Chemistry & Biology, Vol. 10, 511–519, June, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S1074-5521(03)00123-6

Progress Toward an Expanded Eukaryotic Genetic Code

Jason W. Chin,[1,2] T. Ashton Cropp,[1] Stephanie Chu,[1] Eric Meggers,[1,3] and Peter G. Schultz[1,*]
[1]Department of Chemistry
Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, California 92037

Summary

Expanding the eukaryotic genetic code to include unnatural amino acids with novel properties would provide powerful tools for manipulating protein function in eukaryotic cells. Toward this goal, a general approach with potential for isolating aminoacyl-tRNA synthetases that incorporate unnatural amino acids with high fidelity into proteins in *Saccharomyces cerevisiae* is described. The method is based on activation of GAL4-responsive *HIS3*, *URA3*, or *lacZ* reporter genes by suppression of amber codons in GAL4. The optimization of *GAL4* reporters is described, and the positive and negative selection of active *Escherichia coli* tyrosyl-tRNA synthetase (EcTyrRS)/$tRNA_{CUA}$ is demonstrated. Importantly, both selections can be performed on a single cell and with a range of stringencies. This method will facilitate the isolation of a range of aminoacyl-tRNA synthetase (aaRS)/$tRNA_{CUA}$ activities from large libraries of mutant synthetases.

Introduction

The recent addition of unnatural amino acids to the genetic code of *Escherichia coli* (*E. coli*) provides a powerful new approach for analyzing and manipulating protein structure and function both in vitro [1, 2] and in vivo [3]. Amino acids with photoaffinity labels, heavy atoms, keto and olefinic groups, and chromophores have been incorporated into proteins in *E. coli* with an efficiency and fidelity rivaling that of the common twenty amino acids [1–6]. However, currently these methods for incorporation of unnatural amino acids can only be applied in *E. coli*. The extension of this methodology to eukaryotic cells would provide powerful tools for analyzing and controlling processes in these cells at the molecular level.

Unnatural amino acids have been introduced previously into the nicotinic acetylcholine receptor in *Xenopus* oocytes [7] by microinjection of a chemically misacylated *Tetrahymena thermophila* tRNA [8] and the relevant mRNA. This has allowed detailed biophysical studies of the receptor in oocytes by the introduction of amino acids containing side chains with unique physical or chemical properties [9]. Unfortunately, this methodology is limited to proteins in cells that can be microinjected, and because the tRNA is chemically acylated in vitro and cannot be reacylated, the yields of protein are very low. This, in turn, necessitates sensitive techniques to assay protein function.

We, and others [10–12], are interested in using genetically encoded aminoacyl-tRNA synthetase/tRNA pairs for the incorporation of unnatural amino acids into proteins in eukaryotic cells in response to an amber codon. This would have significant technical and practical advantages, since tRNAs would be reacylated with the unnatural amino acid by their cognate synthetases, leading to large amounts of mutant protein. Moreover, genetically encoded aminoacyl-tRNA synthetases and tRNAs are, in principle, heritable, allowing the unnatural amino acid to be incorporated into proteins through many cell divisions without exponential dilution.

The steps necessary to add new amino acids to the genetic code of *E. coli* have been described [13, 14], and similar principles are likely to be useful for expanding the genetic code of eukaryotes. In the first step, an orthogonal aminoacyl-tRNA synthetase (aaRS)/$tRNA_{CUA}$ pair is identified. This pair must function with the host cell's translational machinery, but the aaRS must not charge any endogenous tRNAs with an amino acid, and the $tRNA_{CUA}$ must not be aminoacylated by any endogenous synthetases [15]. In a second step, those aaRS/tRNA pairs that are capable of using only the unnatural amino acid are selected from a library of mutant aaRSs. In *E. coli*, the selection of unnatural amino acid utilizing variants of MjTyrRS was carried out using two-step "double sieve" selections [13]. A comparable selection method is required in eukaryotic cells.

Saccharomyces cerevisiae (*S. cerevisiae*) was chosen as the eukaryotic host organism, as it is unicellular and has a rapid generation time, as well as relatively well-characterized genetics [16]. Moreover, since the translational machinery of eukaryotes is well conserved [17–19], it is likely that aaRSs genes for the incorporation of unnatural amino acids discovered in *S. cerevisiae* can be "cut and pasted" into higher eukaryotic organisms and used in partnership with cognate tRNAs [12, 20] to incorporate unnatural amino acids. The expansion of the genetic code of *S. cerevisiae* is therefore a gateway to expanding the genetic code of complex multicellular eukaryotic organisms [21]. Unfortunately, the tyrosyl pair derived from *Methanococcus jannaschii* TyrRS (MjTyrRS)/tRNA [22], which we have previously used to expand the genetic code of *E. coli*, is not orthogonal in eukaryotic organisms [23], and a new orthogonal pair is required to expand the eukaryotic genetic code. Several potential orthogonal pairs have been described in *S. cerevisiae* [11, 24–26]. Schimmel and coworkers have shown that the *E. coli* tyrosyl-tRNA synthetase (EcTyrRS)/$tRNA_{CUA}$ pair suppresses amber codons in *S. cerevisiae* with tyrosine. They have also shown that *E. coli* $tRNA_{CUA}$ is not charged by endogenous aminoacyl

*Correspondence: schultz@scripps.edu
[2]Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, UK.
[3]Present address: Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104.

