),[26] azidonorleucine (**2,** ANL),[18] biotin-PEO-propargylamide **4,**[17] biotin-PEO-cyclooctyne **5**,[27] and the tris(triazolyl)amine ligand **3**[19,28] were prepared as described previously. CuBr (99.999% purity) was purchased from Aldrich.

Methionyl-tRNA Synthetase (MetRS) Library Construction. Restriction enzymes were from New England Biolabs or Roche. T4 DNA ligase was purchased from Invitrogen or New England Biolabs. The plasmid pAJL-20 was used as a template for construction of the MetRS library. pAJL-20 has been previously described[18] and encodes a variant of outer membrane protein C (OmpC) under the control of the isopropyl-β-thiogalactoside (IPTG)-inducible T5 promoter. In addition, pAJL-20 contains a cassette encoding the *E. coli* methionyl-tRNA synthetase (MetRS) under the control of its natural promoter. Oligonucleotides encoding degenerate NNK (N = A,T,G,C; K = G,T) codons at the sites corresponding to leucine 13, proline 257, tyrosine 260, and histidine 301 in *E. coli* MetRS were obtained from Qiagen. Four separate PCR reactions were performed with PicoMaxx polymerase (Stratagene) using pAJL-20 as a template and the following pairs of primers: Nhe lib for and L13 rev, L13 for and P257 rev, P257 for and H301 rev, H301 for and lib rev. The P257 for and P257 rev primers span both the proline 257 and tyrosine 260 codons. Sequences for all primers may be found in the Supporting Information. The DNA fragments obtained from these PCR reactions were electrophoresed and purified using Zymo-spin column (Zymo Research Corp).

Equimolar quantities (ca. 10 fmol) of the fragments were mixed and subjected to ten rounds of PCR (95 ºC for 1 min, 55 ºC for 1 min, 72 ºC for 2 min) using PicoMaxx as the polymerase. The primers Nhe lib for and lib rev were subsequently added, and the reaction mixture was subjected to 30 more rounds of PCR using the same parameters described above. The resulting 1.4 kb PCR product was digested with *Not* I and *Bsr* GI restriction enzymes and ligated into pAJL-20 digested with the same enzymes. The ligation mixture was transformed into chemically-competent XL-1 Blue cells (Stratagene) yielding 10^6 independent transformants. The plasmid DNA from the pooled transformants was isolated via Maxiprep (Qiagen), diluted 10-fold with water, and used to transform the methionine auxotroph M15MA via electroporation, yielding 5 x 10^6 independent clones. Aliquots of cultures containing the pooled transformants (500 μL) were mixed with 500 μL frozen stock solution (65 % glycerol, 25 mM tris, 100 mM $MgSO_4$) and stored at -80 ºC until needed.

Other Plasmids. XL-1 Blue *E. coli* (Stratagene) was used for all recombinant DNA manipulations. The plasmid pAJL-61 is functionally equivalent to the previously described pQE-15 MRS[3] except that it also encodes a copy of the gene for the *lac* repressor protein *lacI*^{q}. pAJL-61 was constructed by digesting the 0.6 kb dihydrofolate reductase (DHFR) coding region from pQE-15 MRS with *Bam* HI and *Hind* III restriction enzymes. This fragment was ligated into pQE-80L (Qiagen) digested with the same enzymes to generate pAJL-60. The 2.5 kb cassette encoding MetRS was isolated from pQE-15 MRS upon digestion with *Nhe* I and was ligated into pAJL-60 linearized by *Nhe* I to generate pAJL-61. The plasmid pAJL-80 encodes the 551-aa monomeric variant

of MetRS bearing a C-terminal 6xHis-tag for affinity purification. A 1.7 kb DNA fragment was amplified from pAJL-20 using the primers Cterm MRS for and Cterm MRS rev (see supporting information for sequences). This fragment was digested with *Bsa* I and *Bgl* II and subsequently ligated to pQE-60 (Qiagen) digested with *Nco* I and *Bgl* II to generate pAJL-80. The plasmid pAJL-84 encodes the L13G mutant of MetRS and was generated using QuikChange site-directed mutagenesis with the primers L13G for and L13G rev (see Supporting Information). The integrity of all constructs was confirmed by DNA sequencing.

OmpC Expression and Cell Surface Labeling. The expression of recombinant OmpC containing either AHA or ANL was carried out essentially as previously described.[17] Briefly, 30 mL of M9 medium (M9 salts, 0.2 % glucose, 1 mM $MgSO_4$, 25 mg/L thiamine) supplemented with 40 mg/L of each of the twenty canonical amino acids, 200 mg/L ampicillin, and 35 mg/L kanamycin was inoculated with either 400 μL of an overnight culture of M15MA[pAJL-20] or with a 1 mL aliquot of M15MA transformed with the MetRS library. When the OD_{600} of these cultures reached 0.9 to 1.0, the cells were pelleted by centrifugation at 5000 x *g* for 5 minutes. The cells were resuspended in 30 mL of M9 medium supplemented with 19 amino acids (no methionine) and incubated with shaking for 10 min at 37 ºC. Following this incubation, the cells were pelleted again and resuspended in fresh M9 medium supplemented with 19 amino acids. The resulting cultures were divided into four 5 mL aliquots that were supplemented with the following: methionine (40 mg/L), AHA (40 mg/L), ANL (8 mM, 1.38 g/L), or no analog. Expression of recombinant OmpC was induced for 3 h at 37 ºC upon addition of IPTG to

