Supporting Information

Tanrikulu et al. 10.1073/pnas.0905735106

SI Text

Nomenclature for Populations and Clones Isolated Through Library Screens. The naïve saturation mutagenesis library was constructed using the plasmid pAJL-20, which allows the constitutive expression of wild-type MetRS, as a template. The set of plasmids obtained by the introduction of the library into pAJL-20 is designated pAJL-20-LYH.1.0. The population of clones produced by the transformation of pAJL-20-LYH.1.0 into the expression host M15MA[pREP4] is referred to as LYH.1.0.

Several screens were applied to LYH.1.0, giving rise to different cell populations. Names and descriptions of these populations are listed below. With the exception of LYH.6.2, all populations were obtained by screening LYH.1.0 in one or more rounds for clones exhibiting the highest fluorescence labeling.

LYH.1.0. Population carrying the naïve library (Fig. 3*A*)
LYH.1.1a. The naïve library screened at 1.0 mM Anl
LYH.1.2. LYH.1.1a screened at 1.0 mM Anl (Fig. 3*B*)
LYH.2.1b. The naïve library screened at 0.1 mM Anl
LYH.2.2. LYH.2.1b screened at 0.1 mM Anl (Fig. 3*D*)
LYH.2.3. LYH.2.2 screened at 0.1 mM Anl
LYH.3.1. The naïve library screened at 0.3 mM Anl
LYH.3.2. LYH.3.1 screened at 0.3 mM Anl (Fig. 3*C*)
LYH.3.3. LYH.3.2 screened at 0.3 mM Anl
LYH.4.3. LYH.3.2 screened at 0.1 mM Anl
LYH.5.3c. Mixture of LYH.1.1a, LYH.2.1b, and LYH.3.1 screened for two rounds at 1.0 mM Anl and 0.03 to 0.1 mM Met
LYH.5.3d. The naïve library screened at 1.0 mM Anl and 0.01 to 0.1 mM Met for three rounds (Fig. 3*E*)
LYH.5.4b. Mixture of LYH.1.1a, LYH.2.1b, and LYH.3.1 screened for three rounds at 0.3 mM Anl and 0.01 to 0.03 mM Met
LYH.6.2. Result of sorting a section of the population on the low-fluorescence arm of LYH.1.1a

Individual clones isolated from these populations were reanalyzed by flow cytometry to determine the extent of cell-surface labeling supported by the MetRS mutants they carry. The pAJL-20 variant isolated from each clone is named based on the MetRS mutant it codes for, such that the plasmid pAJL-20-GML bears the gene for the MetRS-GML mutant. Variants of pAJL-61 and pMTY21 used in this study were named similarly.

Among the MetRS variants isolated from the populations listed above, mutants that enable the incorporation of Anl at 1.0 mM concentration are presented and the populations that yielded each mutant are listed in Table S2. Mutants discussed in the main article are indicated in bold face.