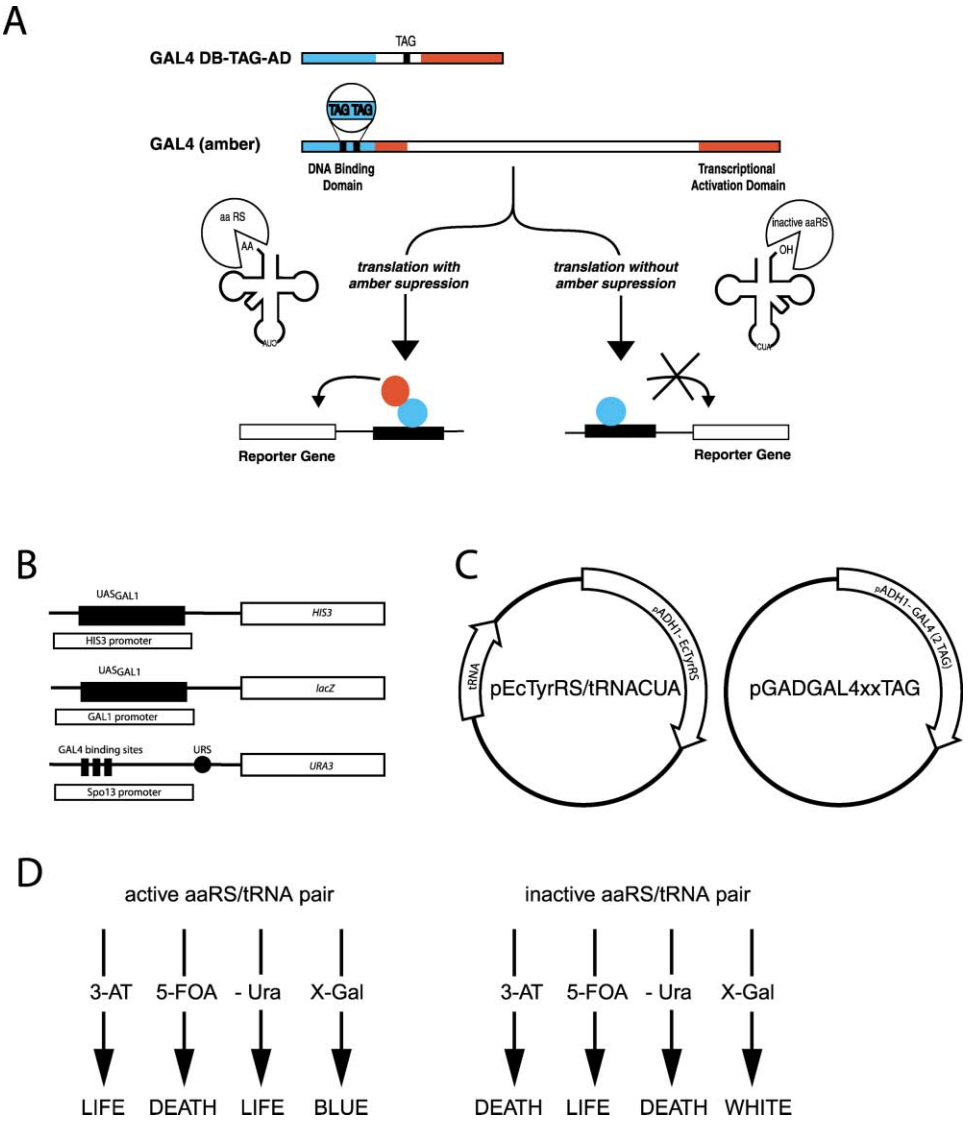

Figure 1. A General Positive and Negative Selection Scheme for Expanding the Genetic Code of *S. cerevisiae*

(A) A schematic of the first generation (top) and second generation (below) GAL4 reporters (the properties of which are shown in Figures 2 and 3, respectively). The first generation reporter contains the DNA binding domain of GAL4 (1–147, shown in blue) linked by the sequence SRSNQTSLYKKAGX (shown in white) to the GAL4 activation domain (768–881, shown in red), where X is the amino acid inserted in response to an amber codon. Second-generation GAL4 reporters are derived from full-length GAL4 (1–881) and contain amber codons at positions in the DNA binding domain. Activated transcription of genes (described in [B]) is driven by amber suppression of TAG codons in GAL 4 by an active aminoacyl tRNA synthetase (aaRS)/tRNA pair.

(B) The genes in the yeast cell line MaV203 whose transcription is activated by GAL4; all genes are genomically integrated in the same cell line.

(C) Two-micrometer plasmids that encode (left) the Ec TyrRS and cognate suppressor $tRNA_{CUA}$ and (right) the GAL4 amber mutants used in data presented in Figure 3 and beyond.

(D) A schematic of the phenotypes expected with active and inactive orthogonal aaRS/tRNA pairs on the selective media indicated.

tRNA synthetases in the yeast cytosol [24, 25] (Figure 2). In addition, numerous groups [18, 26, 27] have shown that EcTyrRS does not charge yeast tRNA in vitro. Thus, the EcTyrRS/$tRNA_{CUA}$ pair is a likely candidate for an orthogonal pair in *S. cerevisiae* as well as in higher eukaryotes [20].

To broaden the substrate specificity of EcTyrRS in *E. coli*, Nishimura and coworkers screened an error-prone PCR-generated library of mutants of EcTyrRS and discovered a mutant with an improved ability to incorporate 3-azatyrosine [28]. However, this amino acid is incorporated throughout the proteome of *E. coli*, and the evolved enzyme still prefers tyrosine as a substrate. Yokoyama and coworkers screened a small collection of designed active site variants of EcTyrRS in an in vitro translation system and discovered an EcTyrRS variant that utilizes 3-iodotyrosine more effectively than tyrosine [29]. In contrast to the enzymes we have evolved in *E. coli* [1, 2, 6, 30], this enzyme still incorporates tyrosine in the absence of the unnatural amino acid [29]. Recently, Yokoyama and coworkers have also demonstrated that this EcTyrRS mutant functions with a $tRNA_{CUA}$ from *Bacillus*

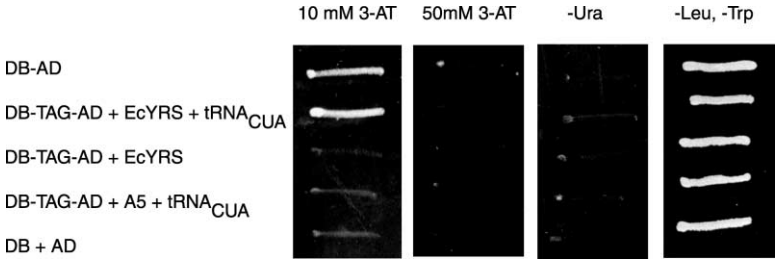

Figure 2. The EcTyrRS- and $tRNA_{CUA}$-Dependent Phenotypes of First-Generation GAL4 Reporters on Selective Media

DB-AD is a fusion between the GAL4 DNA binding domain and activation domain. DB-TAG-AD has a TAG codon replacing a tyrosine codon in the synthetic linker between DB and AD (as described in the legend to Figure 1). A5 is an inactive version of EcTyrRS in which five residues in the active site (detailed in Experimental Procedures) have been mutated to alanine.

stearothermophilus to suppress amber codons in mammalian cells [12].

A critical requirement is that any amino acid added to the eukaryotic genetic code be incorporated with a fidelity similar to that of the common twenty amino acids. To accomplish this, we have pursued a general, in vivo selection method for the discovery of EcTyrRS/$tRNA_{CUA}$ variants that function in *S. cerevisiae* to incorporate unnatural amino acids but none of the common amino acids in response to the amber codon TAG. A major advantage of a selection is that enzymes that selectively incorporate unnatural amino acids can be rapidly selected and enriched from libraries of 10^8 EcTyrRS active site variants, 6–7 orders of magnitude more diversity than has been screened in vitro [29]. This increase in diversity vastly increases the likelihood of isolating EcTyrRS variants for the incorporation of a diverse range of useful functionality with very high fidelity [22].