a final concentration of 1 mM. Following OmpC expression, a 1 mL sample of each culture was pelleted and washed with 1 mL of sterile phosphate-buffered saline (PBS, pH 7.4). The cell surface of these cells was tagged either via Cu-catalyzed azide-alkyne ligation or by treatment of the cells with biotin-PEO-cyclooctyne. For the Cu-catalyzed reaction, the cells were treated with 50 μM biotin-PEO-propargylamide, 200 μM triazole ligand, and 100 μM CuBr, which was delivered as an aqueous suspension as previously described.[18] The Cu-catalzyed reaction proceeded for 16 h at 4 ºC with agitation. Alternatively, 1 mL of the washed cells was treated with 100 μM biotin-PEO-cyclooctyne for 16 h at 37 ºC with agitation. Following the labeling reaction, the cells were washed twice more with 1 mL of PBS and treated with 2.5 μL of a 1 mg/mL solution of an avidin-Alexa Fluor 488 conjugate (Molecular Probes) for 2 h at 4 ºC with agitation. The cells were washed three more times with 1 mL PBS to remove non-specifically bound avidin.

Flow Cytometry and Cell Sorting. All flow cytometric analyses were carried out on a DakoCytomation MoFlo cell sorter (DakoCytomation, Ft. Collins, CO) equipped with an argon ion laser emitting at 488 nm. All sorting was performed in sort single mode. Sort gates were set and data analysis was performed with Summit software (DakoCytomation). Highly fluorescent cells were sorted from libraries expressing OmpC containing ANL by setting a gate in the fluorescence channel corresponding to the top 1% of the cells. Additional gates were set in the forward scatter and side scatter channels to exclude events with unusually large size. 10^5 cells were collected in a typical experiment, and 2-3 x 10^3 of the sorted cells were reanalyzed to check the quality of the

sort. The remainder of the cells was rescued upon inoculation into 10 mL of 2xYT medium supplemented with 200 mg/L ampicillin and 35 mg/L kanamycin. The regrown pools of cells were mixed 1:1 with frozen stock buffer and stored at -80 ºC. Alternatively, the sorted cells were plated on 2xYT agar supplemented with the same antibiotics to facilitate analysis of individual clones.

Recombinant DHFR Expression, Purification, and Analysis. Dihydrofolate reductase (DHFR) containing azidonorleucine was produced from M15MA[pAJL-61] and its variants expressing selected MetRS mutants using the same culture conditions described above for the expression of OmpC. Supplementation of ANL was decreased to 1 mM (172 mg/L) for production of DHFR using the L13G mutant of MetRS. Following protein expression, 10 mL of culture was pelleted and resuspended in 8M urea. DHFR was purified under denaturing conditions using Ni-NTA spin columns (Qiagen) as directed by the manufacturer. N-terminal sequencing was performed by the Peptide and Protein Microanalysis Laboratory at Caltech. Protein yields were determined by measuring the A_{280} of a solution of the purified protein using an extinction coefficient of 32693 M^{-1} cm^{-1}. For MALDI-MS analysis, purified DHFR in 8 M urea was diluted 10-fold in 75 mM NH_4HCO_3 and porcine trypsin (Promega) was added to a final concentration of 4 ng/μL. Following two hours of digestion at 37 ºC, the resulting peptide mixtures were desalted using either C_{18} ZipTips (Millipore) or PepClean C_{18} columns (Pierce). The peptide mixtures were dissolved in a saturated solution of the matrix α-cyano hydroxycinnamic acid (Fluka) in 1:1 water:acetonitrile and 0.1 % trifluoroacetic acid, and subjected to MALDI-MS analysis.

MetRS Expression, Purification, and Activation Assays. 2 L of Superbroth medium were inoculated with 10 mL of an overnight culture of XL-1 Blue[pAJL-84]. When the culture reached an OD_{600} of 1.0, IPTG was added to a final concentration of 1 mM, and protein expression was induced at 37 ºC for 5 h. The cells were harvested by centrifugation (6000 x*g* for 15 minutes). Histidine-tagged MetRS was isolated under native conditions using Ni-NTA agarose (Qiagen) according to the manufacturer's recommendations. The column eluent was buffer-exchanged into storage buffer (50 mM Tris, 1 mM DTT) using Bio-Rad PD-10 columns. The eluent was added to an equal mass of glycerol, mixed thoroughly, and frozen at -80 ºC until needed. Activation assays were carried out as described previously.[29] The methionine concentrations tested ranged from 150 μM to 5 mM while the azidonorleucine concentrations ranged from 625 μM to 20 mM. The L13G mutant of MetRS was added to all reactions at a concentration of 100 nM. Data were fit to the Michaelis-Menten model using SigmaPlot (SPSS) software.

Results and Discussion

Cell Surface Labeling with Biotin-PEO-Cyclooctyne. We have previously labeled azide-functionalized *E. coli* cell surfaces via copper-catalyzed azide-alkyne ligation using the triazole ligand **3** and biotin-PEO-propargylamide **4** as outlined in Figure 2. Because of the toxicity of the requisite copper catalyst toward *E. coli*, we also investigated the efficiency of cell surface labeling with biotin-PEO-cyclooctyne **5**. Cells displaying OmpC containing AHA (OmpC-AHA) were generated and were covalently biotinylated by treatment with either 100 μM CuBr, 200 μM **3**, and 50 μM **4** for 16 h at 4 ºC or with

100 μM **5** for 16 h at 37 ºC. Following staining with fluorescent avidin, the cells were subjected to flow cytometry. The median fluorescence of cells tagged with **5** is roughly 1/3 of the median fluorescence of cells tagged via the copper-catalyzed reaction (Figure 3). The median fluorescence of cells tagged with **5** still represents a 25-fold increase over the background fluorescence of *E. coli*. While cells subjected to Cu-catalyzed azide-alkyne ligation are unable to divide, cells that have been tagged with **5** and subjected to cell sorting are readily regrown in selective rich media.