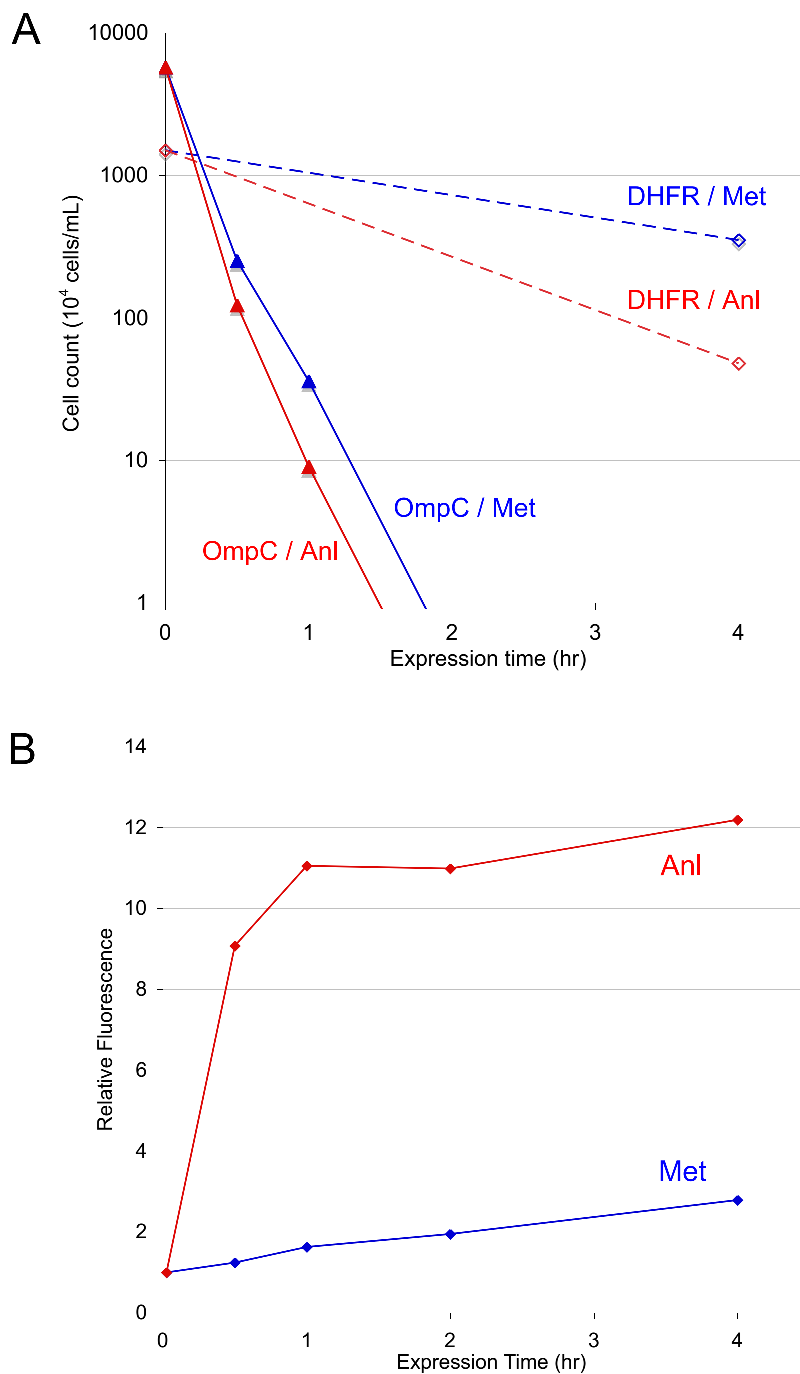

Fig. S1. Effect of the duration of OmpC expression on cell viability and cell-surface labeling. Aliquots were taken from cultures expressing either OmpC or mDHFR in media containing either Anl or Met at different times during protein expression. After cell-surface labeling, cells were plated on agar and the resulting colonies were counted to determine the number of viable cells. (*A*) Counts of viable colonies during expression. Data from OmpC- and DHFR- expressing cultures are marked with filled triangles (▲) and open diamonds (◇), respectively. Data from Anl-treated cultures are shown in red. No viable colonies could be observed for cultures that expressed OmpC for 2 h or more. Data were extrapolated beyond 1 h to emphasize this fact. Results suggest that the toxicity due to OmpC expression markedly decreases cell viability. (*B*) Total cell fluorescence at different times during OmpC expression. Aliquots taken during protein expression were fluorescently labeled and their total cell fluorescence was determined. Data are shown relative to the fluorescence read from uninduced cells. The fluorescence labeling is near complete after only 0.5 h of OmpC expression.

