Review TRENDS in Genetics Vol.20 No.12 December 2004

An expanding genetic code

T. Ashton Cropp[1] and Peter G. Schultz[2]

[1]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA
[2]Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

More than 30 novel amino acids have been genetically encoded in response to unique triplet and quadruplet codons including fluorescent, photoreactive and redox active amino acids, glycosylated and heavy atom derived amino acids in addition to those with keto, azido and acetylenic chains. In this article, we describe recent advances that make it possible to add new building blocks systematically to the genetic codes of bacteria, yeast and mammalian cells. Taken together these tools will enable the detailed investigation of protein structure and function, which is not possible with conventional mutagenesis. Moreover, by lifting the constraints of the existing 20-amino-acid code, it should be possible to generate proteins and perhaps entire organisms with new or enhanced properties.

With the rare exceptions of selenocysteine [1] and pyrrolysine [2] the genetic codes of all known organisms specify the same 20 amino-acid building blocks. These building blocks contain a limited number of functional groups including carboxylic acids and amides, a thiol, a thiol ether, alcohols, basic amines, alkyl and aryl groups. Although various arguments have been put forth to explain the nature and the number of amino acids in the code, it is clear that proteins require additional chemistries to perform their natural functions. These groups are provided through post-translational modifications including phosphorylation, glycosylation, hydroxylation and cofactors such as flavins, pyridoxal and metal ions. The need for additional factors in so many protein-mediated processes suggests that although a 20-amino-acid code is sufficient for life, it might not be optimal. Consequently, the development of methods that enabled the systematic expansion of the genetic code might enable the evolution of proteins, or even entire organisms, with new or enhanced properties. Moreover, such a method would dramatically increase our ability to manipulate rationally protein structure and function *in vitro* and *in vivo*. This in turn would enable a more classic, chemical approach to the study of proteins, in which carefully defined changes in the steric or electronic properties of an amino acid are correlated with the changes in protein structure and function. In this article, we review recent advances that have made it possible to add new amino acids to the genetic codes of prokaryotic and eukaryotic organisms.

Corresponding author: Peter G. Schultz (Schultz@scripps.edu).
Available online 13 October 2004

Reassignment of nonsense codons with orthogonal tRNA-synthetase pairs

To encode additional amino acids in an organism, one requires codons that uniquely specify the unnatural amino acids of interest. The natural genetic code consists of 64 unique codons comprising all possible three-nucleotide permutations. Sixty one encode the 20 common amino acids; the remaining three, TAG, TAA and TGA, lack a corresponding tRNA and instead bind to termination factors and act as stop codons to end protein translation. The amber stop codon (TAG) is the least used of these in *Escherichia coli* and in yeast, and is therefore the most obvious candidate for reassignment to an unnatural amino acid. In addition, some strains of *E. coli* contain naturally occurring amber-suppressor tRNAs that are substrates for endogenous aminoacyl-tRNA synthetases and efficiently deliver a natural amino acid in response to TAG. Similar suppressor tRNAs have been discovered for higher organisms [3,4], and, in general, amber suppression is well tolerated and does not affect host growth rates significantly.

The use of an amber codon to encode unnatural amino acids requires the generation of a new set of translational components, namely a suppressor tRNA and a cognate aminoacyl-tRNA synthetase. In order for an amber-suppressor tRNA to incorporate an unnatural amino acid efficiently and with high fidelity, the tRNA must be expressed and processed into a functional form, but it can not be aminoacylated by any of the host tRNA synthetases (hereafter called synthetases) with endogenous amino acids. Similarly, the cognate synthetase must not transfer amino acids to host tRNAs, otherwise the unnatural amino acid will be misincorporated throughout the proteome. Such an independent tRNA-synthetase pair is said to be orthogonal to the host. The substrate specificity of the orthogonal synthetase must then be evolved to selectively charge the unnatural amino acid and not any endogenous amino acids. An important requirement for such a system is that the unnatural amino acid of interest must not be toxic to the cells, and must be efficiently transported from the growth medium into the cytoplasm or be biosynthesized by the organism.

