ELSEVIER

Adding amino acids to the genetic repertoire

Jianming Xie and Peter G Schultz

Considerable progress has been made in expanding the number and nature of genetically encoded amino acids in *Escherichia coli*, yeast and mammalian cells in the past four years. To date, over 30 unnatural amino acids have been cotranslationally incorporated into proteins with high fidelity and efficiency by means of a unique codon and corresponding orthogonal tRNA–aminoacyl-tRNA synthetase pair. The incorporated amino acids contain spectroscopic probes, post-translational modifications, metal chelators, photoaffinity labels and unique functional groups. The ability to genetically encode additional amino acids, beyond the common 20, provides a powerful approach for probing protein structure and function both *in vitro* and *in vivo*, as well as generating proteins with new or enhanced properties.

Addresses
Department of Chemistry and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Corresponding author: Schultz, Peter G (schultz@scripps.edu)

Current Opinion in Chemical Biology 2005, **9**:548–554

This review comes from a themed issue on
Biopolymers
Edited by Scott A Strobel and Tom W Muir

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2005.10.011

Introduction

The ability to genetically encode unnatural amino acids with defined physical, chemical or biological properties provides a powerful new method to investigate and manipulate protein structure and function. In 2001, our group first reported the selective cotranslational incorporation of an unnatural amino acid, *O*-methyl-L-tyrosine, into proteins in *Escherichia coli* in response to an amber nonsense codon [1••]. Further development of this methodology during the past several years has made it possible to systematically incorporate a large number of structurally diverse unnatural amino acids into proteins in *E. coli* [1••], yeast [2••] and mammalian cells. To date, more than 30 unnatural amino acids have been incorporated into proteins in response to unique triplet or quadruplet codons with high fidelity and in yields of up to 1 g/l [3•,4]. In this article, we review recent advances in this novel methodology.

Methodology

To cotranslationally incorporate unnatural amino acids into proteins at specified sites, new components of the protein biosynthetic machinery are required. These include a codon that uniquely designates the unnatural amino acid, and an orthogonal tRNA–aminoacyl-tRNA synthetase pair that can specifically incorporate the unnatural amino acid into proteins in response to the cognate codon. The amber stop codon TAG is the least used among the three stop codons in *E. coli* and yeast, rarely terminates essential genes, and is efficiently translated by amber suppressor tRNAs *in vivo* and *in vitro*. Therefore, the use of TAG to encode novel amino acids is not expected to significantly perturb the growth of a host organism (which indeed is the case). One can also use the opal stop codon TGA, rare codons such as AGG, and codons made up of four nucleotides to encode new amino acids. Recently, a combination of amber and frameshift codons (TAG and AGGA, respectively) were used to simultaneously incorporate two different unnatural amino acids at distinct sites in the same protein with high fidelity [5•].

The first orthogonal tRNA–synthetase pair in *E. coli* was derived from a tyrosyl-tRNA synthetase (TyrRS)–$tRNA^{Tyr}$ pair from the archaea *Methanococcus jannaschii* (Mj). Archaeal tRNAs have distinct aminoacyl-tRNA synthetase recognition elements relative to their *E. coli* counterparts, and therefore do not cross-react with the endogenous synthetases of the latter. The Mj TyrRS also has a minimalist anticodon loop binding domain [6] which makes it possible to alter the anti-codon loop of its cognate tRNA to CUA (to suppress UAG) with little loss in affinity by the synthetase. In addition, it lacks an editing mechanism that could deacylate an unnatural amino acid. Finally, this pair can be expressed in high levels in functional form in *E. coli*. Indeed, an orthogonal tRNA–synthetase pair evolved from the Mj tyrosyl pair has been used to incorporate more than 20 new amino acids into protein in *E. coli*. Recently, several additional orthogonal *E. coli* pairs were generated from archaea using a consensus-based strategy. These pairs include a $tRNA^{Lys}$–LysRS pair from the archaea *Pyrococcus horikoshii* [5], a $tRNA^{Glu}$–GluRS pair from *Methanosarcina mazei* [7], and a heterologous pair consisting of a leucyl-tRNA synthetase from *Methanobacterium thermoautotrophicum* and a mutant leucyl tRNA derived from *Halobacterium sp* [8]. These and other new orthogonal pairs in *E. coli* will probably increase the structural diversity and number of unnatural amino acids that can be incorporated into proteins by this method.