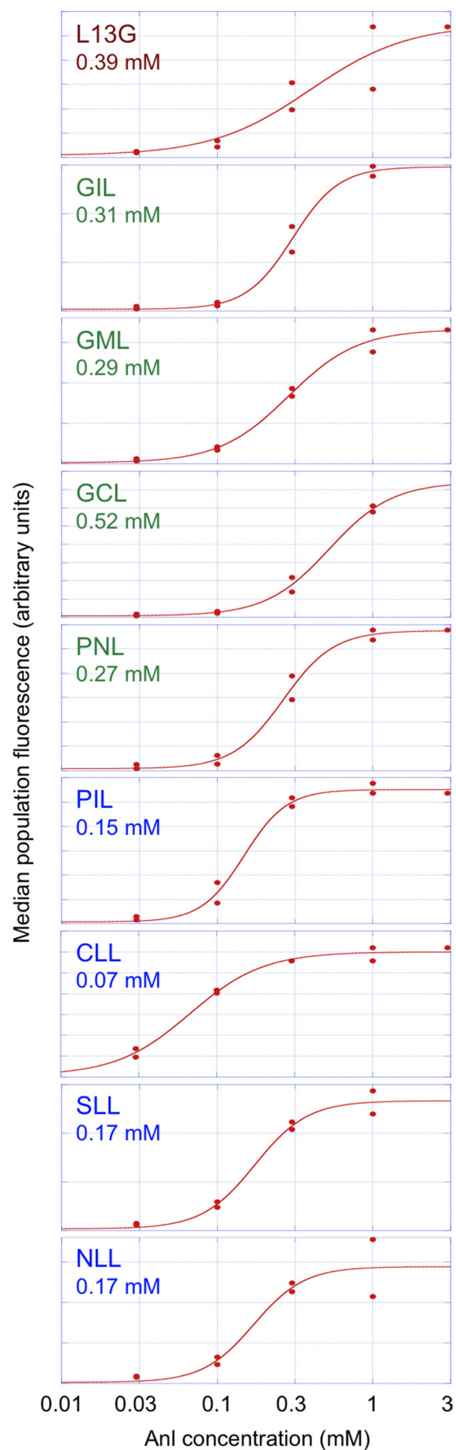

Fig. S2. Extent of fluorescence labeling at different Anl concentrations on cells expressing various MetRS mutants. OmpC expression was carried out in M9 + 19aa media supplemented with 0.03, 0.01, 0.3, 1.0, or 3.0 mM Anl. Median fluorescence of the labeled cell population was determined by flow cytometry. The median fluorescence for cells treated at each Anl concentration (red circles), and the Hill equation fit to the data (red lines) are displayed for each MetRS mutant tested. In cases where a fluorescence drop was observed going from 1.0 to 3.0 mM Anl, the final data point was not included in equation fitting. The EC_{50} value obtained from the Hill equation is indicated on each plot in millimolar units. Mutants selected at 0.3 mM Anl (labeled in blue) exhibit lower EC_{50} values than mutants selected at 1.0 mM Anl (labeled in green).