The most successful strategy to date for generating orthogonal tRNA-synthetase pairs is to import both of them from another organism that contains different tRNA-synthetase-recognition elements. Eukaryotic and archeabacterial synthetases generally do not charge *E. coli* tRNAs [5] and the same is true for bacterial synthetases in a eukaryotic host. Among the orthogonal

 0168-9525/$ - see front matter doi:10.1016/j.tig.2004.09.013

pairs first created for use in *E. coli*, the most efficient was the tyrosyl pair that was derived from the archeabacterium *Methanococcus jannaschii* [5,6]. This pair is expressed in an active form and inserts tyrosine in response to the amber nonsense codon in *E. coli*. However, the tRNA is recognized to some degree by the endogenous *E. coli* synthetases. To further increase the orthogonality of this tRNA-synthetase pair, the tRNA coding sequence was randomized partially using PCR-based mutagenesis and, in the absence of the *Methanococcus jannaschii* tyrosyl-tRNA synthetase (MjYRS), the resulting library was passed through a negative selection in *E. coli*, which was based on suppression of in-frame amber stop codons in the toxic gene that encodes barnase. Any tRNAs that can be aminoacylated by endogenous synthetases are thus removed. Survivors were then positively selected based on the suppression of amber mutations in the gene encoding chloramphenicol resistance (CmR) on media containing chloramphenicol [7,8]. Any surviving clones can be aminoacylated by the cognate archeal synthetase and function efficiently in translation. The result was a new tRNA that is a poorer substrate for the endogenous *E. coli* synthetases, yet retains the original activity with the cognate MjYRS and its ability to function in translation. Recently, additional orthogonal leucyl, lysyl and glutamyl tRNA-synthetase pairs have been developed for use in *E. coli* using a 'consensus suppressor strategy' based on conserved homology among all archeal tRNAs [9]. These additional pairs might represent a better starting point for inserting unnatural amino acids that are structurally different from tyrosine and might possibly function using alternative codons (opal, TGA or four-base AGGA) to be used in combination with the existing tyrosyl amber-suppressor pair.

Although archeal tRNA-synthetase pairs can be used to encode novel amino acids in *E. coli*, they are not orthogonal in eukaryotes. To encode unnatural amino acids genetically in higher organisms, additional orthogonal pairs are needed. Because the translational components of eukaryotes are well conserved, it should be possible to develop tRNA-synthetase pairs in S*accharomyces cerevisiae* that can be later transferred to more-complex organisms. Research to date has therefore largely focused on developing bacterial pairs that, based on their distinct tRNA-synthetase-identity elements, are orthogonal in yeast. It has been shown that the *E. coli* tyrosyl-tRNA synthetase (EcYRS) does not aminoacylate yeast tRNAs [10]. Furthermore, a derivative of the cognate *E. coli* tyrosyl-$tRNA_{CUA}$ can be expressed in yeast, modified to an active form and function as an amber suppressor [11,12]. Moreover, we have shown that this tRNA is not aminoacylated by endogenous yeast synthetases, demonstrating that the *E. coli* tyrosyl-tRNA synthetase functions as an orthogonal pair. In addition to the tyrosine pair, a variant of the *E. coli* leucyl tRNA-synthetase pair has recently been generated that is orthogonal to yeast and can efficiently deliver leucine in response to an amber codon.

Evolution of tRNA-synthetase substrate specificity

A systematic approach was developed to evolve synthetase mutants with specificities for a desired unnatural amino acid, and none for the endogenous amino acids. This scheme involves directed evolution experiments, in which large numbers of synthetase variants are passed through a series of stringent positive and negative genetic selection steps. Based on the crystal structure of the homologous *Bacillus stearathermophilus* tyrosyl-tRNA synthetase [13], a MjYRS-active-site library was designed in which all residues interacting with the bound tyrosine substrate were randomized [14]. This library, which consisted of $>10^8$ mutants, was then passed through a 'double-sieve' selection that consists of both positive and negative enrichment steps (Figure 1). To isolate synthetase variants that are capable of charging the unnatural amino acid, the mutant synthetases are first transferred into cells containing the *CmR* with amber mutations at sites permissive to mutation, and grown in media containing chloramphenicol and the unnatural amino acid. The synthetase clones surviving this first round of selection are then passed into a strain containing an amber mutant of the toxic barnase gene and grown in the absence of the unnatural amino acid. All clones that continue to charge endogenous amino acids will produce full-length barnase protein and will die, resulting in the enrichment of only those synthetases that are specific for the desired unnatural amino acid.