The first orthogonal pair for use in yeast was generated from an *E. coli* TyrRS–$tRNA^{Tyr}$ pair. This pair has been evolved to incorporate 10 unnatural amino acids into proteins in response to the amber codon [2,9]. Another orthogonal pair derived from an *E. coli* $tRNA^{Leu}$–LeuRS pair was recently used to incorporate fluorescent and photocaged amino acids into proteins in yeast [10]. RajBhandary and co-workers have reported that human initiator tRNA and *E. coli* glutaminyl-tRNA synthetase (GlnRS) is also an orthogonal pair in yeast [11].

To alter the specificity of the orthogonal synthetase to acylate the cognate tRNA with the unnatural amino acid, and not any endogenous amino acids, a directed evolution approach was developed in which large libraries of synthetase variants were passed through a series of stringent positive and negative selections. For the tyrosyl system in *E. coli* [1••], the library was initially generated by randomizing five residues in the amino acid binding site of the TyrRS (Figure 1a). To identify synthetase variants specific for unnatural amino acids, the libraries were first transferred into cells containing chloramphenicol acetyl transferase (*CmR*) with an amber mutation at a permissive site, and grown in media containing chloramphenicol and the unnatural amino acid. Survivors contained synthetase variants that incorporate either the unnatural or an endogenous amino acid in response to the amber codon. Selected synthetase clones were then transferred into cells containing a toxic barnase gene with amber mutations at permissive sites, and grown in the absence of unnatural amino acid. All clones that charged endogenous amino acids produced full-length barnase protein and died [1••]. Repeated rounds of positive and negative selections resulted in the isolation of mutant synthetases that can specifically incorporate the unnatural amino acid in response to the amber codon (Figure 1b). In a simplified genetic selection system, an amber-T7/GFPuv was used as an additional reporter in the positive selection [12]. Suppression of amber codons introduced at permissive sites in the T7 RNA polymerase gene produces full-length T7 RNAP, which drives the expression of GFPuv.

A similar double-sieve selection strategy (Figure 1c) was developed for yeast [2••,13]. Two amber codons were introduced at permissive sites in the gene encoding the transcriptional activator GAL4. The reporter genes HIS3, URA3 and lacZ, under the control of a GAL4 promoter, were inserted into the chromosome of yeast MaV203. Suppression of amber codons in the GAL4 gene leads to the expression of full-length GAL4, which in turn activates the transcription of the reporter genes. Positive selection of an active-site library of *E. coli* TyrRS mutants (an orthogonal synthetase in yeast) was carried out in the presence of the unnatural amino acid and absence of histidine or uracil. Survivors encoded mutant TyrRSs that incorporate either the unnatural amino acid or an endogenous amino acid into the GAL4 protein. For the negative selection, survivors were plated on media lacking the unnatural amino acid but containing the protoxin 5-fluoroorotic acid (5-FOA). Those cells containing mutant synthetases that recognize an endogenous amino acid died as a result of the expression of the URA3 gene product, which converts 5-FOA to a toxin.

Recently, this methodology has been extended to mammalian cells. Yokoyama and co-workers showed that an *E. coli* TyrRS mutant and the *Bascillus stearothermophilus* amber suppressor $tRNA^{Tyr}$ can be used together to incorporate 3-iodotyrosine [14•] and *p*-benzoyl-L-phenylalanine [15] into proteins in response to the TAG codon in mammalian CHO cells. Similarly, we have used *E. coli* TyrRS mutants evolved in yeast (to accept *p*-azidophenylalanine, *p*-benzoyl-L-phenylalanine, *p*-iodophenylalanine, *p*-acetylphenylalanine and *p*-methoxyphenylalanine) together with the *B. stearothermophilus* amber suppressor $tRNA^{Tyr}$ to insert the unnatural amino acids in response to an amber codon in the gene encoding firefly luciferase in mammalian 293T cells (PG Schultz, unpublished data). Another orthogonal mammalian tRNA–synthetase pair has been developed from the *Bacillus subtilis* tryptophanyl tRNA–synthetase pair. A rationally designed point mutant of this synthetase has been used to incorporate 5-hydroxytryptophan into proteins in mammalian 293T cells in response to the opal TGA codon [16•].