To extend the selection approach to *S. cerevisiae*, we took advantage of the transcriptional activator protein, GAL4 [31–34] (Figure 1). The N-terminal 147 amino acids of this 881 amino acid protein form a DNA binding domain (DBD) that binds DNA sequence specifically [35, 36]. The DBD is linked by an intervening protein sequence to a C-terminal 113 amino acid activation domain (AD) that can activate transcription when bound to DNA [37, 38]. We envisioned that by placing amber codons toward the N-terminal DBD of a single polypeptide that contained *both* the N-terminal DBD of GAL4 and its C-terminal AD, it should be possible to link amber suppression by the EcTyrRS/$tRNA_{CUA}$ pair to transcriptional activation by GAL4 (Figure 1A). By the choice of appropriate GAL4-activated reporter genes, it should be possible to perform both positive and negative selections on synthetase or suppressor tRNA activity (Figure 1B). While many reporter genes based on complementing the amino acid auxotrophy of a cell can be used for positive selections (eg., *URA3, LEU2, HIS3, LYS2*), the *HIS3* gene is particularly attractive, as the activity of the protein it encodes (imidazole glycerol phosphate dehydratase) can be modulated in a dose-dependent manner by addition of 3-aminotriazole (3-AT) [39]. In *S. cerevisiae*, fewer genes have been used for negative selections. One of several negative selections [40–43] that has been successfully used is the *URA3*/5-flurorotic acid (5-FOA) negative selection [44] system described in the "reverse two-hybrid" system developed by Vidal and coworkers [42, 43]. In the reverse two-hybrid system, a genomically integrated *URA3* reporter is placed under a tightly controlled promoter that contains GAL4 DNA binding sites. When two proteins that interact are produced as fusions to the GAL4 DBD and GAL4 AD, they reconstitute the activity of GAL4 and activate transcription of *URA3*. In the presence of 5-FOA, the URA3 gene product converts 5-FOA to a toxic product, killing the cell [44]. This selection has been used to select for proteins that disrupt a protein-protein interaction and for mutations that disrupt a protein-protein interaction. A variant for screening small molecule inhibitors of protein-protein interactions has also been described [45].

We now demonstrate that the appropriate choice of amber codons in full-length GAL4 allows efficient positive selections for active EcTyrRS variants using either *HIS3* or *URA3* GAL4-activated reporters to complement histidine or uracil auxotrophy in yeast cells. Moreover, the *URA3* reporter can be used in negative selections for inactive EcTyrRS variants in the presence of 5-FOA. In addition, colorometric assays using *lacZ* can be used to read out aminoacyl-tRNA synthetase activity in yeast cells.

Results and Discussion

The EcTyrRS gene was expressed under the control of the constitutive ADH1 promoter, and the *E. coli* tyrosyl $tRNA_{CUA}$ gene was expressed from the same high-copy yeast plasmid (pEcTyrRS/tRNACUA, Figure 1C). Upon cotransformation of pEcTyrRS/tRNACUA and a low-copy reporter that contains a single amber mutation between the DNA binding domain and activation domain of a chimeric GAL4 construct (GAL4 DB-TAG-AD, Figure 1A, top) into MaV203, cells grew on selective media lacking histidine and containing 10–20 mM 3-AT (Figure 2). When MaV203 cells were transformed with the same GAL4 construct and either an inactive synthetase mutant (A5) or a construct lacking the *Ect*RNA_{CUA} gene, no growth was observed on 10 mM 3-AT (Figure 2). These experiments establish that EcTyrRS can be constitutively expressed in a functional form from the ADH1 promoter, that there is minimal endogenous amber suppression in MaV203, and that there is little charging of EctRNA_{CUA} by yeast synthetases in this system [24, 25]. Since EcTyrRS does not charge *S. cerevisiae* tRNA [18, 26, 27], these experiments confirm that EcTyrRS/EctRNA_{CUA} are an orthogonal pair in *S. cerevisiae*.

While the first-generation GAL4 chimera was able to activate transcription of the weak *HIS3* reporter, it was unable to activate transcription of the *URA3* reporter in MaV203 sufficiently to allow significant growth on concentrations of 3-AT greater than 20 mM or on −Ura plates (Figure 2). For the purposes of selection of

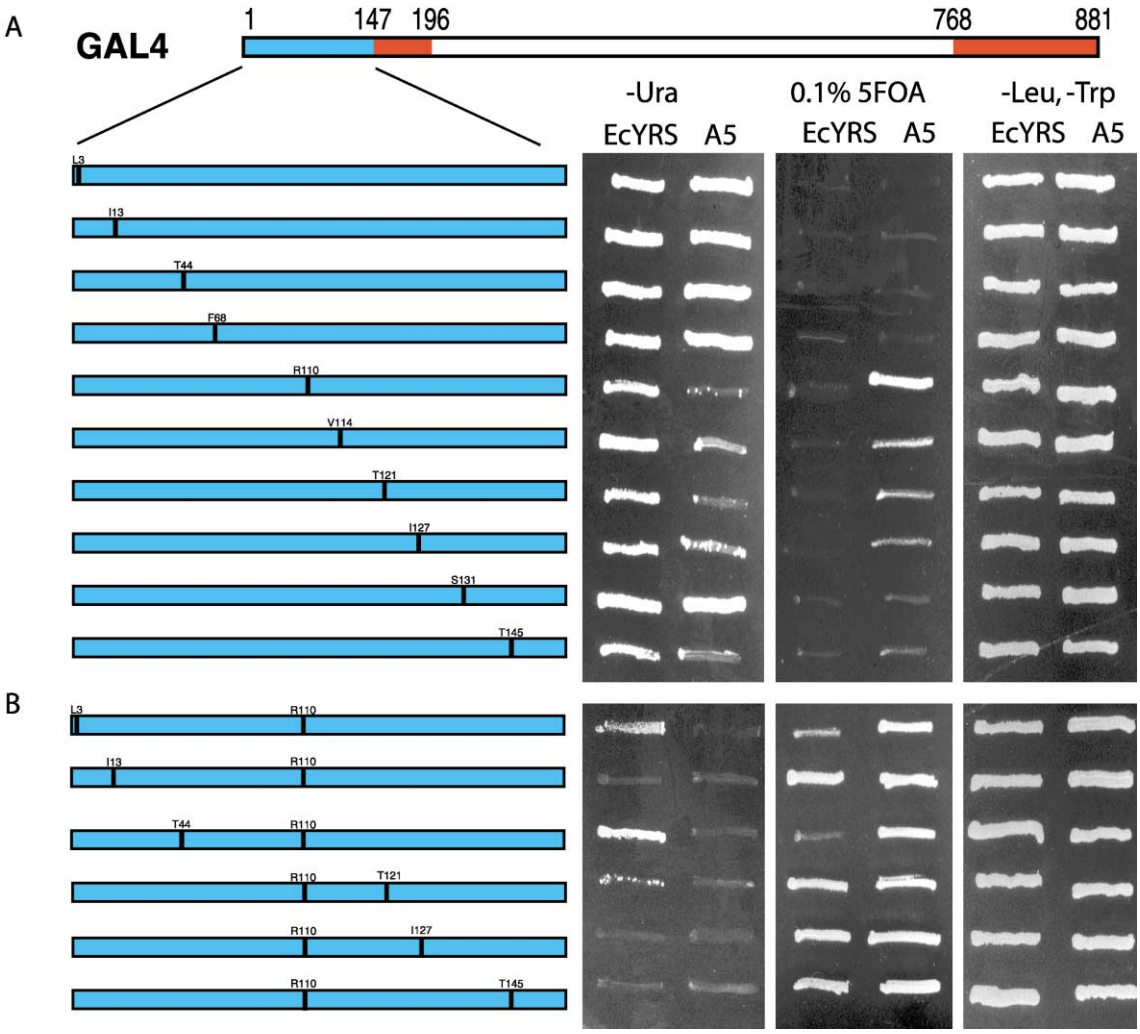

Figure 3. The EcTyrRS- and $tRNA_{CUA}$-Dependent Phenotypes of Second-Generation GAL4 Reporters on the Selective Media Indicated

A schematic of GAL4 is shown at the top of the figure. The DNA binding domain is indicated in blue, and the major and cryptic activation domains are red. The DNA binding domain mutants are indicated (left), and the growth of MaV203 transformed with active or inactive EcTyrRS and cognate tRNA on positive and negative selective media is shown.