An analogous yeast-selection system has also been developed to evolve tRNA-synthetase specificities based on the suppression of two permisive amber codons in the gene encoding the transcriptional activator GAL4 [15–17]. Suppression of these two stop codons by a functional tRNA-synthetase pair results in the translation of full-length GAL4 protein, which activates transcription of the *HIS3, URA3* and *lacZ* reporter genes that have been inserted into the chromosome of the yeast two-hybrid strain MaV203 (available from Invitrogen; http://www.invitrogen.com) (Figure 1). Positive selection of an active-site library of the EcYRS (an orthogonal pair in yeast) in the presence of an unnatural amino acid is accomplished by plating on agar lacking histidine or uracil (these auxotrophies are complemented by the activated transcription of *HIS3* and *URA3*, respectively). Conveniently, negative selections that remove synthetases that are not specific for the unnatural amino acid are accomplished by plating the remnants of the first selection on agar that lacks the unnatural amino acid but contains protoxin 5-fluoroorotic acid (5-FOA). This compound is converted to a toxic product by the *URA3* gene product.

Adding amino acids to the genetic code

Using the double-sieve selections described previously, it has been possible to evolve the substrate specificities of the *M. jannaschii* and *E. coli* tyrosyl- and *E. coli* leucyl-tRNA synthetases to accept a wide variety of unnatural amino acids (Figure 2). These amino acids are usually incorporated with high translational fidelity and in good yields (i.e. 20–50% of wild-type protein). Some of these amino acids provide novel biophysical probes of protein structure and function both *in vitro* and *in vivo*. For example, two photocrosslinking amino acids, *p*-azido-L-phenylalanine [18] and *p*-benzoyl-L-phenylalanine [19], have been genetically encoded in both prokaryotes and