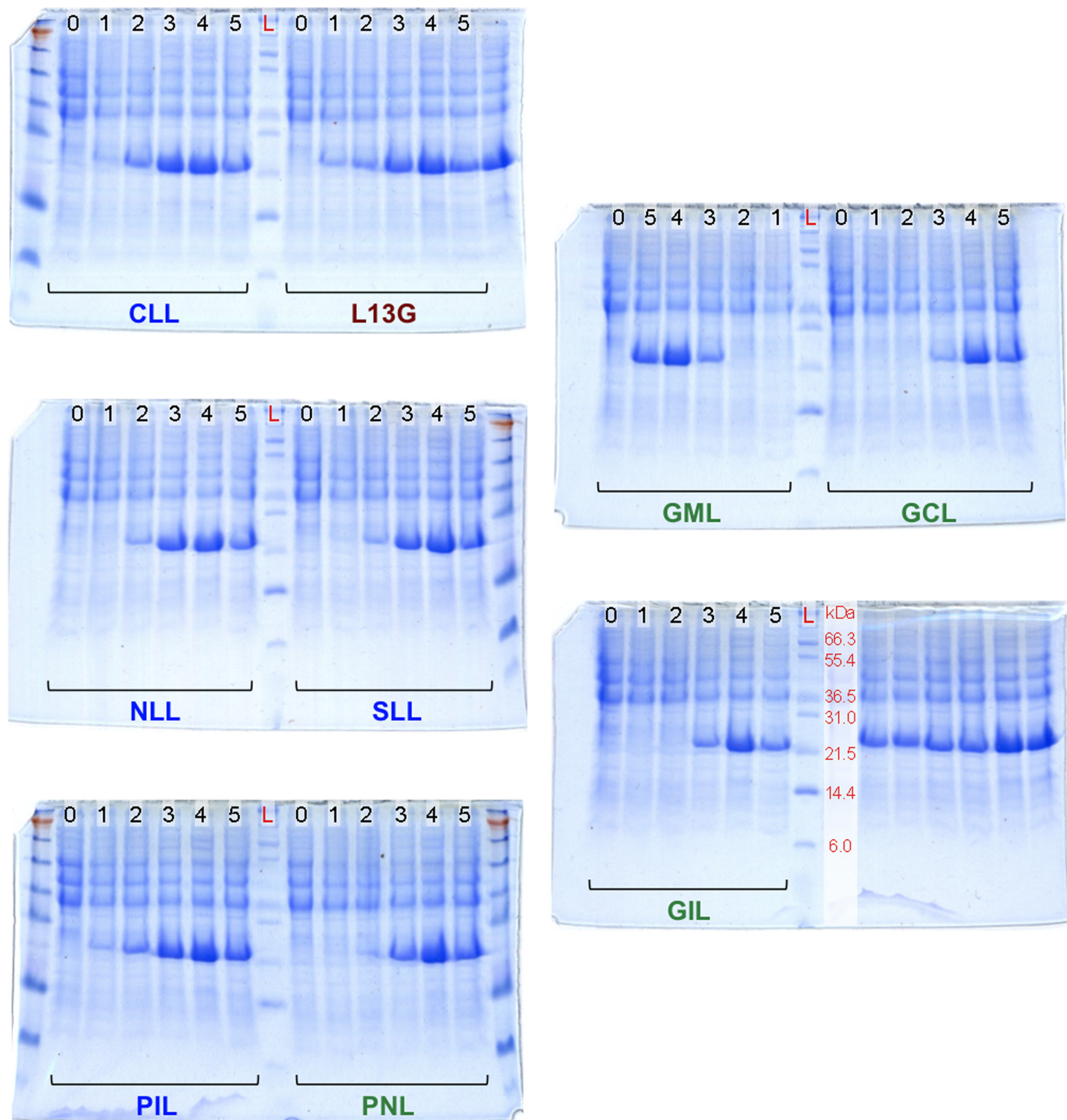

Fig. S3. Unmodified images of SDS/PAGE gels of cell lysates showing mDHFR synthesis at various expression conditions. Expression of mDHFR was performed in M15MA[pREP4/pAJL-61] cells encoding the MetRS mutants obtained from screens carried out at 0.3 mM and 1.0 mM Anl, and the L13G mutant. Cells grown to midlog phase in M9 + 20aa medium were then shifted into M9 + 19aa medium (no Met) supplemented with either 0.1, 0.3, or 1.0 mM Anl (lanes 2, 3 and 4, respectively), 40 mg/L Met (lane 5), or no 20th amino acid (lane 1) before induction of mDHFR synthesis. Cell aliquots obtained before induction are shown in lane 0. Mark12 (Invitrogen) ladder was used in all gels (lane L). mDHFR has a molecular mass of 24.0 kDa.