EcTyrRS variants, a second-generation GAL4 construct was made. This GAL4 reporter was designed to be more active, have a greater dynamic range, and avoid the accumulation of revertants. To increase the activity of the GAL4 reporters, full-length GAL4 was used (which has a transcriptional activiation activity twice that of a DBD-AD fusion [37]) under the control of a strong ADH1 promoter, and a high-copy 2 μm plasmid (with a copy number 10–30 times that of the centromeric plasmid of the initial GAL4 chimera) was used. An increase in both the copy number of the plasmid and the activity of the protein it encodes should extend the dynamic range of the reporters. Amber mutations were targeted to the region of the GAL4 gene that encodes amino acid residues 2 and 147 (Figure 3). This region is sufficient for sequence-specific DNA binding [35] and lies to the 5′ side of the first cryptic activation domain in the GAL4 gene [37], such that the truncated products produced in the absence of amber suppression are not anticipated to activate transcription. We were guided in our choice of amino acid codons to mutate by previous saturation mutagenesis selections on GAL4 [46], as well as the X-ray structures of the N-terminal DNA binding domain of GAL4 [47, 48] and the NMR structure of its dimerization region [49].

Full-length GAL4 was cloned into a small pUC-based vector to allow the rapid construction of ten single amber mutants (at the codons for amino acids L3, I13, T44, F68, R110, V114, T121, I127, S131, and T145) by site-directed mutageneisis. GAL 4 and the resulting amber mutants were then subcloned into a 2 μm yeast vector under the control of the full-length ADH1 promoter to create pGADGAL4 and a series of amber mutants denoted pGADGAL4 (xxTAG) (Figure 1C), where xx denotes the amino acid codon in the GAL4 gene that was mutated to the amber codon. Each GAL4 mutant was cotransformed with either EcTyrRS/$tRNA_{CUA}$ or A5/$tRNA_{CUA}$ into MaV203 cells, converting transformants to leucine and tryptophan protrophy. pGADGAL4 itself transformed with very low efficiency ($<10^{-3}$ times that

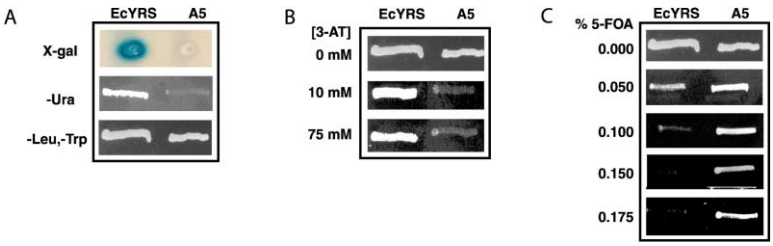

Figure 4. pGADGAL4(T44TAG, R110TAG) with and without EcTyrRS Spans the Dynamic Range of the Reporters in MaV203

of the GAL4 amber mutants) and is presumably deleterious to MaV203 cells at such high copy; no such effect was observed with the amber mutants of GAL4.

The phenotypes of MaV203 cells containing GAL4 reporters, in the presence of $tRNA_{CUA}$ and an active or inactive synthetase, were assayed on −Ura plates and 0.1% 5-FOA plates (Figure 3A). Five GAL4 mutants (L3TAG, I13TAG, T44TAG, F68TAG, and S131TAG) grew on −Ura plates and failed to grow on 0.1% 5-FOA in the presence of either a wild-type or inactive EcTyrRS. In these amber mutants, endogenous suppression is apparently sufficient to push the EcTyrRS/$tRNA_{CUA}$-mediated suppression beyond the dynamic range of the *URA3* reporter in MaV203. Five GAL4 single amber mutants (R110TAG, V114TAG, T121TAG, I127TAG, and T145TAG) grew in the absence of uracil and in the presence of EcTyrRS/$tRNA_{CUA}$ (but not A5/$tRNA_{CUA}$) and showed the reverse phenotype on 5-FOA. These mutants show EcTyrRS-dependent phenotypes that fall within the dynamic range of the *URA3* reporter in MaV203. The cleanest EcTyrRS-dependent phenotype on both −Ura and 0.1% 5-FOA was observed with the R110 TAG mutant of GAL4. However, this mutant showed some blue color in X-GAL assays when cotransformed with A5 (see Supplemental Data at http://www.chembiol.com/cgi/content/full/10/6/511/DC1). To further improve the dynamic range, we made a series of six double amber mutants of GAL4 containing R110 TAG (Figure 3B) (L3TAG, R110TAG; I13TAG, R110TAG; T44TAG, R110TAG; R110TAG, T121TAG; R110TAG, I127TAG; and R110TAG, T145TAG). Four of these double mutants (I13TAG, R110TAG; R110TAG, T121TAG; R110TAG, I127TAG and T145TAG, R110TAG) were unable to grow in the absence of uracil and grew on 0.1% 5-FOA. These double mutants have activities outside (below) the dynamic range of the plate assays. Two of the double mutants (L3TAG, R110TAG and T44TAG, R110TAG) grew in the presence of wild-type EcTyrRS/$tRNA_{CUA}$ but not with A5/$tRNA_{CUA}$ on −Ura plates; these mutants also showed the expected reciprocal phenotypes on 5-FOA. PGADGAL4(T44TAG, R110TAG), the more active of these two GAL4 mutants, was selected for a more detailed characterization (Figure 4). MaV203 containing pGADGAL4(T44TAG, R110TAG)/pEcTyrRS/tRNACUA was blue on X-GAL, but the corresponding strain containing pA5/tRNACUA was not. Similarly, MaV203 containing pGADGAL4(T44TAG, R110TAG) and pEcTyrRS/tRNACUA grew robustly on plates with 3-AT concentrations up to 75 mM and on −Ura plates, but the corresponding strain containing pA5/tRNACUA failed to grow on 10 mM 3AT or in the absence of uracil. Taken together, these observations suggest that the EcTyrRS-dependent phenotypes of pGADGAL4(T44TAG, R110TAG) span the dynamic range of the *URA3, HIS3*, and *lacZ* reporters in MaV203. Due to the toxicity of the wild-type GAL4 expression construct, we were unable to compare the activity of the mutant proteins to that of the wild-type. However, previous experiments by Schimmel and coworkers suggest that the protein yield may be decreased somewhat by amber suppression [25]. Several factors may contribute to decreased protein yields in gene expression using amber suppression, including competition of translational release factor with tRNAs containing amber anticodons for binding to the mRNA termination codons [50], and nonsense-mediated mRNA decay [19]. Nonetheless, the efficiency of amber suppression in yeast appears to be sufficient to produce good yields of protein [25].