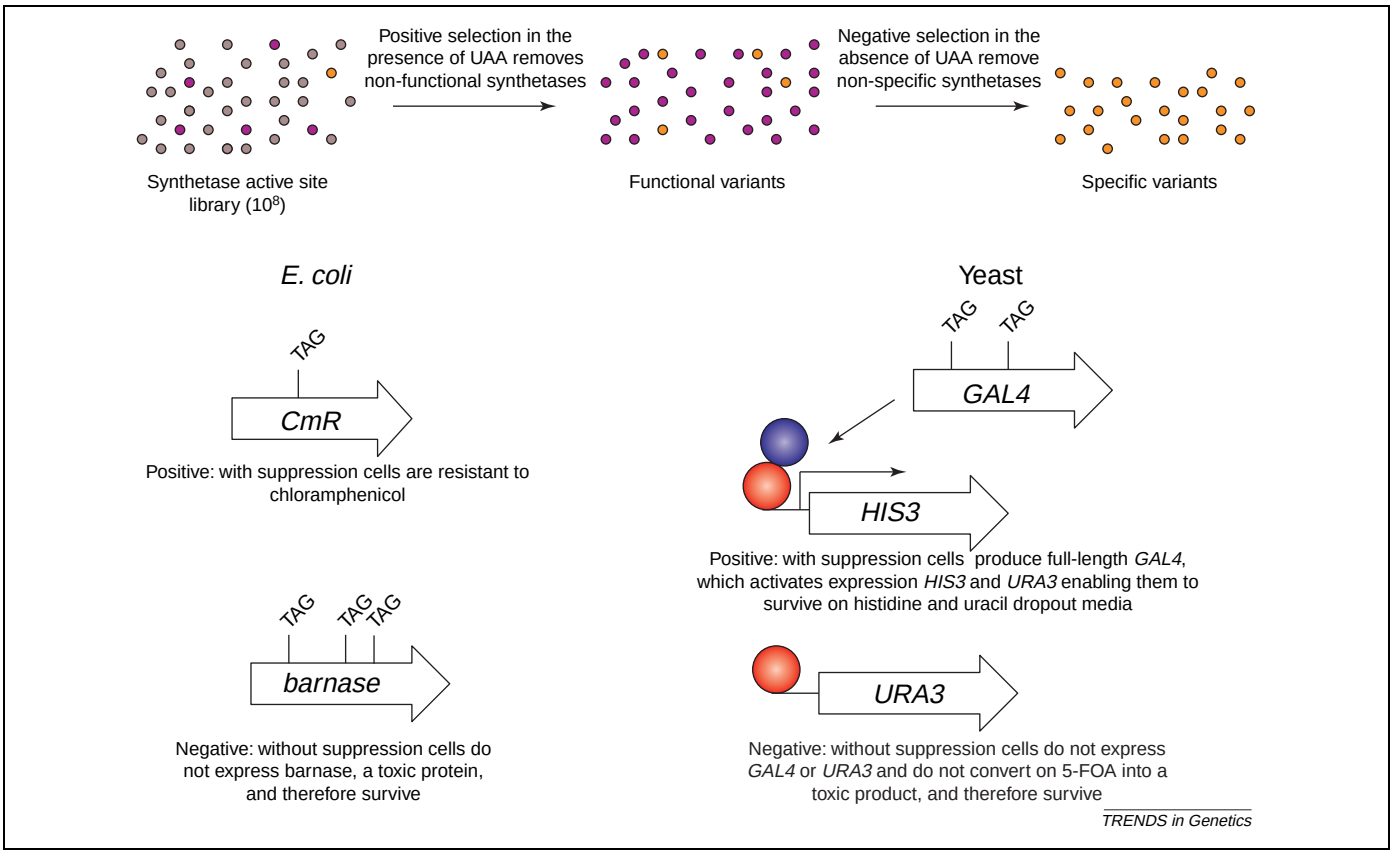

Figure 1. Double-sieve selection for aminoacyl-tRNA-synthetase substrate specificity. The synthetase libraries are first passed through a positive selection with the unnatural amino acid to remove non-functional variants, leaving active enzymes that charge the tRNA with both natural and unnatural amino acids. Those synthetases that accept natural amino acids are removed by negative selection in the absence of the unnatural amino acid. Abbreviation: 5-FOA, protoxin 5-fluoroorotic acid.

eukaryotes, enabling one to covalently capture protein–protein or protein–nucleic acid interactions *in vitro* and *in vivo* [20]. Amino acids with photo-isomerizable or photo-cleavable side chains have also been selectively incorporated into proteins and make it possible to photo-regulate protein activity *in vivo* [21]. Fluorescent amino acids with dansyl- and 7-hydroxycoumarin- side chains have also been genetically encoded in bacteria and yeast and can be used to probe conformational changes, protein–ligand interactions and protein localization, by either environmental effects on fluorescence or fluorescence resonance energy transfer (FRET) (P.G. Schultz, unpublished). These small fluorescent amino acids provide more control over the location of the fluorophore than green fluorescent protein (GFP) and are likely to cause less perturbation of protein function. Finally, the heavy atom-containing *p*-iodophenylalanine has been used for single wavelength anomalous dispersion (SAD), phasing in protein-structure determination [22].

Unnatural amino acids such as *p*-acetylphenylalanine [23], *m*-acetylphenylalanine [24], *p*-propargyloxyphenylalanine and *p*-azidophenylalanine, all of which contain functional groups that are not known to be present in any natural protein, have also been genetically encoded in both bacteria and yeast. These amino acids can be used to selectively modify proteins with nonpeptidic molecules including dyes, drugs, polymers and carbohydrates. The ketone side-chain selectively reacts with both hydrazides and hydroxylamine derivatives to form hydrazone and oxime linkages, respectively, which are stable under physiological conditions. The azide and alkyne functionalities of *p*-azidophenylalanine and *p*-propargyloxyphenylalanine can be selectively modified in copper(I) catalyzed [3+2] cycloaddition reactions with the corresponding alkyne and azide derivatives, respectively [25]. These chemistries have been used to modify proteins with fluorophores, tags and polyethylene glycols (PEG) with exquisite specificity and in some cases in living cells, and will probably be useful for the generation of diagnostic and therapeutic proteins.