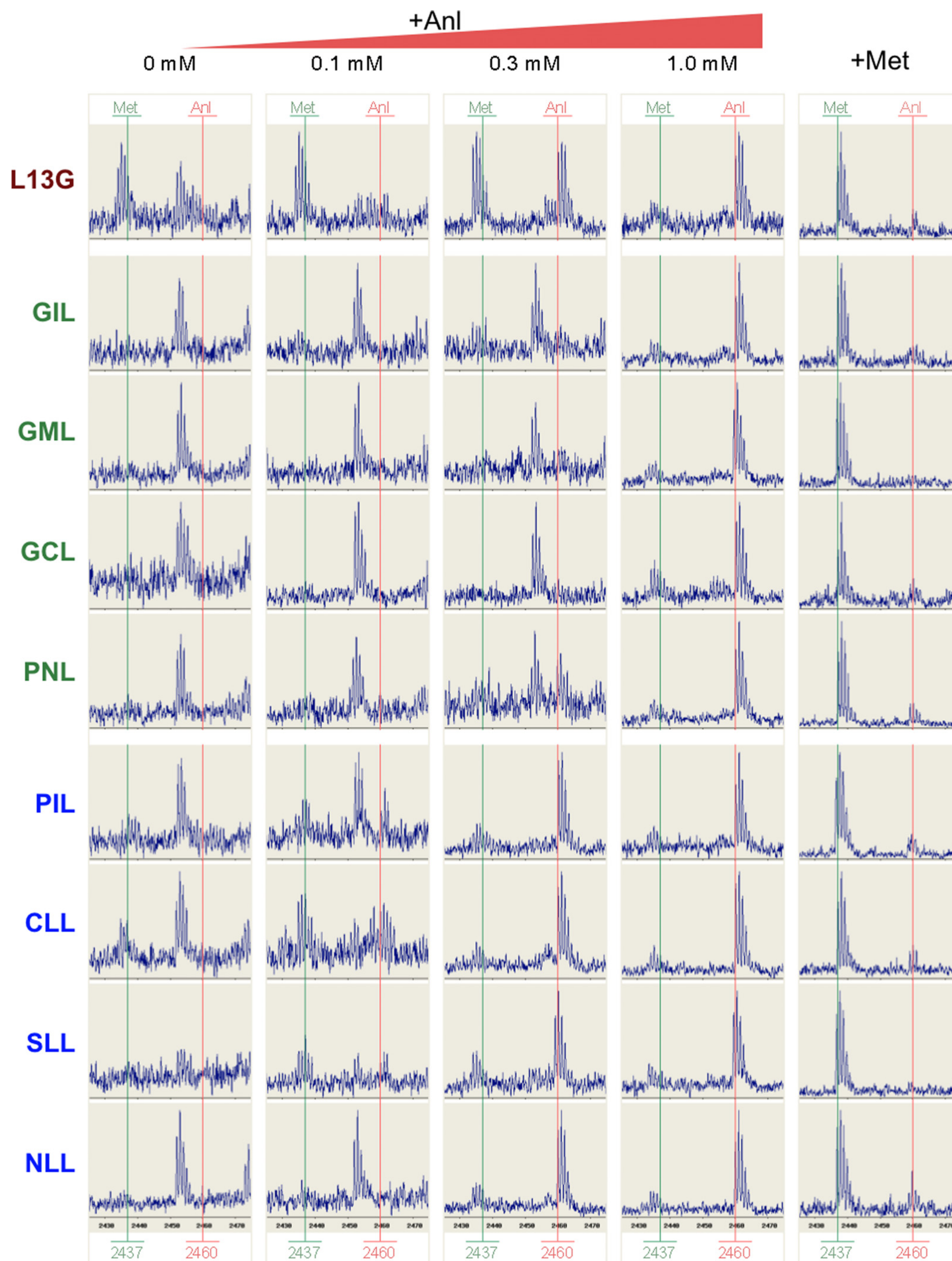

Fig. S4. MALDI-MS analysis of tryptic peptides from mDHFR expressed at varying concentrations of Anl in cells bearing MetRS mutants. Expression of mDHFR was performed in M15MA[pREP4/pAJL-61] cells encoding the MetRS mutants obtained from screens carried out at 0.3 mM and at 1.0 mM Anl, and the L13G mutant. Expression was performed in M9 + 19aa (no Met) media supplemented with either 0.1, 0.3, or 1.0 mM Anl, 40 mg/L Met, or no 20th amino acid. The expressed mDHFR was purified, and its tryptic digests were analyzed by MALDI-MS. The peaks belonging to the tryptic fragment IMQEFESDTFFPEIDLGKYK (2,437.16 Da) are shown in the range 2,425–2,475 Da. Replacement of Met by Anl results in a 23.05 Da increase in peptide mass. The expected locations of the Met- and Anl-containing peptides are marked on the spectra by green and red lines, respectively. Consistent with SDS/PAGE (Fig. 4*B* and Fig. S3), the MALDI-MS data shows that mutants selected at 0.3 mM Anl (labeled in blue) allow near complete replacement of Met by Anl at 0.3 mM Anl, whereas little or no Anl incorporation was observed at this concentration with mutants selected at 1.0 mM Anl (labeled in green). In the absence of Met and at low Anl concentrations the L13G mutant permits the incorporation of an alternate amino acid 3 Da lighter than Met. Such a mass shift corresponds to the natural amino acids glutamine or lysine. Identical incorporation patterns were observed when the mass of an alternate tryptic peptide (QNLVINGR; 930.52 Da) was monitored.