It was of interest to determine the transcriptional activation activity of GAL4 mutants in which T44 or R110 were substituted with amino acids other than tyrosine, since the ability to substitute varied amino acids without altering the activity of GAL4 is likely to be a key to selection of mutant aminoacyl-tRNA synthetases that can incorporate unnatural amino acids into proteins. A series of five mutants of residue T44 in GAL4 (T44Y, T44W, T44F, T44D, and T44K) were constructed in pGADGAL4(R110TAG), since pGADGAL4 is itself toxic. A similar series of mutants at position R110 in GAL4 (R110Y, R110W, R110F, R110D, and R110K) in pGADGAL4(T44TAG) was constructed. These mutants are biased toward the large hydrophobic amino acid side chains that we are interested in incorporating into proteins, but also contain a positively and negatively charged residue as a stringent test of permissiveness. Each mutant was cotransformed with pEcTyrRS/tRNACUA into MaV203 cells and Leu^+ Trp^+ isolates assayed for lacZ production by ortho-nitrophenyl-β-D-galactopyranoside (ONPG) hydrolysis (Figure 5). The variation in activity between cells containing GAL4 with different amino acids substituted for either T44 or R110 was less than 3-fold in all cases. This minimal variability demonstrates the permissiveness of these sites to amino acid substitution without altering the transcriptional activity of GAL4. As expected from the growth of the single amber mutants assayed on selective media, mutants of T44 made in the GAL4(R110TAG) background lead to slower rates of ONPG hydrolysis than mutants of R110 made in the GAL4(T44TAG) background. We cannot rule out that the incorporation of unnatural amino acids at both T44 and R110 would not affect the activity of GAL4.

Model enrichment studies were performed to examine the ability of the system to select an active synthetase from a large excess of inactive synthetases (Table 1A;

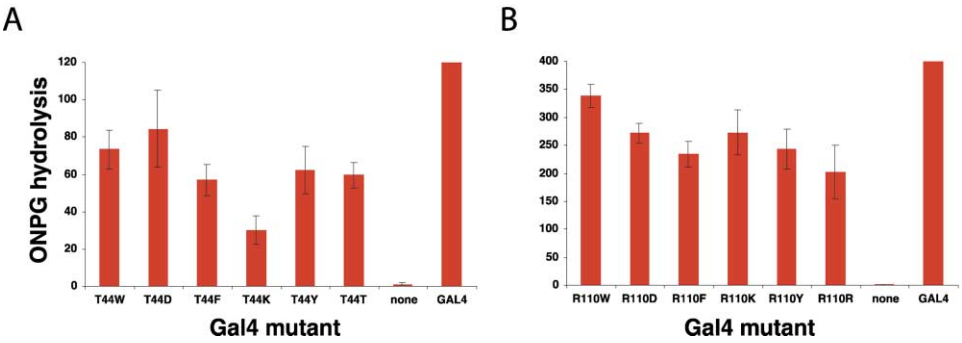

Figure 5. GAL4 Residues T44 and R110 Are Permissive Sites

(A), T44; (B), R110. ONPG hydrolysis units are as defined in Experimental Procedures. "GAL4" is MaV203 transformed with pCL1 and was offscale: 600 ONPG hydrolysis units; "none" is MaV203 transformed with plasmids encoding the GAL4 DBD and GAL4 AD separately.

Figure 6). This selection models the ability to select active synthetases from a library of variants in the presence of an unnatural amino acid. MaV203 cells containing GAL4(T44, R110) and EcTyrRS/$tRNA_{CUA}$ were mixed with a 10- to 10^6-fold excess of cells containing GAL4(T44TAG, R110TAG) and A5/$tRNA_{CUA}$ as judged by both OD_{660} and the fraction of colonies that turned blue when cells were plated on nonselective −Leu, −Trp media and assayed by X-GAL overlay. Those cells able to survive on 50 mM 3-AT or in the absence of uracil were selected. The ratio of cells surviving on 3-AT or −Ura that were blue in the X-GAL assay to those that were white, when compared to the same ratio in the absence of selection, clearly demonstrates that the positive selections can enrich active synthetases from inactive synthetases by a factor greater than 10^5 (Table 1A). Measurement of accurate enrichments for starting ratios greater than 1:10^5 was not possible, because single colonies could not be counted when more than 10^6 cells were plated.

After a positive selection in the presence of unnatural amino acid, the selected cells will contain synthetases able to use natural amino acids and, perhaps more rarely, able to use an added unnatural amino acid. To isolate those synthetases capable of using only the unnatural amino acid, cells encoding synthetases that use natural amino acids must be deleted from the selected clones. This can be accomplished with a negative selection in which the unnatural amino acid is withheld and those synthetases that function with a natural amino acid are removed. A model negative selection was performed in an analogous manner to the model positive selection. EcTyrRS/$tRNA_{CUA}$ was mixed with a 10- to 10^5-fold excess of A5/$tRNA_{CUA}$, and selection was performed on 0.1% 5-FOA. The ratio of cells surviving on 0.1% 5-FOA that were white in the X-GAL assay to those that were blue, compared to the same ratio under nonselective conditions (Table 1B), makes it clear that the negative selections can enrich inactive synthetases from active synthetases by a factor of at least 0.6×10^4. Measurement of accurate enrichments for starting ratios greater than 1:10^4 was not possible, because isolated single colonies could not be counted when more than 10^5 cells were plated.

Significance

We have developed a general approach that allows both positive selection of active aaRSs and negative selection of aaRSs that recognize natural amino acids. By varying the stringencies of the selection, it should be possible to isolate a variety of synthetase activities.

Table 1. Model Selections

A. Model Selections for Functional EcTyrRS

Starting Ratio, EcYRS:A5[a]	1:10	1:10^2	1:10^3		1:10^4		1:10^5	
Cell dilution	10^3	10^3	10^2	10^3	10	10^3	1	10^3
−Leu,Trp (#blue[b])	1360 (81)	1262 (0)	$>10^3$ (1)	1774 (0)	$>10^4$ (-)	1092 (0)	$>10^4$ (-)	1472 (0)
−Ura (#blue[b])	152 (152)	9 (9)	8 (8)	0 (-)	5 (5)	0 (-)	16 (14)	0 (-)
−His + 50 mM 3AT (#blue[b])	135 (135)	7 (7)	0 (-)	0 (-)	3 (3)	0 (-)	10 (10)	0 (-)
Enrichment factor	>10	$>10^2$	$>10^3$		$>10^4$		$>10^5$	

B. Model Selections for Nonfunctional EcTyrRS

StartingRatio, A5:EcYRS[a]	1:10	1:10^2	1:10^3	1:10^4	
Cell dilution	10^3	10^2	10^2	10^2	10
−Leu,Trp (#white[b])	353 (22)	1401 (31)	1336 (2)	1375 (0)	$>10^4$
0.1% 5-FOA (#white[b])	16 (16)	41 (41)	4 (4)	0 (-)	2 (2)
Enrichment factor	>10	>45	>600	$> 0.67 \times 10^4$	

[a] Determined by OD_{660}.
[b] On X-GAL.