It has also been possible to genetically encode products of common eukaryotic post-translational modifications in bacteria. For example, the glycosylated amino acids β-GlcNAc-serine [26] and α-GalNAc-threonine (P.G. Schultz, unpublished) have been introduced into proteins in response to the amber codon and, subsequently, elaborated into more-complex carbohydrates using glycosyl transferases. This methodology enables the synthesis of homogenous glycoproteins, which can be used both to probe the role of protein glycosylation in cellular function and to make selectively glycosylated proteins for therapeutic use. The redox-active amino acid dihydroxyphenylalanine (which results from the post-translational oxidation of tyrosine) has also been selectively incorporated into proteins, and provides a means to monitor and to control electron transfer in proteins [27]. Clearly, the

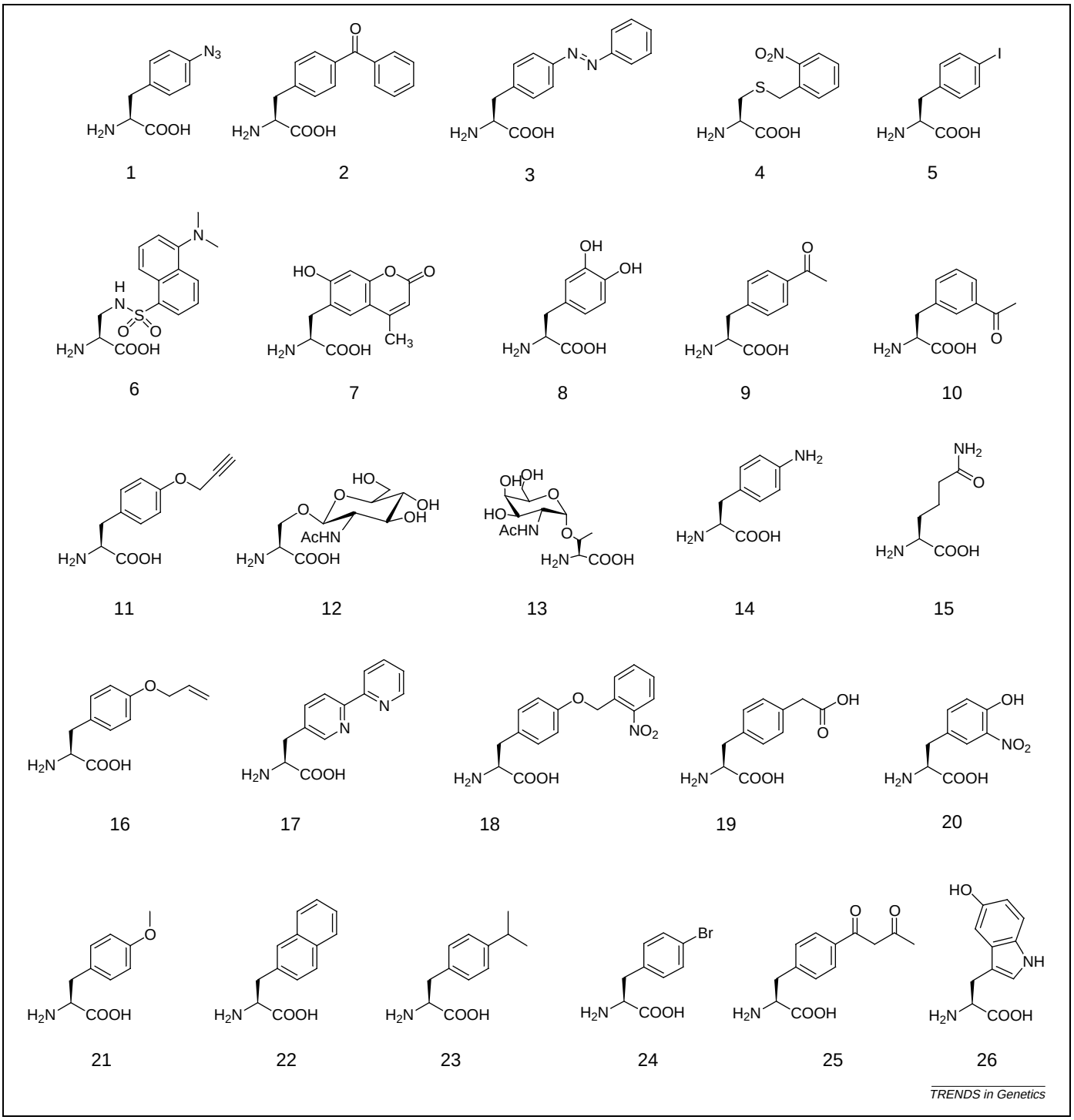

Figure 2. Examples of unnatural amino acids added to the genetic code. These include: 1, *p*-azido-L-phenylalanine, 2, benzoyl-L-phenylalanine; 3, amino acids with photo-isomerizable side chains; 4 and 18, amino acids with photo-cleavable side chains; 5, heavy atom-containing *p*-iodophenylalanine; 6, amino acids with dansyl- side chains; 7, amino acids with 7-hydroxycoumarin side chains; 8, the redox-active amino acid dihydroxyphenylalanine; 9, *p*-acetylphenylalanine; 10, *m*-acetylphenylalanine; 11, *p*-propargyloxyphenylalanine; 12, β-GlcNac-serine; 13, α-GalNAc-threonine; 14, *p*-aminophenylalanine; 15, homo-glutamine; 16, *O*-allyltyrosine; 17, a metal-binding amino acid; 20, *m*-nitrotyrosine; 21, *O*-methyltyrosine; 22, napthylalanine; 23 *p*-isopropylphenlalanine; 24, *p*-bromophenylalanine; 25, *p*-acetoacetylphenylalanine; 26, 5-hydroxytrptophan.