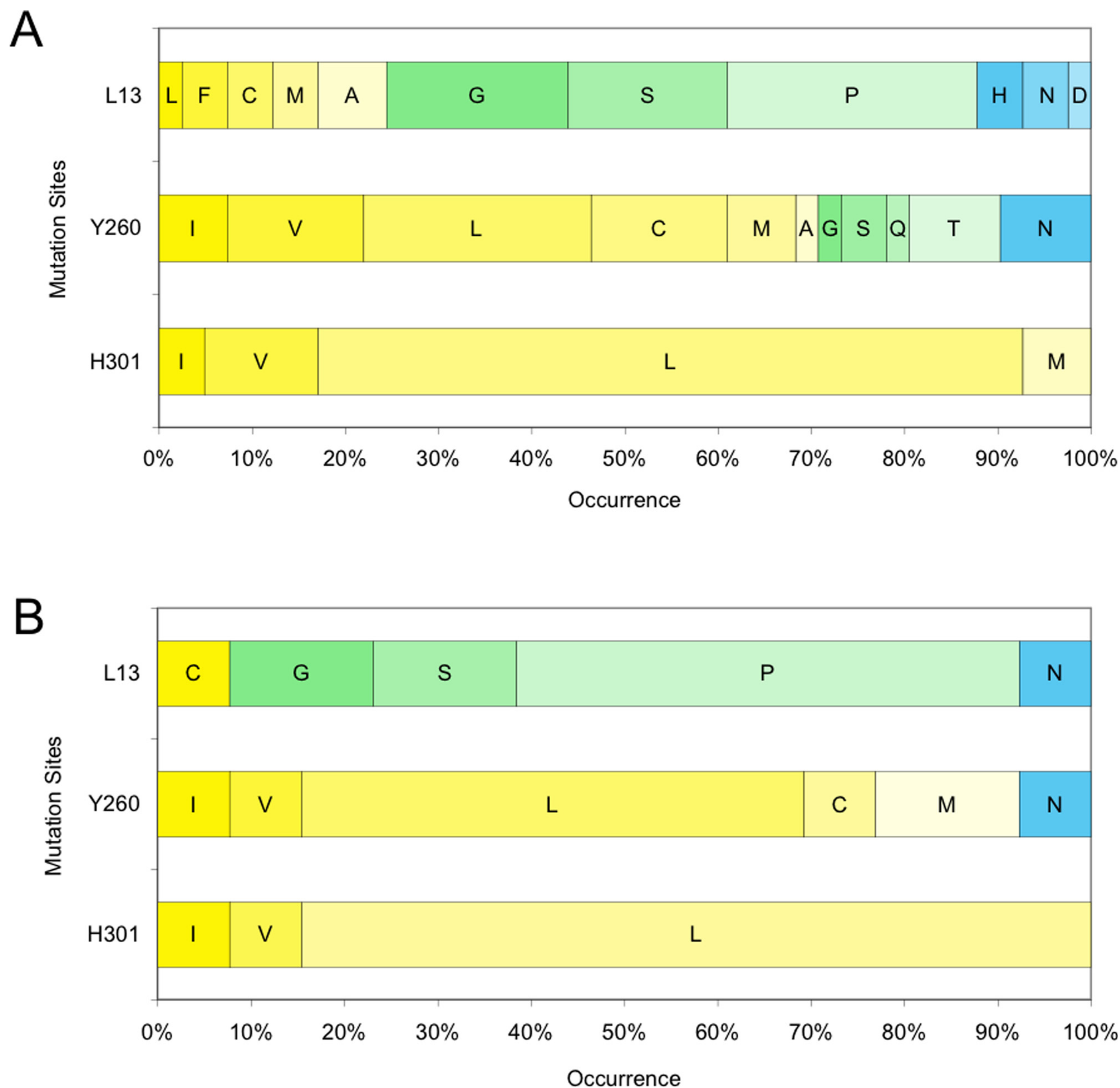

Fig. S5. Distribution of the selected mutations across the randomized sites on MetRS. From 162 active clones isolated through screening the naïve library, 41 distinct MetRS mutants that show activity toward Anl were identified (Table S2). Distribution of mutations observed at each randomized site among the 41 mutants is displayed. Substituted residues at each site are ordered by decreasing hydrophobicity according to the Kyte-Doolittle scale [Kyte J, Doolittle RF (1982) *J Mol Biol* 157:105–132]. Based on this scale, residues are grouped as hydrophobic (>1.0; shades of yellow) or hydrophilic (<−2.0; shades of blue). Residues that rank between these two groups are shown in shades of green. Mutation distributions compiled (*A*) from all distinct mutants (n = 41) isolated, and (*B*) from mutants (n = 13) observed in high stringency screens (0.3 mM Anl or better) are displayed. Results show that mutations to aliphatic nonpolar residues at positions 260 and 301 are compatible with Anl activity. In contrast, small residues with more hydrophilic character are preferred at position 13.