A final consideration is that the unnatural amino acid to be incorporated should be either passively or actively transported into the host organism, or biosynthesized *in vivo*, and be stable inside the cell. LC-MS-based assays [17•] have shown that most unnatural amino acids, with the exception of highly charged amino acids (e.g., *p*-phosphonomethylphenylalanine), are transported into *E. coli* and yeast cytoplasm, consistent with the relatively broad substrate specificity of the amine and amino acid transporters. Some amino acids such as α-hydroxy acids are metabolized before incorporation and therefore require deletion of the corresponding enzymes if they are to be biosynthetically incorporated. In one case, an *E. coli* strain was engineered to both biosynthesize and cotranslationally incorporate the unnatural amino acid *p*-amino-L-phenylalanine in response to an amber codon into proteins [17•].

When the above system is used to express proteins containing unnatural amino acids in *E. coli*, typically 5–20 mg/l of each protein can be purified from minimal media. In general, suppression efficiencies range from 25% to 50% of wild type protein and translational fidelity is $> 99\%$. Recently, the system was optimized by using stronger promoters to drive the transcription of tRNA and mutant synthetase genes. By using the optimized systems in rich media, up to 1 g/l of mutant protein has been expressed in *E. coli* (H Cho *et al.*, unpublished data) and 75 mg/l in yeast (Y Ryu *et al.*, unpublished data).