MetRS Library Design and Construction. In order to select residues in the binding pocket of MetRS for saturation mutagenesis, the crystal structure of methionine-bound MetRS[21] was examined using Swiss PDB Viewer software. Five residues were found within a radius of 4 Å of either the sulfur atom or the methyl group of bound methionine: leucine 13, tryptophan 253, proline 257, tyrosine 260, and histidine 301. A multiple alignment of MetRS sequences of different organisms (see Figure S1) reveals that W253, Y260, and H301 are universally conserved while L13 and P257 are highly conserved. These residues have been implicated as being critical for methionine binding: the backbone amide of L13 and the hydroxyl group of Y260 form hydrogen bonds with the sulfur atom of bound methionine.[21] H301 is also believed to interact with the sulfur atom of methionine while W253 forms hydrophobic contacts with the methionine side chain.[22] Four of these five residues, excluding W253, were selected for mutagenesis to all other possible amino acids, and their orientation with respect to bound methionine is shown in Figure 4. The library was built via a modified PCR gene assembly process (see Materials and Methods) using primers containing degenerate NNK (N = all bases, K = G,T) codons

at the mutagenesis sites. The use of NNK codons permits complete coverage of the possible amino acid space while eliminating two of the nonsense codons and decreasing the theoretical size of the library for complete sequence coverage. The library was subsequently ligated into the plasmid pAJL-20, which encodes a variant of OmpC containing six additional surface-exposed methionine residues under inducible control, and the resulting plasmids were introduced to the methionine auxotroph M15MA.

Screening and Identification of Active Mutants. The MetRS library was screened for the ability to support efficient protein synthesis in minimal medium supplemented with the long chain amino acid azidonorleucine (ANL). We have previously reported very low levels of incorporation of this analog using elevated wild-type MetRS activity,[18] but the amounts of protein produced in these experiments is vanishingly small. We thus sought to identify a MetRS mutant from the library that would efficiently incorporate ANL into recombinant proteins. An aliquot of M15MA transformed with the MetRS library in pAJL-20 was used to inoculate minimal medium supplemented with all twenty canonical amino acids. Upon reaching mid-log phase, the cells were washed and transferred to minimal medium lacking methionine. This culture was supplemented with high levels (8 mM) of ANL, and OmpC expression was induced by the addition of IPTG. The cells were treated sequentially with biotin-PEO-cyclooctyne **5** and fluorescent avidin to render fluorescent the surface of those cells that had successfully incorporated ANL. The same cells were also biotinylated via Cu-catalyzed azide-alkyne ligation for comparison. Cells treated with **5** were analyzed by flow cytometry, and the top 1 % of the cells in the fluorescence channel was sorted and regrown in selective liquid medium

to ensure plasmid retention. A portion of the sorted cell population was also rescued by plating on selective agar plates, allowing analysis of single clones. The sorted pool of cells was subjected to another round of OmpC expression in ANL-supplemented media and biotinylated via azide-alkyne ligation to analyze the extent of enrichment of highly-fluorescent clones. While only 1 % of the cells in the original library resided in the high-fluorescence regime (>80 arbitrary fluorescence units), 40 % of the sorted cells fell in this range, demonstrating the ability to selectively enrich for these highly-fluorescent clones (Figure 5, panels A and B). Two individual clones from this population (clone 2.6.1 and clone 2.6.2) retained the high fluorescence characteristic of extensive azide functionalization of the cell surface (Figure 5, panels C and D) and thus were analyzed further. Additionally, the plasmid DNA from the pooled sorted cells was isolated and the MetRS genes were moved into a variant of the plasmid pAJL-20 encoding an OmpC with two additional methionine sites in its exposed loops as opposed to the six additional sites previously used. These plasmids were reintroduced to M15MA, and the screening protocol was carried out as described except that ANL was supplemented at 3.2 mM. This screen yielded one additional highly-fluorescent mutant, clone 3.2.7.

The mutations to MetRS found in each of these three clones are presented in Table 1. The net effect of these mutations is an expansion of the binding pocket to create room for the bulky substrate azidonorleucine. The character of the binding pocket remains mostly hydrophobic. Most striking is the recurrence of the L13G mutation, which is present in all three mutants analyzed. The L13G MetRS single mutant was thus constructed and tested for its ability to incorporate azidonorleucine as described below.

Recombinant Protein Production with MetRS Mutants. Each of the four MetRS mutants (clones 2.6.1, 2.6.2, 3.2.7, and L13G) was transferred to the plasmid pAJL-61, which encodes a 6xHis-tagged dihyrdofolate reductase (DHFR) under inducible control as well as the repressor protein *lacI*^{q}. The resulting plasmids were transformed into the expression host M15MA, which also contains a plasmid-borne copy of *lacI*^{q}. The high levels of *lacI*^{q} were introduced to minimize leaky expression of DHFR, since any DHFR produced before the induction period will not contain azidonorleucine. Mid-log phase cultures of M15MA[pAJL-61] were shifted to minimal medium lacking methionine but supplemented with 8 mM azidonorleucine as described in the Materials and Methods section. Following three hours of protein induction, the cells were lysed and the DHFR was purified using Ni-NTA chromatography. Replacement of methionine with azidonorleucine was readily observed in MALDI-MS analyses of the tryptic peptides derived from these samples of DHFR (see Figure S2). The yield of protein produced per liter of culture was calculated by measuring the A_{280} value of a solution of the purified protein and the extent of incorporation of azidonorleucine was estimated by N-terminal sequencing (Table 2). The L13G mutant is clearly the most efficient synthetase for the incorporation of azidonorleucine. SDS-PAGE analysis of whole cell lysates from M15MA[pAJL-20] cells harboring the L13G MetRS mutant demonstrates that comparable amounts of DHFR are produced when the cells are supplemented with either methionine or azidonorleucine (Figure 6). The yields of protein obtained from experiments using the L13G MetRS mutant are sufficient for demanding applications such as biomaterials synthesis.