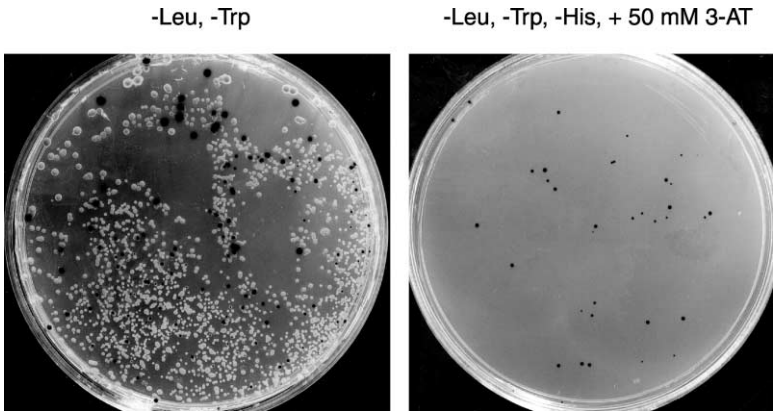

Figure 6. Selection of Active EcTyrRS Clones

MaV203 containing a 1:10 mixture of pEcTyrRS-tRNACUA:pA5-tRNACUA were plated at a 10^3 dilution on −Leu, −Trp plates (left) or −Leu, −Trp, −His + 50 mM 3-AT and processed using X-GAL overlay as described in Experimental Procedures.

Application of this method to a model selection using variants of EcTyrRS showed enrichments of greater than 10^5 in a single round of positive selection and greater than 0.6×10^4 in a single round of negative selection. These observations suggest that this method will provide rapid access to orthogonal aminoacyl-tRNA synthetases that function to site specifically incorporate unnatural amino acids with a diversity of side chains into proteins in *S. cerevisiae*. Moreover, it is likely that enzymes evolved in *S. cerevisiae* can be used in higher eukaryotes.

Experimental Procedures

Vector Construction

The *E. coli* tyrosyl $tRNA_{CUA}$ gene was amplified by PCR using the primers tRNA5′-GGGGGGACCGGTAAGCTTCCCGATAAGGGAGCAGGCCAGTAAAAAGCATTACCCCGTGGTGGGTTCCCGA and tRNA3′-GGCGGCGCTAGCGGGAAGTTCAGGGACTTTTGAAAAAAATGGTGGTGGGGGAAGGAT from pESCSU3URA (a gift from Thomas J. Magliery). This and all other PCR reactions were performed using the Expand PCR kit from Roche according to the manufacturer's instructions. After restriction endonuclease digestion with NheI and AgeI, this tRNA gene was inserted between the same sites in the 2 μm vector pESCTrp (Stratagene) to yield ptRNACUA. The full-length ADH1 promoter was amplified by PCR from pDBLeu (Invitrogen) with the primers PADHf, IGGGGGGACCGGTCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATC AAGG and pADHr, GGGGGGGAATTCAGTTGATTGTATGCTTGGTATAGCTTGAAATATTGTGCAGAAAAAGAAAC and digested with AgeI and EcoRI. EcTyrRS was amplified with the primers pESCTrp1, TCATAACGAGAATTCATGGCAAGCAGTAACTTG and pESCTrp2, TTACTACGTGCGGCCGCTTATTTCCAGCAAATCAGAC. The EcTyrRS PCR product were digested with EcoRI and NotI. ptRNACUA was then digested with AgeI and NotI. A triple ligation of these three DNAs yielded pEcTyrRS/tRNACUA. Plasmid pA5/tRNACUA, in which amino acid residues 37, 126, 182, 183, and 186 in the active site are mutated to alanine, was created by overlap PCR using the oligonucleotides F37Afwd, CCGATCGCGCTCGCTTGCGGCTTCGATC; N126Afwd, ATCGCGGCGAACGCCTATGACTGGTTC; 182,183,186A, GTTGCAGGGTTATGCCGCCGCCTGTGCGAACAAACAGTAC and their reverse complements, as well as the flanking oligonucleotides 4783, GCCGCTTTGCTATCAAGTATAAATAG; 3256, CAAGCCGACAACCTTGATTGG and pEcTyrRS/tRNACUA as a template. The PCR product was digested with EcoRI and NotI and ligated into the large fragment of pEcTyrRS/tRNACUA released upon digestion with the same enzymes.

To construct 1st generation DB-AD reporters, the GAL4 DNA binding domain was PCR amplified from pGADT7 (Clontech) using the forward primer pADfwd: GGGGACAAGTTTGTACAAAAAAGCAGGCTACGCCAATTTTAATCAAAGTGGGAATATTGC or pADfwd(TAG) GGGGACAAGTTTGTACAAAAAAGCAGGCTAGGCCAATTTTAATCAAAGTGGGAATATTGC and ADrev: GGGGACCACTTTGTACAAGAAAGCTGGGTTACTCTTTTTTTGGGTTTGGTGGGGTATC. These PCR products were cloned into the vector pDEST3-2 (invitrogen) using the Clonase procedure, according to the manufacturer's instructions, yielding pDB-AD and pDB-(TAG)-AD.

To construct PGADGAL4 and variants, the GAL4 gene was amplified from pCL1 (Clontech) by PCR using the primers ADH1428-1429, AAGCTATACCAAGCATACAATC and GAL4C, ACAAGGCCTTGCTAGCTTACTCTTTTTTTGGGTTTGGTGGGGTATCTTC. This fragment was cloned into the vector pCR2.1 TOPO (Invitrogen) according to the manufacturer's instructions. A clone containing the GAL4 gene (pCR2.1 TOPOGAL4) was digested with HindIII, and the 2.7 kb GAL4 fragment was gel purified and ligated to the large fragment of pGADT7 that had been digested with HindIII, treated with calf intestinal phosphotase, and gel purified. Variants of the GAL4 gene were created by Quikchange reactions (Stratagene) carried out according to the manufacturer's instructions on pCR2.1 TOPO GAL4 using primers listed in the Supplemental Data (http://www.chembiol.com/cgi/content/full/10/6/511/DC1). GAL4 mutants were cloned into pGADT7 in the same manner as the wild-type GAL4 gene. All final constructs were confirmed by DNA sequencing.

Yeast Media and Manipulations

S. cerevisiae strain MaV203 (Invitrogen) is MATα; *leu2-3,112; trp1-109*; *his3Δ200*; *ade2-101*; *cyh2^R; cyh1^R*; *GAL4Δ; gal80Δ*; *GAL1::lacZ; $HIS3_{UASGAL1}$::HIS3@LYS2; $SPAL10_{UASGAL1}$::URA3*. Yeast media were purchased from Clontech, 5-FOA and X-GAL were from Invitrogen, and 3-AT was from BIO 101. YPER (yeast protein extraction reagent) and ONPG were purchased from Pierce Chemicals. Plasmid transformations were performed by the PEG/Lithium actetate method [16], and transformants were selected on the appropriate synthetic complete dropout media.