Table S1. Kinetic parameters for the activation of Anl and Met by the MetRS variants identified in this study and the L13G mutant

MetRS	EC50, mM	Anl activation				Met activation				Selectivity[†]
		K_m, mM	k_{cat}, s^{-1}	k_{cat}/K_m, M^{-1} s^{-1}	k_{cat}/K_m, rel*	K_m, mM	k_{cat}, s^{-1}	k_{cat}/K_m, M^{-1} s^{-1}	k_{cat}/K_m, rel*	
wt[‡]	-	-	-	-	-	0.024	13.30	550,000	1	-
L13G	0.39	5.1 ± 1.3	0.85 ± 0.07	170 ± 40	1:3,200	0.16 ± 0.06	0.87 ± 0.18	5,600 ± 1,500	1:92	0.03
PLL	-	1.5 ± 0.7	0.92 ± 0.22	650 ± 150	1:850	5.3 ± 1.9	1.0 ± 0.2	200 ± 50	1:2,700	3.2
NLL	0.17	2.2 ± 0.8	0.87 ± 0.11	410 ± 80	1:1,400	2.6 ± 0.5	0.86 ± 0.07	350 ± 70	1:1,600	1.2
SLL	0.17	4.2 ± 1.1	0.90 ± 0.13	220 ± 20	1:2,600	1.0 ± 0.3	0.89 ± 0.16	900 ± 80	1:610	0.24
CLL	0.07	2.0 ± 0.7	0.98 ± 0.16	520 ± 120	1:1,100	0.14 ± 0.04	0.94 ± 0.18	7,100 ± 600	1:77	0.07
PLI	-	2.4 ± 1.0	0.89 ± 0.18	410 ± 140	1:1,300	-	-	-	-	-
PIL	0.15	6.6 ± 1.8	0.84 ± 0.08	130 ± 40	1:4,100	-	-	-	-	-
PNL	0.27	5.3 ± 2.2	0.47 ± 0.10	98 ± 37	1:5,600	-	-	-	-	-
GML	0.29	11 ± 5	0.48 ± 0.11	47 ± 13	1:12,000	-	-	-	-	-
GIL	0.31	15 ± 3	0.22 ± 0.01	14 ± 3	1:38,000	-	-	-	-	-
GCL	0.52	12 ± 3	0.12 ± 0.01	11 ± 2	1:51,000	-	-	-	-	-

*Relative to k_{cat}/K_m of the wild-type enzyme for Met.
[†]Selectivity is defined as the ratio of k_{cat}/K_m for Anl to that for Met.
[‡]Activation parameters for wild-type MetRS are taken from ref. 1.

1. Kiick KL, Tirrell DA (2000) Protein engineering by *in vivo* incorporation of non-natural amino acids: Control of incorporation of methionine analogues by methionyl-tRNA synthetase. *Tetrahedron* 56:9487–9493.

Table S2. MetRS mutants identified through library screening to allow Anl incorporation at 1.0 mM concentration in vivo