***In Vitro* Activation Kinetics of L13G MetRS.** Previous work in our group indicates that the rate of activation of a non-canonical amino acid correlates strongly with its ability to be incorporated into recombinant proteins efficiently.[30,31] The kinetic parameters for activation of both methionine and azidonorleucine were determined using purified L13G MetRS (Table 3). Methionine is still activated by the mutant MetRS, albeit nearly 300-fold more sluggishly than the wild-type enzyme. The activation rate of azidonorleucine by the L13G mutant compares favorably with the rates previously determined for the activation of the excellent methionine surrogates azidohomoalanine and homopropargylglycine by the wild-type MetRS.[31]

Conclusions

We have presented a rapid, high-throughput screening methodology for the identification of new aminoacyl-tRNA synthetase activities *in vivo*. Using this screening protocol, we have identified three mutants of the *E. coli* methionyl-tRNA synthetase (MetRS) that can incorporate the long-chain amino acid azidonorleucine into recombinant proteins in response to methionine codons. The L13G mutation occurs in each of these three mutants, and MetRS harboring this single amino acid mutation proves to be the most efficient synthetase for the incorporation of azidonorleucine into recombinant proteins. The expanded binding pocket of L13G MetRS may permit the efficient incorporation of other long-chain amino acids into recombinant proteins, and we are currently screening the L13G MetRS against a panel of amino acids not incorporated by the wild-type translational machinery. The screening protocol described herein need not be limited to azide-bearing amino acids and may be readily adapted toward any

reactive side-chain moieties that can be covalently modified in aqueous conditions at physiological temperatures.

References

(1) Budisa, N. *Angew. Chem.-Int. Edit.* **2004**, *43*, 6426-6463.

(2) Link, A. J.; Mock, M. L.; Tirrell, D. A. *Current Opinion in Biotechnology* **2003**, *14*, 603-609.

3) Kiick, K. L.; van Hest, J. C. M.; Tirrell, D. A. *Angew. Chem.-Int. Edit.* **2000**, *39*, 2148-2152.

(4) Kiick, K. L.; Weberskirch, R.; Tirrell, D. A. *FEBS Lett.* **2001**, *502*, 25-30.

(5) Tang, Y.; Tirrell, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 11089-11090.

(6) Wang, P.; Fichera, A.; Kumar, K.; Tirrell, D. A. *Angew. Chem.-Int. Edit.* **2004**, *43*, 3664-3666.

(7) Ibba, M.; Hennecke, H. *FEBS Lett.* **1995**, *364*, 272-275.

(8) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. *FEBS Lett.* **2000**, *467*, 37-40.

(9) Bentin, T.; Hamzavi, R.; Salomonsson, J.; Roy, H.; Ibba, M.; Nielsen, P. E. *J. Biol. Chem.* **2004**, *279*, 19839-19845.

(10) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. *Chembiochem* **2002**, *3*, 235-237.

(11) Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 5652-5653.

(12) Kim, W. Y.; George, A.; Evans, M.; Conticello, V. P. *Chembiochem* **2004**, *5*, 928-936.

(13) Tang, Y.; Tirrell, D. A. *Biochemistry* **2002**, *41*, 10635-10645.

(14) Doring, V.; Mootz, H. D.; Nangle, L. A.; Hendrickson, T. L.; de Crecy-Lagard, V.; Schimmel, P.; Marliere, P. *Science* **2001**, *292*, 501-504.

(15) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498-500.

(16) Wang, L.; Schultz, P. G. *Angew. Chem.-Int. Edit.* **2005**, *44*, 34-66.

(17) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164-11165.

(18) Link, A. J.; Vink, M. K. S.; Tirrell, D. A. *J. Am. Chem. Soc.* **2004**, *126*, 10598-10602.

(19) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.

(20) Mechulam, Y.; Schmitt, E.; Maveyraud, L.; Zelwer, C.; Nureki, O.; Yokoyama, S.; Konno, M.; Blanquet, S. *J. Mol. Biol.* **1999**, *294*, 1287-1297.

(21) Serre, L.; Verdon, G.; Choinowski, T.; Hervouet, N.; Risler, J. L.; Zelwer, C. *J. Mol. Biol.* **2001**, *306*, 863-876.

(22) Crepin, T.; Schmitt, E.; Mechulam, Y.; Sampson, P. B.; Vaughan, M. D.; Honek, J. F.; Blanquet, S. *J. Mol. Biol.* **2003**, *332*, 59-72.

(23) Fersht, A. R.; Kaethner, M. M. *Biochemistry* **1976**, *15*, 3342-3346.

(24) Fersht, A. R. *Biochemistry* **1977**, *16*, 1025-1030.