To test the phenotypes conferred by various plasmid combinations on MaV203, yeast colonies from synthetic complete dropout plates of each transformation were resuspended in 15 μl of sterile water and streaked on the selective media of interest. Each phenotype was confirmed with at least five independent colonies.

X-GAL assays were performed by the agarose overlay method [51]. Briefly, colonies or cell patches were lysed on agar plates by several additions of neat chlorofom. After chloroform evaporation, 1% agarose containing 0.25 g/l of X-GAL and buffered with 0.1 M Na_2PO_4 (pH 7.0) was applied to the plate surface. Once the agarose was set, plates were incubated at 37°C for 12 hr.

For ONPG assays, 1 ml of SD −Leu, −Trp in a 96-well block was inoculated with a single colony and incubated at 30°C with shaking. The OD_{660} of 100 μl of cells and several dilutions of cells were recorded in parallel in a 96-well microtiter plate. Cells (100 μl) were mixed with 100 μl of YPER:ONPG (1× PBS, 50% v/v YPER, 20 mM $MgCl_2$, 0.25% v/v β-mercaptoethanol, and 3 mM ONPG) and incubated with shaking at 37°C. Upon color development, cells were pelleted by centrifugation, the supernatant was transferred to a clean 96-well microtiter plate [51] (Nunclon, catalog no. 167008), and the A_{420} was recorded. All data shown are the mean of trials

from at least four independent clones, and the error bars shown represent the standard deviation. ONPG hydrolysis was calculated using the equation β-galactosidase units = $1000 \cdot A_{420}/(V.t.OD_{660})$, where V is the volume of cells in milliliters, t is the time of incubation in minutes [51]. One β-galactosidase unit corresponds to the hydrolysis of 1 μmol of ONPG per min per cell [51]. Spectrophotometric readings were performed on a SPECTRAmax190 plate reader.

Model Selections

Positive Selections

Two overnight cultures were grown in SD −Leu, −Trp. One contained MaV203 harboring pEcTyrRS/tRNACUA and pGADGAL4(T44, R110TAG), and the other pA5-tRNACUA and pGADGAL4(T44, R110TAG). These cells were harvested by centrifugation and resuspended in 0.9% NaCl by vortexing. The two cell solutions were then diluted to identical OD_{660}s. MaV203 harboring pEcTyrRS/tRNACUA and pGADGAL4(T44,R110TAG) were serially diluted over 7 orders of magnitude and each dilution was then mixed 1:1 vol:vol with undiluted MaV203 harboring pA5/tRNACUA and pGADGAL4(T44, R110TAG) to afford defined ratios of cells containing active and inactive tyrosyl-tRNA synthetase. For each ratio, a second serial dilution was performed in which the concentration of cells was decreased but the ratio of cells harboring pEcTyrRS/tRNACUA and pGADGAL4(T44,R110TAG) to cells harboring pA5/tRNACUA and pGADGAL4(T44,R110TAG) was maintained. These dilutions were plated on SD −Leu, −Trp; SD −Leu, −Trp,−Ura; and SD −Leu, −Trp, −His + 50 mM 3-AT. After 60 hr, the number of colonies on each plate was counted using an Eagle Eye CCD camera (Stratagene), and the phenotype of survivors was confirmed with a X-GAL β-galactosidase assay. Cells from several individual blue or white colonies were isolated and grown to saturation in SD −Leu, −Trp and the plasmid DNA isolated by standard methods. The identity of the EcTyrRS variant was confirmed by DNA sequencing.

Negative Selection

The model negative selection was performed in an analogous manner to the positive selection, except that MaV203 harboring pA5/tRNACUA and pGADGAL4(T44,R110TAG) were serially diluted and mixed with a fixed density of MaV203 harboring pEcTyrRS/tRNACUA and pGADGAL4(T44,R110TAG). Cells were plated on SD −Leu, −Trp + 0.1% 5-FOA, the number of colonies was counted after 48 hr, and the plates were processed as described above.

Acknowledgments

J.W.C. was a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (grant DRG 1707-02). This work was supported by the National Institute of Health (1 R01 GM 62159-02) and the Department of Energy (DE-FG0300ER45812). J.W.C. thanks Dr. Stanley Fields (U. Washington) for information on pCL1 and Dr. Aseem Ansari (UW Madison) for helpful discussions. We thank Drs. Thomas J. Magliery and Stephen W. Santoro for materials. This is manuscript 15527-CH of The Scripps Research Institute.

Received: January 22, 2003
Revised: April 18, 2003
Accepted: April 28, 2003
Published: June 20, 2003

References

1. Chin, J.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. (2002). Addition of a photocrosslinker to the genetic code of Escherichia coli. Proc. Natl. Acad. Sci. USA *99*, 11020–11024.
2. Chin, J.W., Santoro, S.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. (2002). Addition of p-Azido-L-phenylalanine to the genetic code of Escherichia coli. J. Am. Chem. Soc. *124*, 9026–9027.
3. Chin, J.W., and Schultz, P.G. (2002). Invivo photocrosslinking with unnatural amino acid mutagenesis. Chembiochem *11*, 1135–1137.
4. Zhang, Z., Wang, L., Brock, A., and Schultz, P.G. (2002). The selective incorporation of alkenes into proteins in Escherichia coli. Angew. Chem. Int. Ed. Engl. *41*, 2840–2842.
5. Wang, L., Zhang, Z., Brock, A., and Schultz, P.G. (2003). Addition of a keto functional group to the genetic code of *Escherichia coli*. Proc. Natl. Acad. Sci. USA *100*, 56–61.
6. Wang, L., Brock, A., and Schultz, P.G. (2002). Adding L-3-(2-naphthyl)alanine to the genetic code of E-coli. J. Am. Chem. Soc. *124*, 1836–1837.
7. Nowak, M.W., Gallivan, J.P., Silverman, S.K., Labarca, C.G., Dougherty, D.A., and Lester, H.A. (1998). In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. Methods Enzymol. *293*, 504–529.
8. Saks, M.E., Sampson, J.R., Nowak, M.W., Kearney, P.C., Du, F., Abelson, J.N., Lester, H.A., and Dougherty, D.A. (1996). An engineered Tetrahymena tRNAGln for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression. J. Biol. Chem. *271*, 23169–23175.
9. Dougherty, D.A. (2000). Unnatural amino acids as probes of protein structure and function. Curr. Opin. Chem. Biol. *4*, 645–652.
10. Drabkin, H.J., Park, H.J., and RajBhandary, U.L. (1996). Amber suppression in mammalian cells dependent upon expression of an Escherichia coli aminoacyl-tRNA synthetase gene. Mol. Cell. Biol. *16*, 907–913.
11. Kowal, A.K., Kohrer, C., and RajBhandary, U.L. (2001). Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. Proc. Natl. Acad. Sci. USA *98*, 2268–2273.
12. Sakamoto, K., Hayashi, A., Sakamoto, A., Kiga, D., Nakayama, H., Soma, A., Kobayashi, T., Kitabatake, M., Takio, K., Saito, K., et al. (2002). Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. Nucleic Acids Res. *30*, 4692–4699.
13. Liu, D.R., and Schultz, P.G. (1999). Progress toward the evolution of an organism with an expanded genetic code. Proc. Natl. Acad. Sci. USA *96*, 4780–4785.
14. Furter, R. (1998). Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. Protein Sci. *7*, 419–426.
15. Liu, D.R., Magliery, T.J., Pasternak, M., and Schultz, P.G. (1997). Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. Proc. Natl. Acad. Sci. USA *94*, 10092–10097.
16. Burke, D., Dawson, D., and Stearns, T. (2000). Methods in Yeast Genetics. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
17. Translational Control, J.W.B. Hershey, M.B. Matthews, and N. Sonenberg, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
18. Kwok, Y., and Wong, J.T. (1980). Evolutionary relationship between Halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. Can. J. Biochem. *58*, 213–218.
19. Gonzalez, C.I., Wang, W., and Peltz, S.W. (2001). Nonsense-mediated mRNA decay in Saccharomyces cerevisiae: a quality control mechanism that degrades transcripts harboring premature termination codons. In The Ribosome (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 321–328.
20. Kohrer, C., Xie, L., Kellerer, S., Varshney, U., and RajBhandary, U.L. (2001). Import of amber and ochre suppressor tRNAs into mammalian cells: a general approach to site-specific insertion of amino acid analogues into proteins. Proc. Natl. Acad. Sci. USA *98*, 14310–14315.
21. Buvoli, M., Buvoli, A., and Leinwand, L.A. (2000). Suppression of nonsense mutations in cell culture and mice by multimerized suppressor tRNA genes. Mol. Cell. Biol. *20*, 3116–3124.
22. Wang, L., and Schultz, P.G. (2002). Expanding the genetic code. Chem. Comm. *7*, 1–11.
23. Fechter, P., Rudinger-Thirion, J., Tukalo, M., and Giege, R. (2001). Major tyrosine identity determinants in Methanococcus jannaschii and Saccharomyces cerevisiae tRNA(Tyr) are conserved but expressed differently. Eur. J. Biochem. *268*, 761–767.
24. Edwards, H., and Schimmel, P. (1990). A bacterial amber suppressor in Saccharomyces cerevisiae is selectively recognized