number and structural diversity of additional amino acids that have been genetically encoded in *E. coli* and yeast suggests that it will be possible to add many more novel amino acids to the code, including amino acids with modified backbones, metal-binding amino acids and spin-labeled amino acids.

Engineering more-complex organisms

In the experiments described previously, the unnatural amino acids are added exogenously to the growth media. Recently, a completely autonomous version of *E. coli* was generated with the ability to biosynthesize *p*-aminophenylalanine from common precursors and to selectively

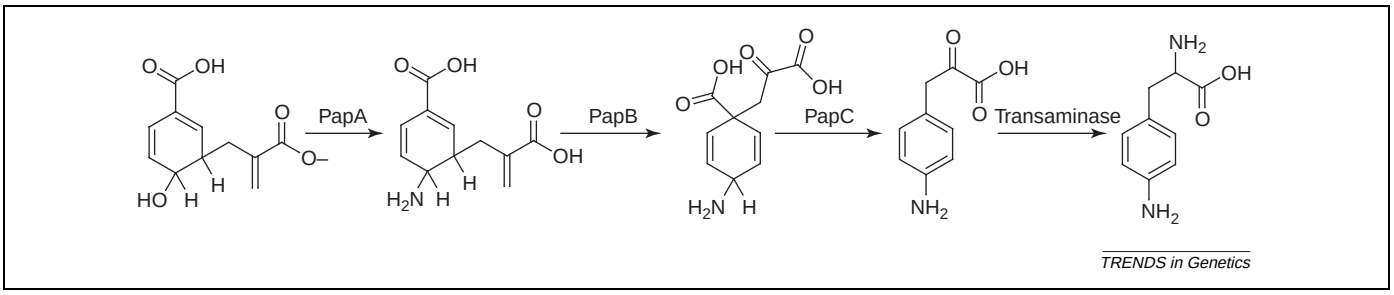

Figure 3. Biosynthesis of *p*-aminophenylalanine by a completely autonomous *Escherichia coli*. This strain is capable of producing an unnatural amino acid and inserting it into proteins in response to the amber nonsense codon, TAG.

insert it into proteins in response to the TAG codon [28]. This strain contains an orthogonal tRNA-synthetase pair that is specific for *p*-aminophenylalanine, in addition to a biosynthetic gene cluster from *Streptomyces venzualae* that converts chorismate to *p*-aminophenylalanine through a four-step pathway catalyzed by PapA, PapB, PapC and a host transaminase (Figure 3). This bacterium presents an excellent opportunity to study the evolvability of organisms with a 21-amino-acid code, and to determine whether additional amino acids provide any fitness advantage when bacteria are placed under selective pressures.