(25) Englisch, S.; Englisch, U.; Vonderhaar, F.; Cramer, F. *Nucleic Acids Res.* **1986**, *14*, 7529-7539.

(26) Mangold, J. B.; Mischke, M. R.; Lavelle, J. M. *Mutation Research* **1989**, *216*, 27-33.

(27) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.

(28) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.* **2004**, *6*, 2853-2855.

(29) Hendrickson, T. L.; Nomanbhoy, T. K.; Schimmel, P. *Biochemistry* **2000**, *39*, 8180-8186.

(30) Kiick, K. L.; Tirrell, D. A. *Tetrahedron* **2000**, *56*, 9487-9493.

(31) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 19-24.

Supporting Information

Primer Sequences. Following are sequences of all primers described in this chapter. For degenerate nucleotides, N = A,T,G,C and K = G,T. Oligonucleotides were purchased from either Qiagen (Operon) or Integrated DNA Technologies.

Nhe lib for: 5'-GGTGAGAATCCAAGCTAGCTCTAGAGACGTCCG-3'

L13 for: 5'-GGTGACGTGCGCANNKCCGTACGCTAACGG-3'

L13 rev: 5'-CCGTTAGCGTACGGKNNTGCGCACGTCACC-3'

P257 for: 5'-CTACGTCTGGCTGGACGCANNKATTGGCNNKATGGGTTCTTTC-3'

P257 rev: 5'-GAAAGAACCCATKNNGCCAATKNNTGCGTCCAGCCAGACGTAG-3'

H301 for: 5'-TATTGTTTACTTCNNKAGCCTGTTCTGGCC-3'

H301 rev: 5'-GGCCAGAACAGGCTKNNGAAGTAAACAATA-3'

lib rev: 5'-CAGTGAAGGTTTTGTACAACTGCGGGTCAG-3'

Cterm MRS for: 5'-CTATTAAGAAGTAAGGTCTCCTATGACTCAAGTCGCG-3'

Cterm MRS rev: 5'-GGCAGCGGCTTTTACAGATCTTTTAGAGGCTTCC-3'

L13G for: 5'- CTGGTGACGTGCGCACTGCCGTACGCTAACGGC-3'

L13G rev: 5'- GCCGTTAGCGTACGGCAGTGCGCACGTCACCAG-3'

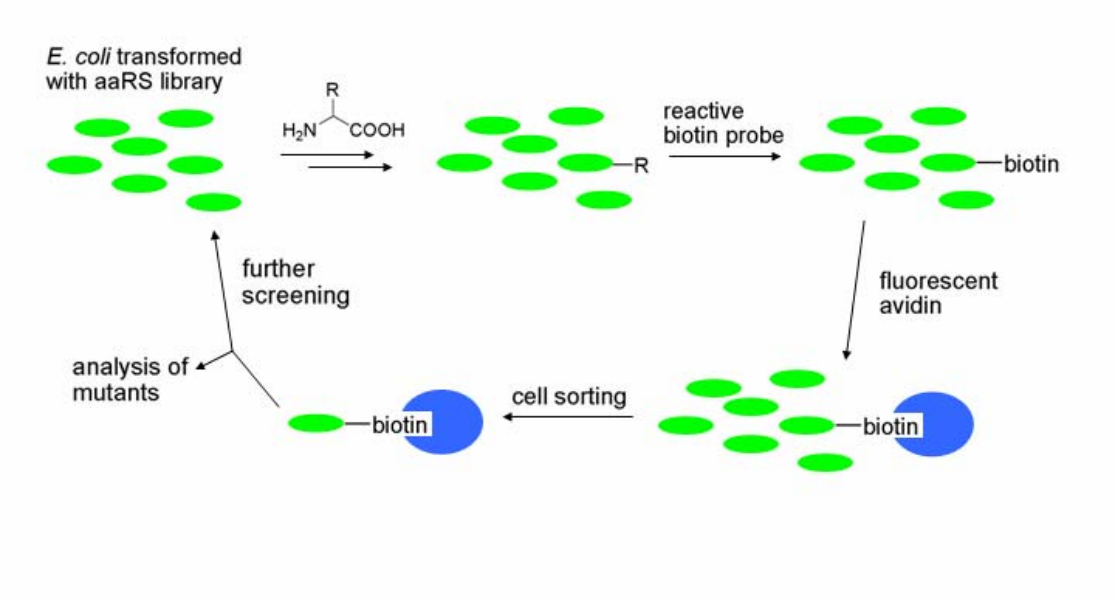

Figure 1: Protocol for screening libraries of mutant aminoacyl-tRNA synthetases (aaRS). Cells transformed with an aaRS library are induced to express outer membrane protein C (OmpC) in medium supplemented with a non-canonical bearing a reactive side-chain. Cells that successfully incorporate the non-canonical amino thus display the reactive side-chain on their surface. The reactive moiety is covalently captured by an appropriate biotin probe and subsequently stained with fluorescent avidin. Flow cytometric cell sorting isolates these tagged cells from the remainder of the library. Sorted cells may be subjected to additional rounds of screening or analyzed immediately.