by a bacterial aminoacyl-tRNA synthetase. Mol. Cell. Biol. *10*, 1633–1641.
25. Edwards, H., Trezeguet, V., and Schimmel, P. (1991). An Escherichia coli tyrosine transfer RNA is a leucine-specific transfer RNA in the yeast Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *88*, 1153–1156.
26. Wakasugi, K., Quinn, C.L., Tao, N., and Schimmel, P. (1998). Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. EMBO J. *17*, 297–305.
27. Doctor, B.P., Loebel, J.E., and Kellogg, D.A. (1966). Studies on the species specificity of yeast and E. coli tyrosine tRNAs. Cold Spring Harb. Symp. Quant. Biol. *31*, 543–548.
28. Hamano-Takaku, F., Iwama, T., Saito-Yano, S., Takaku, K., Monden, Y., Kitabatake, M., Soll, D., and Nishimura, S. (2000). A mutant Escherichia coli tyrosyl-tRNA synthetase utilizes the unnatural amino acid azatyrosine more efficiently than tyrosine. J. Biol. Chem. *275*, 40324–40328.
29. Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., et al. (2002). An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. Proc. Natl. Acad. Sci. USA *99*, 9715–9720.
30. Wang, L., Brock, A., Herberich, B., and Schultz, P.G. (2001). Expanding the genetic code of Escherichia coli. Science *292*, 498–500.
31. Laughon, A., Driscoll, R., Wills, N., and Gesteland, R.F. (1984). Identification of two proteins encoded by the Saccharomyces cerevisiae GAL4 gene. Mol. Cell. Biol. *4*, 268–275.
32. Laughon, A., and Gesteland, R.F. (1984). Primary structure of the Saccharomyces cerevisiae GAL4 gene. Mol. Cell. Biol. *4*, 260–267.
33. Keegan, L., Gill, G., and Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science *231*, 699–704.
34. Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683–689.
35. Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F., and Ptashne, M. (1989). An amino-terminal fragment of GAL4 binds DNA as a dimer. J. Mol. Biol. *209*, 423–432.
36. Giniger, E., Varnum, S.M., and Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell *40*, 767–774.
37. Ma, J., and Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. Cell *48*, 847–853.
38. Ma, J., and Ptashne, M. (1987). The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. Cell *50*, 137–142.
39. Kishore, G.M., and Shah, D.M. (1988). Amino acid biosynthesis inhibitors as herbicides. Annu. Rev. Biochem. *57*, 627–663.
40. DeMaggio, A.J., Goldman, P., Shih, H.M., Goodman, R.H., and Hoekstra, M.F. (2000). The yeast split-hybrid system. Methods Enzymol. *328*, 128–137.
41. Shih, H.M., Goldman, P.S., DeMaggio, A.J., Hollenberg, S.M., Goodman, R.H., and Hoekstra, M.F. (1996). A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP. Proc. Natl. Acad. Sci. USA *93*, 13896–13901.
42. Vidal, M., Braun, P., Chen, E., Boeke, J.D., and Harlow, E. (1996). Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. Proc. Natl. Acad. Sci. USA *93*, 10321–10326.
43. Vidal, M., Brachmann, R.K., Fattaey, A., Harlow, E., and Boeke, J.D. (1996). Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. Proc. Natl. Acad. Sci. USA *93*, 10315–10320.
44. Boeke, J.D., LaCroute, F., and Fink, G.R. (1984). A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. *197*, 345–346.
45. Huang, J., and Schreiber, S.L. (1997). A yeast genetic system for selecting small molecule inhibitors of protein-protein interactions in nanodroplets. Proc. Natl. Acad. Sci. USA *94*, 13396–13401.
46. Johnston, M., and Dover, J. (1988). Mutational analysis of the GAL4-encoded transcriptional activator protein of Saccharomyces cerevisiae. Genetics *120*, 63–74.
47. Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. Nature *356*, 408–414.
48. Baleja, J.D., Marmorstein, R., Harrison, S.C., and Wagner, G. (1992). Solution structure of the DNA-binding domain of Cd2–GAL4 from S. cerevisiae. Nature *356*, 450–453.
49. Hidalgo, P., Ansari, A.Z., Schmidt, P., Hare, B., Simkovich, N., Farrell, S., Shin, E.J., Ptashne, M., and Wagner, G. (2001). Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. Genes Dev. *15*, 1007–1020.
50. Nakamura, Y., Uno, M., Toyoda, T., Fujiwara, T., and Ito, K. (2001). Protein tRNA mimicry in translation termination. In The Ribosome (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 469-475.
51. Serebriiskii, I.G., and Golemis, E.A. (2000). Uses of lacZ to study gene function: evaluation of beta-galactosidase assays employed in the yeast two-hybrid system. Anal. Biochem. *285*, 1–15.