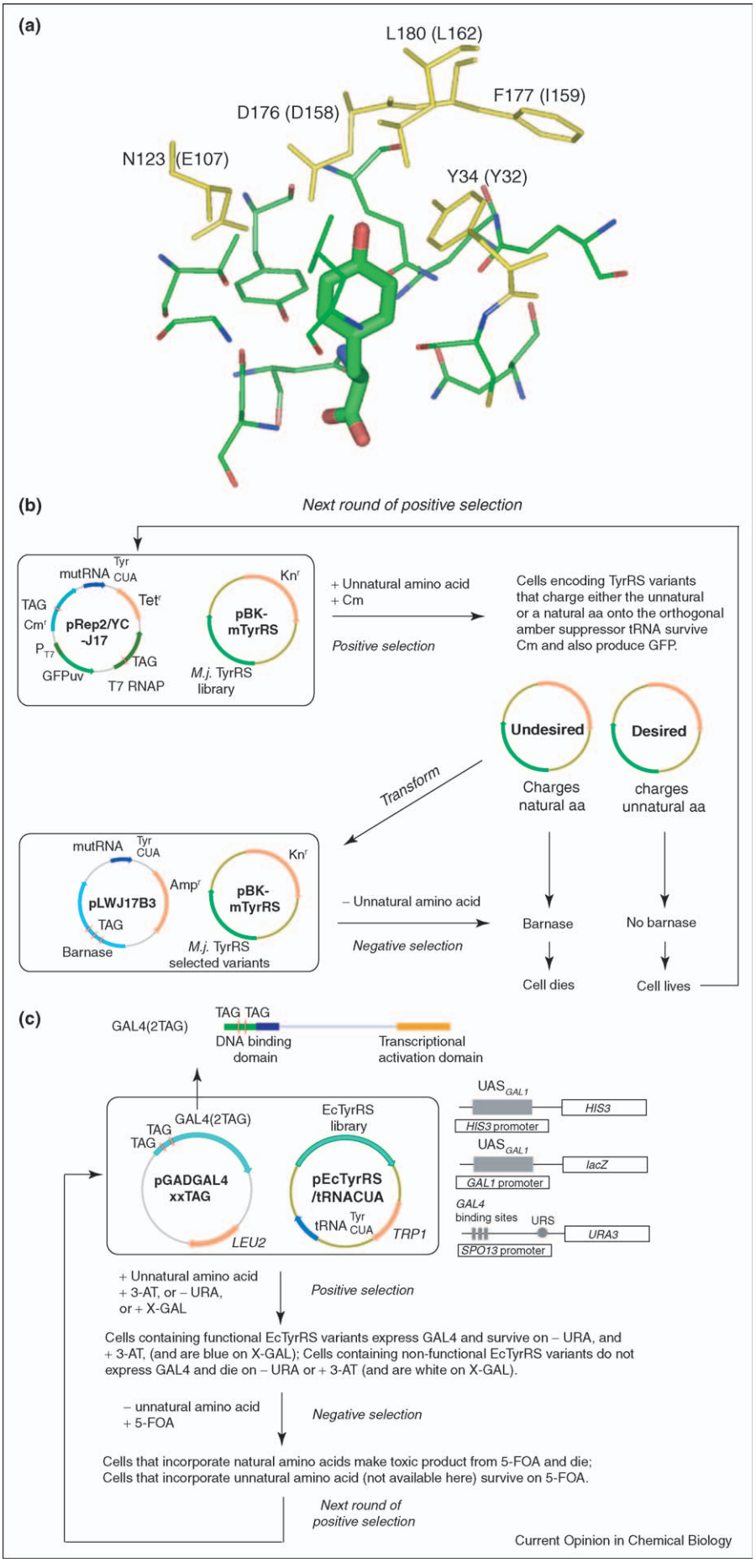
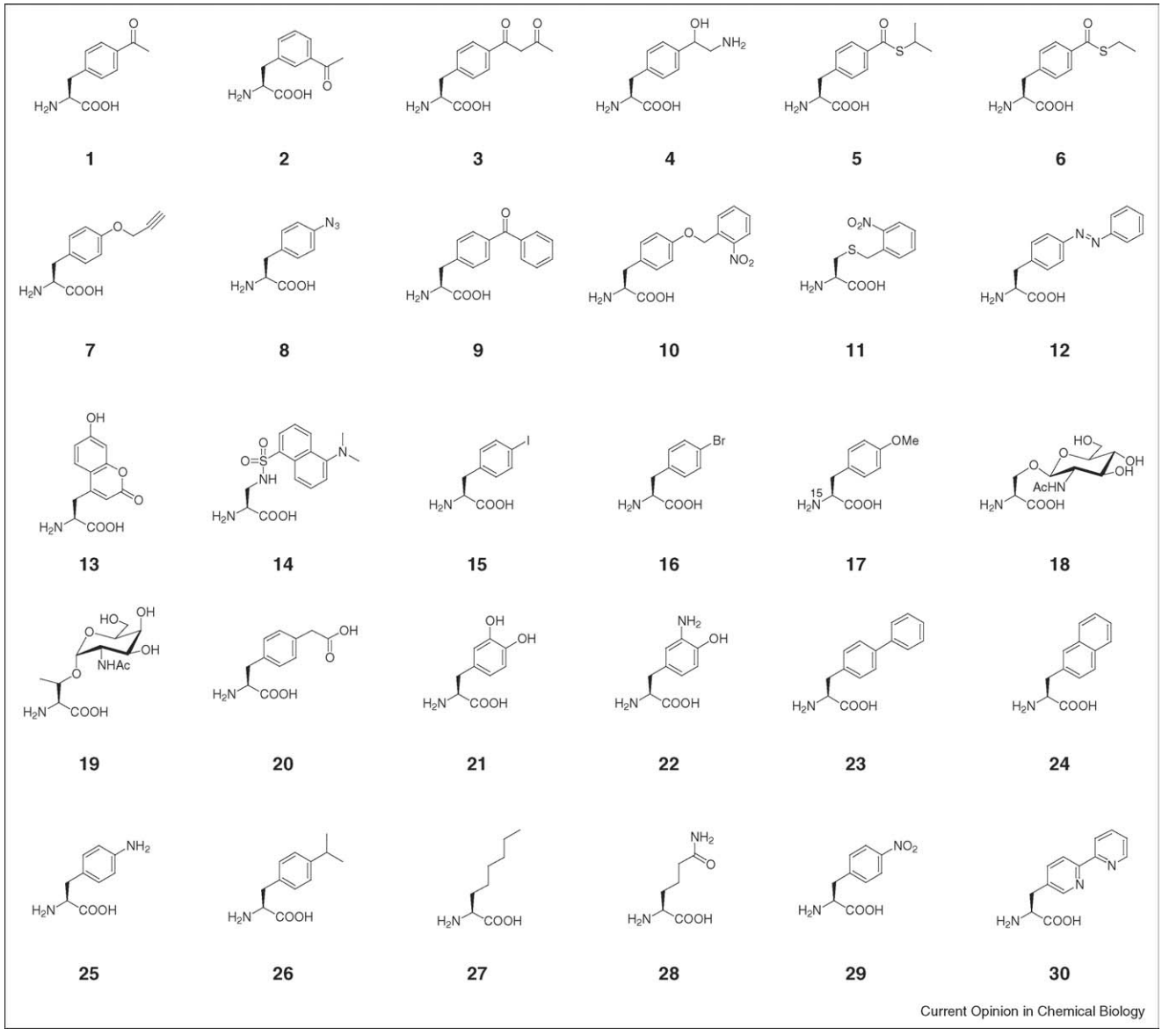