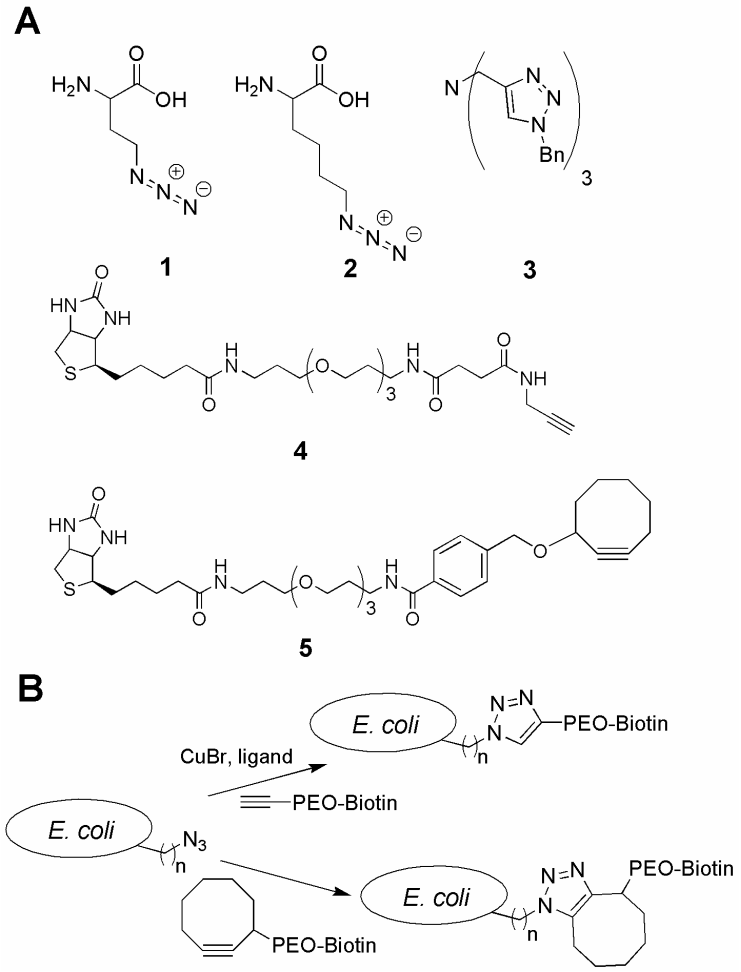

Figure 2 A: Non-canonical amino acids and tagging reagents used in this study. **1**: azidohomoalanine (AHA), **2**: azidonorleucine (ANL), **3**: tris(triazolyl)amine ligand for copper-catalyzed azide-alkyne ligation, **4**: biotin-PEO-propargylamide, **5**: biotin-PEO-cyclooctyne. B: Scheme for biotin tagging of azide functionalized *E. coli* cell surfaces. **1** or **2** is incorporated into outer membrane protein C (OmpC) displaying the azide functionality on the cell surface. Cell surface azides react either by Cu-catalyzed azide-alkyne ligation (top route) or by strain-promoted azide-alkyne ligation (bottom route).

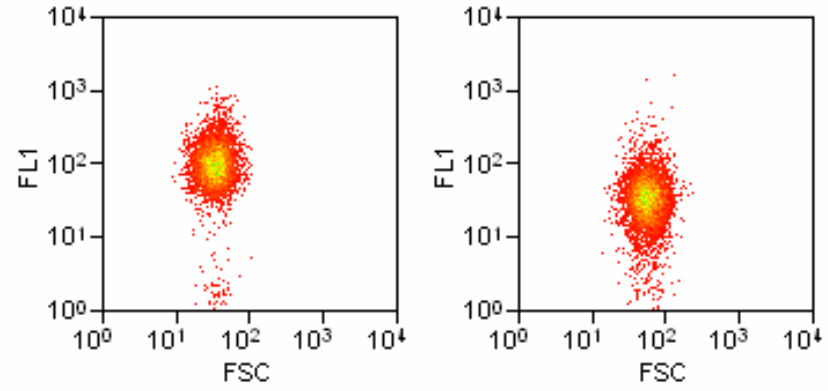

Figure 3: Comparison of the extent of cell surface labeling by Cu-catalyzed azide-alkyne ligation and strain-promoted azide-alkyne ligation. Cells displaying OmpC containing azidohomoalanine (**1**) were biotinylated either via Cu-catalyzed azide-alkyne ligation (50 μM **4**, 200 μM **3**, 100 μM CuBr) or by strain-promoted azide-alkyne ligation (100 μM **5**). The median fluorescence (FL1, y-axis) of cells labeled via Cu-catalyzed azide-alkyne ligation is 3-fold larger than the median fluorescence of cells labeled via the strain-based reaction.

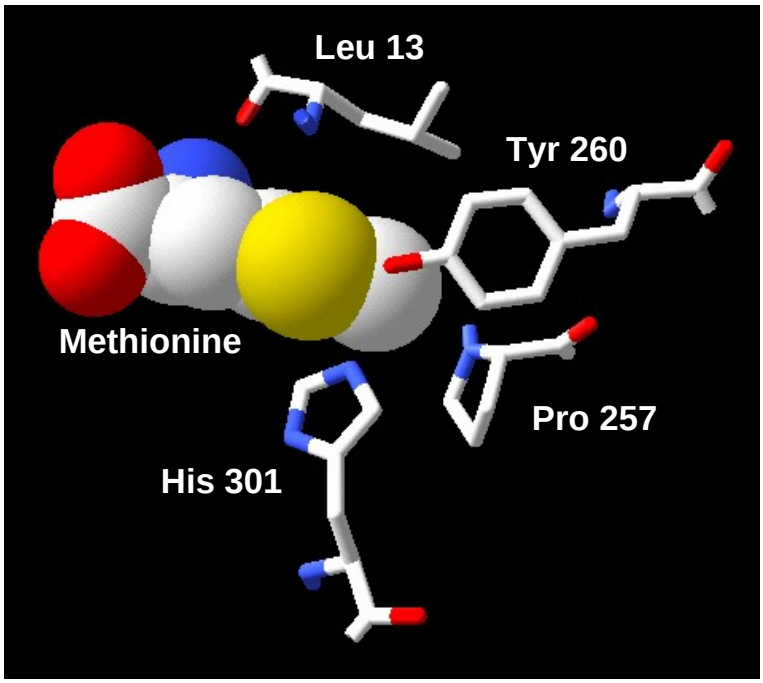

Figure 4: Residues in the methionine binding pocket of MetRS selected for saturation mutagenesis. Four residues (stick models) found within 4 Å of the sulfur atom or methyl group of methionine (space-filling model) were mutated to all other possible amino acids. Drawn from coordinates in ref. 21, PDB file 1F4L.