It is also possible to encode unnatural amino acids using four base codons [29]. Previously, it has been shown that natural frameshift suppressors function efficiently both *in vivo* and *in vitro* [30]. To determine which of the possible quadruplet codons might be used to encode unnatural amino acids in *E. coli*, a β-lactamase reporter containing an in-frame randomized four-base codon was developed. This library of 256 different reporter genes was co-transformed with a corresponding $tRNA^{Ser}$ library containing a randomized and expanded anticodon loop. Selection of the two libraries by growth on agar containing ampicillin revealed that several four-base codons (AGGA, CUAG, UAGA and CCCU) can be suppressed efficiently by the corresponding suppressor tRNAs [31]. In general, these codons are four-base derivatives of either nonsense or rare codons leading to reduced competition of the frameshift suppressor tRNA with natural tRNAs. Recently, an orthogonal AGGA suppressor lysyl tRNA-synthetase pair that was derived from the archean *Pyrococcus horikoshii* was developed, and the synthetase adapted to selectively charge the unnatural amino acid homo-glutamine. This pair, in combination with the previously developed amber-suppressor tyrosyl pair, has been used to encode two different unnatural amino acids in *E. coli* simultaneously. This experiment suggests that neither the number of available triplet codons nor the translational machinery represents a significant barrier to further expansion of the genetic code.

Thus far the methods for evolving new tRNA-synthetase pairs have focused on *E. coli* and yeast. Although these are useful model organisms, it should also be possible to evolve systematically synthetase variants that are capable of incorporating a wide variety of unnatural amino acids in yeast, and then transfer them directly into mammalian cells. Recently, it has been shown that the *E. coli* tyrosyl-tRNA-synthetase pair can be used to incorporate 3-iodotyrosine into proteins in mammalian cells [32]. Similarly, we have shown that a tyrosyl-$tRNA_{CUA}$ derived from *Bacillus stearothermophilus* is expressed efficiently in mammalian cells and can be used with the EcYRS as a functional orthogonal pair in a variety of mammalian cell lines. Indeed, in addition to the *B. stearathermophilus* tyrosyl-$tRNA_{CUA}$, the EcYRS synthetases evolved in yeast to accept then amino acids *p*-azidophenylalanine, benzoyl-L-phenylalanine, *p*-iodophenylalanine, *p*-acetylphenylalanine; *p*-methoxyphenylalanine can suppress an amber stop codon in the gene encoding firefly luciferase in mammalian 293T cells (T.A Cropp and P.G Schultz, unpublished). In addition, the *E. coli* leucyl tRNA-synthetase pair, which has been used to incorporate amino acids with photo-cleavable- or with dansyl side chains in yeast, is most likely to be orthogonal in mammalian cells and might also be transferred in a similar fashion as the *E. coli* tyrosyl pair. An alternative orthogonal pair for use in mammalian cells is based on the *Bacillus subtilis* tryptophanyl tRNA-synthetase pair. A rationally designed point mutant of this synthetase has been used to insert 5-hydroxytryptophan into proteins in mammalian cells in response to the opal nonsense codon, TGA [33]. This amino acid can be electrochemically oxidized to generate an efficient protein–protein crosslinking agent. Again it should be possible to evolve the substrate specificity of this orthogonal tryptophan pair to accommodate other amino acids by transfering it into the yeast-selection system described previously.

Concluding remarks and future directions

Future work will focus on expanding the nature and number of amino acids that can be genetically encoded in both prokaryotic and eukaryotic organisms, including multicellular organisms. Additional orthogonal pairs that suppress three- and four-base-containing codons are being developed. It might even be possible to delete rare redundant codons from the *E. coli* genome and use them instead to encode unnatural amino acids. More mutations, in addition to those in the orthogonal tRNA, synthetase and additional host strains, are likely to lead to further increases in suppression efficiencies and protein yields. The ability to encode unnatural amino acids genetically (e.g. with photoreactive, fluorescent, chemically reactive, isotopically labeled, metal binding or post-traslationally modified side chains) should provide powerful probes, both *in vitro* and *in vivo*, for protein structure and function. It might also enable the rational design or evolution of proteins with novel properties. Examples might include

selectively glycosylated or PEGylated therapeutic proteins with improved pharmacological properties, fluorescent proteins that act as sensors of small molecules and protein–protein interactions in the cell, or proteins whose activity can be photoregulated *in vivo*. It should also be possible to incorporate non-amino-acid building blocks into proteins or perhaps create biopolymers with entirely unnatural backbones (e.g. from α-hydroxy- or β-amino acid building blocks). Finally, the ability to add novel amino acids to the genetic code of organisms will enable us to test experimentally whether there is an evolutionary advantage for organisms with expanded genetic codes.