(a)
L180 (L162)
F177 (I159)
D176 (D158)
N123 (E107)
Y34 (Y32)
(b)
Next round of positive selection
mutRNA Tyr CUA
TAG
Cm^r
pRep2/YC -J17
P T7
GFPuv
Tet^r
TAG
T7 RNAP
Kn^r
pBK-mTyrRS
M.j. TyrRS library
+ Unnatural amino acid
+ Cm
Positive selection
Cells encoding TyrRS variants that charge either the unnatural or a natural aa onto the orthogonal amber suppressor tRNA survive Cm and also produce GFP.
Undesired
Desired
Charges natural aa
charges unnatural aa
Transform
mutRNA Tyr CUA
pLWJ17B3
TAG
Barnase
Amp^r
Kn^r
pBK-mTyrRS
M.j. TyrRS selected variants
– Unnatural amino acid
Negative selection
Barnase
No barnase
Cell dies
Cell lives
(c)
GAL4(2TAG)
TAG TAG
DNA binding domain
Transcriptional activation domain
GAL4(2TAG)
TAG
TAG
pGADGAL4 xxTAG
LEU2
EcTyrRS library
pEcTyrRS /tRNACUA
tRNA Tyr CUA
TRP1
UAS GAL1
HIS3
HIS3 promoter
UAS GAL1
lacZ
GAL1 promoter
GAL4 binding sites
URS
URA3
SPO13 promoter
+ Unnatural amino acid
+ 3-AT, or – URA,
or + X-GAL
Positive selection
Cells containing functional EcTyrRS variants express GAL4 and survive on – URA, and + 3-AT, (and are blue on X-GAL); Cells containing non-functional EcTyrRS variants do not express GAL4 and die on – URA or + 3-AT (and are white on X-GAL).
– unnatural amino acid
+ 5-FOA
Negative selection
Cells that incorporate natural amino acids make toxic product from 5-FOA and die;
Cells that incorporate unnatural amino acid (not available here) survive on 5-FOA.
Next round of positive selection
Current Opinion in Chemical Biology

Figure 2



Unnatural amino acids that have been added to the genetic codes of *E. coli* or yeast.

An expanded amino acid repertoire

More than 30 unnatural amino acids (Figure 2) have been successfully incorporated in *E. coli*, yeast or mammalian cells with the orthogonal tRNA–synthetase pairs and the genetic selections described above. For example, unnatural amino acids with uniquely reactive functional groups have been genetically encoded in *E. coli* and/or yeast [18–21] including: *p*-acetylphenylalanine (**1**), *m*-acetylphenylalanine (**2**), *p*-(3-oxobutanoyl)-L-phenylalanine (**3**), *p*-(2-amino-3-hydroxyethyl)phenylalanine (**4**), *p*-isopropylthiocarbonyl-phenylalanine (**5**), *p*-ethylthiocarbonyl-phenylalanine (**6**), *p*-propargyloxyphenylalanine (**7**) and *p*-azidophenylalanine (**8**). These amino acids can be used to selectively modify proteins under mild conditions with a

(Figure 1 Legend) Modification of the amino acid specificity of an orthogonal TyrRS in *E. coli* and in yeast. **(a)** A library of *M. jannaschii* TyrRS mutants was generated by randomizing five residues (in parentheses). Residues were selected based on observed contacts between the homologous *B. stearothermophilus* TyrRS residues and tyrosyl adenylate in the crystal structure of the *B. stearothermophilus* TyrRS–tyrosyl adenylate complex. A similar library of *E. coli* TyrRS mutants (not shown) was generated for selection in yeast. **(b)** directed evolution of *M. jannaschii* TyrRS variants in *E. coli*. **(c)** Directed evolution of *E. coli* TyrRS variants in yeast. Abbreviation: aaRS, aminoacyl-tRNA synthetase; aa, amino acid; 5-FOA, protoxin 5-fluoroorotic acid.

variety of reagents. For example, the keto- and β-diketo-moieties can selectively react with both hydrazides and hydroxylamine derivatives to form stable hydrazone and oxime linkages, respectively [18,20]. This approach has been used to modify proteins with fluorophores [18,20], biotin [18,20], sugar analogues [22] and polyethylene glycols (PEGs), etc. The azide containing amino acid **8** can be selectively modified by copper(I) catalyzed [3 + 2] cycloadditions with an alkyne derivative [21] (and vice versa), or by a Staudinger ligation with appropriate water-soluble phosphine-containing reagents [23]. The thioester moiety in amino acids **4** and **5** can react with amines to form stable amide bonds.