Occurrence	L13	Y260	H301	Populations from which clones were isolated*
62†	**N**	**L**	**L**	**2.2, 2.3, 3.2, 3.3, 4.3, 5.3c, 5.3d, 5.4b**
14†‡	**P**	**L**	**L**	**2.2, 2.3, 5.3c, 5.3d**
12†	**S**	**L**	**L**	**2.2, 2.3, 3.2, 5.4b**
9†	**P**	**L**	**I**	**2.2, 5.3c, 5.4b**
5	**P**	**N**	**L**	**1.2, 3.3, 4.3**
4	**G**	**M**	**L**	**1.2**
4	G	T	L	1.2, 3.3, 6.2
4	G	V	L	6.2
4†	S	M	L	3.2, 3.3
3	A	V	L	6.2
3	**G**	**I**	**L**	**1.2**
3†	G	N	L	3.2, 3.3, 4.3
2	A	Q	L	1.2, 6.2
2†	**C**	**L**	**L**	**3.2**
2	**G**	**C**	**L**	**1.2**
2†	G	L	L	3.2, 5.3d
2	P	N	V	6.2
2	S	S	L	6.2
1	A	C	L	1.2
1	C	V	L	1.2
1	D	L	I	1.2
1	F	G	L	6.2
1	F	V	L	3.3
1	G	L	V	6.2
1	H	C	L	1.2
1	H	T	M	1.2
1	L	C	L	6.2
1	M	A	L	1.2
1	M	S	L	6.2
1	N	V	L	1.2
1†	P	C	L	3.2
1†	**P**	**I**	**L**	**3.2**
1	P	I	M	1.2
1†	P	L	V	3.2
1†	P	M	L	2.2
1	P	T	V	1.2
1†	P	V	L	5.3d
1	S	C	M	6.2
1	S	L	V	4.3
1	S	N	L	6.2
1	S	T	L	1.2
162	Total			

Mutants discussed in the main article are shown in bold.

*The identifiers for the populations are listed (e.g., population ''2.2'' refers to LYH.2.2).

†These 13 mutants were observed in high stringency screens (0.3 and 0.1 mM Anl screens and screens against Met).

‡Count also includes 12 clones carrying MetRS-PLL with the additional H98N surface mutation.

Table S3. Primers used in this study

Primer name	Sequence
Primers used in library construction	
Lib_fwd2	CTCAGTACCAGTTCGACTTCGGTCTGCGTCCGTCCCTG
L13_lib2-r	GATTGAGCCGTTAGCGTACGG**MNN**TGCGCACGTCACCAG
L13_lib3	GAAAATTCTGGTGACGTGCGCA**NNK**CCGTACGCTAACGGCTCA
Y260_lib3-r	CTTGAAAGAACCCAT**MNN**GCCAATCGGTGCGTCCAGC
Y260_lib3	GCTGGACGCACCGATTGGC**NNK**ATGGGTTCTTTCAAG
H301_lib3-r	CAGCATGGCAGGCCAGAACAGGCT**MNN**GAAGTAAACAATATC
H301_lib2	GATATTGTTTACTTC**NNK**AGCCTGTTCTGGCCTGCCATGC
lib_rev2	CAGTACCGGCTTCAGGTAAGTCATCAGCACGCGGAAC
lib_fwd3	CAGTACCAGTTCGACTTCGGTCTGCGTC
lib_rev3	CTTCAGGTAAGTCATCAGCACGCGGAAC
Primers used in the construction of pMTY21 variants	
MRS_BamHI	TTCCGCGGATCCATGACTCAAGTCGCGAAGAAAATTC
MRS_SalI-r	TTTGGGGTCGACTCATTTAGAGGCTTCCACCAGTG
Site directed mutagenesis primers	
eM_L13G	CTGGTGACGTGCGCA**GGT**CCGTACGCTAACGGCTC
eM_G13C	ATTCTGGTGACGTGCGCA**TGT**CCGTACGCTAAC
eM_G13N	ATTCTGGTGACGTGCGCA**AAC**CCGTACGCTAAC
eM_G13P	ATTCTGGTGACGTGCGCA**CCG**CCGTACGCTAAC
eM_G13S	ATTCTGGTGACGTGCGCA**AGC**CCGTACGCTAAC
eM_Y260C	GACGCACCGATTGGC**TGC**ATGGGTTCTTTCAAG
eM_Y260I	GGACGCACCGATTGGC**ATC**ATGGGTTCTTTCAAG
eM_Y260L	GGACGCACCGATTGGC**CTG**ATGGGTTCTTTCAAG
eM_Y260M	GGACGCACCGATTGGC**ATG**ATGGGTTCTTTCAAG
eM_Y260N	GGACGCACCGATTGGC**AAC**ATGGGTTCTTTCAAG
eM_H301L	GATATTGTTTACTTC**CTG**AGCCTGTTCTGGCCTGC

For mixed nucleotides, N = A, T, G, C, K = G, T and M = C, A. Only the forward sequence is provided for site-directed mutagenesis primers. All sequences are listed in the 5′-to-3′ direction. Mutation sites are indicated in bold.