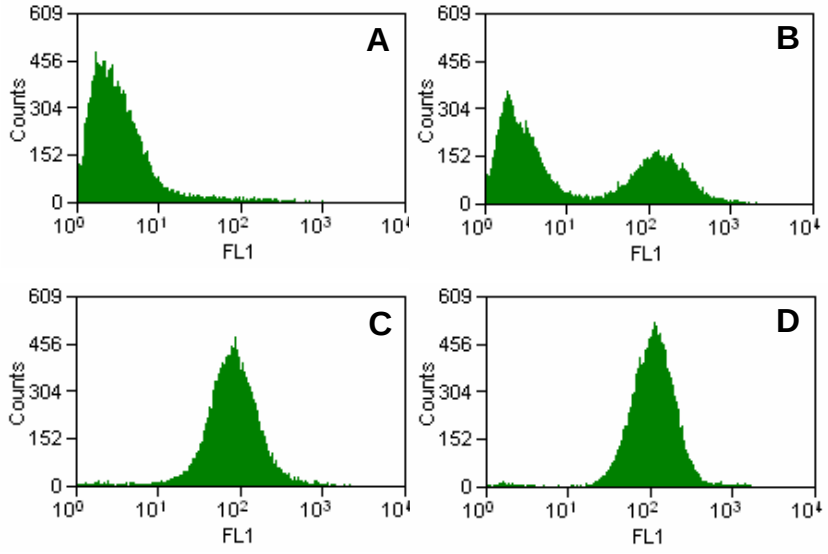

Figure 5: Fluorescence histograms of cells induced to produce OmpC in medium supplemented with azidonorleucine. All cells were biotinylated via Cu-catalyzed azide-alkyne ligation. A: cells harboring naïve MetRS library, B: cells enriched from the top 1 % of cells in A, C: cells harboring clone 2.6.1 MetRS, D: cells harboring clone 2.6.2 MetRS.

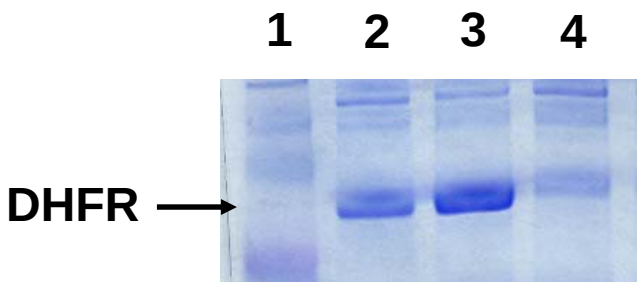

Figure 6: SDS-PAGE analysis of whole-cell lysates of M15MA[pAJL-61] encoding the L13G MetRS mutant. Lane 1: molecular weight marker. Cells were grown to mid-log phase in medium containing all 20 canonical amino acids, then shifted to 19 amino acid medium (no methionine) supplemented with 2: 40 mg/L methionine, 3: 172 mg/L azidonorleucine, 4: no amino acid. The expression levels of DHFR are comparable in the cultures supplemented with either methionine or azidonorleucine.

```
E. coli              --------------------MTQVAKKILVTCALPYANGSIHLGHM 26
S. typhimurium       --------------------MTQVAKKILVTCALPYANGSIHLGHM 26
V. cholerae          VLRVVLSLQFGLYQEKRILSMANDPRKLLVTCALPYANGSIHLGHM 74
S. odeinensis        --------------------MATSQRKILVTSALPYANGPIHLGHM 26
M. janaschii         -------------------------MRYLITTALAYTNGPLHLGHA 21
D. radiodurans       --------MQNPPQHPEAQSPETRDREFFITAAIDYANGTPHIGHV 38
M. tuberculosis      ------------------------MKPYYVTTAIAYPNAAPHVGHA 22
S. cerevisiae        -----------VKPKDSEILPKPNERNILITSALPYVNNVPHLGNI 216
H. sapiens           AWEKGLESLPPLRPQQNPVLPVAGERNVLITSALPYVNNVPHLGNI 284
C. elegans           ------------KKDPKSILPQPGKRNILITAALPYVNNVPHLGNI 55

E. coli              LQQWDISR--DAPYFGFEIPNA--PGKYFYVWLDAPIGYMGSFKNL 268
S. typhimurium       LQQWDISR--DAPYFGFEIPNA--PGKYFYVWLDAPIGYMGSFKNL 268
V. cholerae          LQQWDISR--DAPYFGFEIPGE--KDKFFYVWLDAPIGYMGSFKNL 316
S. odeinensis        LQQWDITR--DAPYFGFEIPDA--PGKYFYVWLDAPIGYMGSFKNL 268
M. janaschii         LHDWDISR--DIS-WGVPIPG---TNQVMYVWLEAPIGYISFTKML 261
D. radiodurans       IGPLSISRPKARVPWGIELPWD--TDHVTYVWFDALLSYLTPLVSQ 262
M. tuberculosis      LDDLSISR--TSFDWGVQVPEH--PDHVMYVWVDALTNYLTGAGFP 242
S. cerevisiae        LKPRCITR---DLVWGTPVPLEKYKDKVLYVWFDATIGYVSITSNY 464
H. sapiens           LKPRCITR---DLKWGTPVPLEGFEDKVFYVWFDATIGYLSITANY 532
C. elegans           LDPRCITR---DLKWGTAVPLDGFEKKVFYVWFDAPIGYLSITKCV 305


E. coli              KKDSTAELYHFIGKDIVYFHSLFWPAMLEGSN-FRKPTNLFVHG-Y 326
S. typhimurium       KKDSDAELYHFIGKDIVYFHSLFWPAMLEGSH-FRKPTNLFVHG-Y 326
V. cholerae          NKDSKTELYHFIGKDIVYFHSLFWPAMLDGSG-FRKPTNVFVHG-Y 373
S. odeinensis        AKDSKAEVYHFIGKDIVYFHSLFWPAMLYGSG-YRQPNSVYAHG-Y 325
M. janaschii         LEK-DTKIYHFIGKDITVHHAVFWPGMLIAHGSFNLPTAVVSGG-Y 313
D. radiodurans       ---------HVIGKDILKPHAVFWPTMLRAAG-LPLYRRLVVHSHI 309
M. tuberculosis      PAD-----LHMIGKDIIRFHAVYWPAFLMSAG-IELPRRIFAHG-F 292
S. cerevisiae        NNPEHVSLYQFMGKDNVPFHTVVFPGSQLGTEENWTMLHHLNTTEY 518
H. sapiens           KNPEQVDLYQFMAKDNVPFHSLVFPCSALGAEDNYTLVSHLIATEY 586
C. elegans           KNPENVELFNFVGKDNVAFHAVMFPCSQLGANDNYTVVNNLCATEY 360
```