Unnatural amino acids containing photoactive side chains have also been selectively incorporated into proteins. For example, two photo-crosslinking amino acids, *p*-azido-L-phenylalanine (**8**) [2••,19] and *p*-benzoyl-L-phenylalanine (*p*BPA) (**9**) [24–26], have been added to the genetic codes of both *E. coli* and yeast. An *E. coli* TyrRS mutant evolved to specifically incorporate *p*BPA in yeast has been used by Yokoyama and co-workers to incorporate *p*BPA into human Grb2 protein in CHO cells. The synthesized Grb2 variant containing *p*BPA was cross-linked with epidermal growth factor (EGF) receptor upon exposure of cells to 365 nm light [15]. Amino acids with photocaged side chains, *O*-(2-nitrobenzyl)tyrosine (**10**) and *S*-(2-nitrobenzyl)cysteine (**11**) [10], have been genetically encoded in *E. coli* and yeast, respectively. In model studies, these amino acids were used to generate light-activated β-galactosidase and caspase 3 mutants. A photoisomerizable amino acid, *p*-azophenyl-phenylalanine (azoPhe) (**12**), has been genetically encoded in *E. coli*. Introduction of this amino acid at specific sites in the cAMP-binding domain of an *E. coli* transcriptional activator CAP (catabolite activator protein) made it possible to photochemically regulate the accessibility of the functional domain of CAP to cAMP, affording a light-controlled genetic switch [27].

Useful probes of protein structure and function have also been selectively incorporated into proteins using this methodology. For example, fluorescent amino acids with 7-hydroxycoumarin (**13**) and dansyl side chains (**14**) have been selectively incorporated into proteins in *E. coli* and in yeast (PG Schultz, unpublished data), providing small fluorescent probes for direct visualization of protein conformational changes, localization and intermolecular interactions. The heavy-atom containing amino acid *p*-iodo-L-phenylalanine (**15**) has been genetically encoded both in *E. coli* and in yeast, and can be used for SAD phasing in structure determination [28•]. In addition, ^{15}N-labeled *O*-methyltyrosine (**17**) has been selectively incorporated into proteins as an NMR probe [29].

Unnatural amino acids corresponding to post-translational modifications have also been cotranslationally incorporated into proteins. For example, the selective incorporation of β-GlcNAc-serine (**18**) and α-GalNAc-threonine (**19**) into proteins provides a method for synthesizing homogenous glycoproteins [30,31•]. *p*-Carboxymethyl-phenylalanine (**20**) has also been incorporated into proteins as a phosphotyrosine mimic. The hydroxylated amino acid 3,4-dihydroxy-L-phenylalanine (DHP) (**21**) has been site-specifically incorporated into proteins in *E. coli* [32]; because this amino acid can be electrochemically oxidized to the semi-quinone radical or quinone state within the protein it may facilitate study of electron transfer in proteins, as well as the engineering of redox proteins with novel properties. More recently, another redox-active amino acid, 3-amino-L-tyrosine (**22**), was incorporated into proteins in *E. coli*. This amino acid can act as a radical trap due to the stability of its oxidized semiquinone form, or can serve as a unique handle for chemical modification of proteins [33].

Unnatural amino acids with aromatic hydrophobic side chains, including biphenylalanine (**23**) and L-3-(2-naphthyl)alanine (**24**) [34], have been added to the genetic code of *E. coli*. They may be useful in modulating hydrophobic packing interactions in proteins or between proteins. Other representative unnatural amino acids that have been added to the genetic code of *E. coli* or yeast include *p*-amino-L-phenylalanine (**25**) [17•], *p*-isopropyl-L-phenylalanine (**26**) [12], α-aminocaprylic acid (**27**) [10], L-homoglutamine (**28**) [5•], and *p*-nitro-L-phenylalanine (**29**). Finally, the bipyridyl-containing amino acid (**30**) has been added to the genetic code of *E. coli* (PG Schultz, unpublished data). This amino acid chelates transition metal ions (e.g., Zn^{2+}, Cu^{2+}, Fe^{2+} and Ru^{2+}), and may facilitate the generation of metalloproteins with novel properties (e.g., metal-dependent protein dimerization, metalloproteases, or *in vivo* imaging agents). Clearly, a large number of structurally diverse building blocks have been added to the genetic code using this approach.