Acknowledgements

This work was supported by grants from the National Institutes of Health and Department of Energy. T.A.C. is supported by a National Research Service Award from NIH.

References

1 Leinfelder, W. *et al*. (1988) Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* 331, 721–725
2 Srinivasan, G. *et al*. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 296, 1459–1462
3 Yesland, K.D. *et al*. (1993) Identity of *Saccharomyces cerevisiae* tRNA(Trp) is not changed by an anticodon mutation that creates an amber suppressor. *J. Biol. Chem.* 268, 217–220
4 Buvoli, M. *et al*. (2000) Suppression of nonsense mutations in cell culture and mice by multimerized suppressor tRNA genes. *Mol. Cell. Biol.* 20, 3116–3124
5 Kwok, Y. *et al*. (1980) Evolutionary relationship between *Halobacterium cutirubrum* and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. *Can. J. Biochem.* 58, 213–218
6 Steer, B.A. *et al*. (1999) Major anticodon-binding region missing from an archaebacterial tRNA synthetase. *J. Biol. Chem.* 274, 35601–35606
7 Wang, L. *et al*. (2000) A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the *in vivo* incorporation of unnatural amino acids into proteins. *J. Am. Chem. Soc.* 122, 5010–5011
8 Wang, L. and Schultz, P.G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* 8, 883–890
9 Anderson, J.C. and Schultz, P.G. (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry* 42, 9598–9608
10 Wakasugi, K. *et al*. (1998) Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. *EMBO J.* 17, 297–305
11 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
12 Edwards, H. *et al*. (1991) An *Escherichia coli* tyrosine transfer RNA is a leucine-specific transfer RNA in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 1153–1156
13 Brick, P. *et al*. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.* 208, 83–98
14 Wang, L. *et al*. (2001) Expanding the genetic code of *Escherichia coli*. *Science* 292, 498–500
15 Chin, J.W. *et al*. (2003) Progress toward an expanded eukaryotic genetic code. *Chem. Biol.* 10, 511–519
16 Vidal, M. *et al*. (1996) Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10321–10326
17 Vidal, M. *et al*. (1996) Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10315–10320
18 Chin, J.W. *et al*. (2002) Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* 124, 9026–9027
19 Chin, J.W. *et al*. (2002) Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11020–11024
20 Chin, J.W. and Schultz, P.G. (2002) *In vivo* photocrosslinking with unnatural amino. Acid mutagenesis. *Chembiochem* 11, 1135–1137
21 Wu, N. *et al*. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* (in press)
22 Xie, J. *et al.* The site-specific incorporation of p-iodophenylalanine for structure determination. (in press)
23 Wang, L. *et al*. (2003) Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 56–61
24 Zhang, Z. *et al*. (2003) A new strategy for the site-specific modification of proteins *in vivo*. *Biochemistry* 42, 6735–6746
25 Deiters, A. *et al*. (2003) Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* 125, 11782–11783
26 Zhang, Z. *et al*. (2004) A new strategy for the synthesis of glycoprotein. *Science* 303, 371–373
27 Alfonta, L. *et al*. (2003) Site-specific incorporation of a redox-active amino acid into proteins. *J. Am. Chem. Soc.* 125, 14662–14663
28 Mehl, R. *et al*. (2003) Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* 125, 935–939
29 Hohsaka, T. *et al*. (2001) Incorporation of nonnatural amino acids into proteins by using various four-base codons in an *Escherichia coli in vitro* translation system. *Biochemistry* 40, 11060–11064
30 Riyasalt, S. and Atkins, J.F. (1968) External suppression of a frame-shift mutant in *Salmonella*. *J. Mol. Biol.* 34, 541–557
31 Magliery, T.J. *et al*. (2001) Expanding the genetic code: selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in *Escherichia coli*. *J. Mol. Biol.* 307, 755–769
32 Sakamoto, K. *et al*. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* 30, 4692–4699
33 Zhang, Z. *et al*. (2004) Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8882–8887