Figure S1: Multiple sequence alignment of MetRS amino acid sequences from several different organisms. Methionine binding pocket residues within 4 Å of either the sulfur atom or methyl group of methionine are highly conserved and are highlighted in yellow.

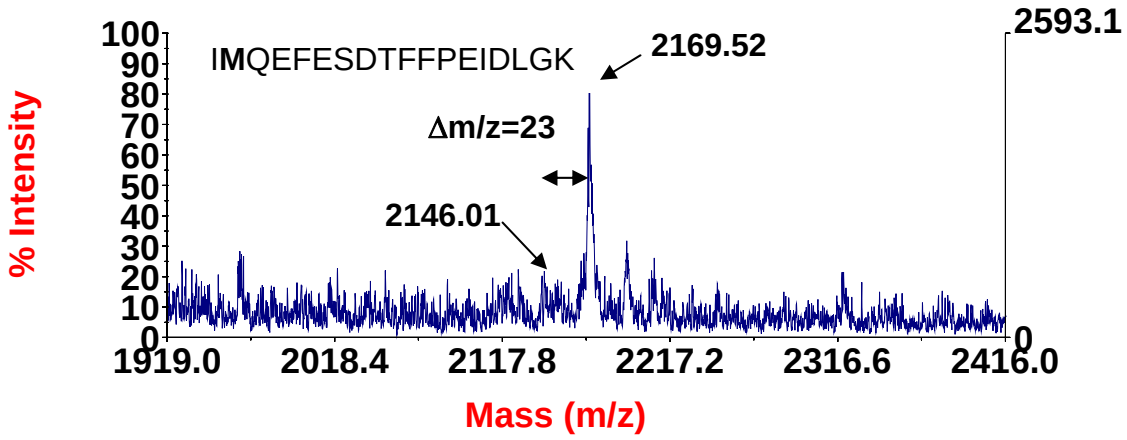

Figure S2: MALDI-MS spectrum of a tryptic peptide from DHFR containing azidonorleucine produced with L13G mutant MetRS. The predominant peak is mass-shifted 23 units from the predicted mass indicating nearly quantitative replacement of methionine with azidonorleucine.

MetRS	Leu 13	Pro 257	Tyr 260	His 301
2.6.1	**G**	**L**	**T**	**A**
2.6.2	**G**	**S**	**T**	**L**
3.2.7	**G**	**L**	**L**	**V**

Table 1: Mutations found in MetRS clones that efficiently incorporate azidonorleucine **2** into recombinant proteins.

MetRS	DHFR yield (mg/L culture)	Extent of incorporation
Wild-type	-*	-†
Clone 2.6.1	4.5	55 %
Clone 2.6.2	3.2	39 %
Clone 3.2.7	1.5	-†
L13G	18.1	95 %

Table 2: Purified protein yield and extent of incorporation of azidonorleucine in dihydrofolate reductase (DHFR) produced in M15MA[pAJL-61] harboring different MetRS variants. *: no protein was detected, †: not determined.

Amino acid	Synthetase	k_{cat}/K_m (μM^{-1} s^{-1})	Relative rate
Methionine	Wild-type	5.47×10^{-1}*	1
Azidohomoalanine (**1**)	Wild-type	1.42×10^{-3}*	1/390
Methionine	L13G	2.02×10^{-3}	1/270
Azidonorleucine (**2**)	L13G	1.56×10^{-3}	1/366

Table 3: Kinetic parameters for activation of methionine, **1**, and **2** by wild-type and L13G MetRS. *: from reference 31.