The structures of the complexes between several mutant *M. jannaschii* TyrRSs and their cognate unnatural amino acids (e.g., naphthylalanine, *p*-iodophenylalanine and *p*-acetylphenylalanine specific) have been determined by X-ray crystallography. Surprisingly, some of the mutant synthetases have significant conformational changes in both the side chains and peptide backbone compared with the wild type enzyme [35]. These conformational changes lead to altered packing and hydrogen bonding interactions that favor binding of the unnatural amino acid, and at the same time disfavor binding of tyrosine. The large degree of structural plasticity in the active site of this enzyme may account for the relative ease with which novel amino acids can be genetically encoded.

Most recently, we have also used this methodology in conjunction with phage display as a general approach to generate polypeptide libraries containing unnatural amino acids [36•]. This should significantly increase

the scope of phage display technology. Other display formats such as ribosome and yeast display may also be extended by unnatural amino acid incorporation.

Concluding remarks

Given the structural diversity and number of unnatural amino acids that have been selectively incorporated into proteins to date, it is likely that a large number of additional amino acids can be genetically encoded, such as spin labels and IR probes, phosphorylated, prenylated, acetylated and methylated amino acids, conformationally restricted amino acids, and even α-hydroxyl or *N*-methyl amino acids. By using quadruplet codons or systematically reassigning degenerate triplet codons, it should be possible to further expand the code. Moreover, we are attempting to generalize this methodology to multicellular organisms. This methodology should provide a powerful new approach toward the study of proteins (both *in vivo* and *in vitro*) in which amino acids with novel physical, chemical, or biological properties can be tailored to address a specific structural/functional question at hand. And finally, the ability to add new building blocks to the genetic code may allow the rational design or evolution of proteins with new or enhanced properties.

Acknowledgements

This work is supported by grants from the Department of Energy (ER46051) and National Institutes of Health (GM62159) and by the Skaggs Institute for Chemical Biology.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Wang L, Brock A, Herberich B, Schultz PG: **Expanding the genetic code of *Escherichia coli***. *Science* 2001, **292**:498-500.
This was the first report that an unnatural amino acid (*O*-methyl-L-tyrosine) could be successfully added to the genetic code of *E. coli* by means of amber-suppression and an engineered orthogonal *M. jannaschii* TyrRS–tRNA pair.

2. •• Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG: **An expanded eukaryotic genetic code**. *Science* 2003, **301**:964-967.
This report expanded this methodology to the genetic code of yeast.

3. • Wang L, Schultz PG: **Expanding the genetic code**. *Angew Chem Int Ed Engl* 2004, **44**:34-66.
A comprehensive review of this methodology as well as other methods for introducing unnatural amino acids into proteins.

4. Cropp TA, Schultz PG: **An expanding genetic code**. *Trends Genet* 2004, **20**:625-630.

5. • Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG: **An expanded genetic code with a functional quadruplet codon**. *Proc Natl Acad Sci USA* 2004, **101**:7566-7571.
A combination of amber and frameshift (AGGA) suppression was used to incorporate two unnatural amino acids into a protein simultaneously with high fidelity.

6. Steer BA, Schimmel P: **Major anticodon-binding region missing from an archaebacterial tRNA synthetase**. *J Biol Chem* 1999, **274**:35601-35606.

7. Santoro SW, Anderson JC, Lakshman V, Schultz PG: **An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli***. *Nucleic Acids Res* 2003, **31**:6700-6709.

8. Anderson JC, Schultz PG: **Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression**. *Biochemistry* 2003, **42**:9598-9608.

9. Deiters A, Cropp TA, Mukherji M, Chin JW, Anderson JC, Schultz PG: **Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae***. *J Am Chem Soc* 2003, **125**:11782-11783.

10. Wu N, Deiters A, Cropp TA, King D, Schultz PG: **A genetically encoded photocaged amino Acid**. *J Am Chem Soc* 2004, **126**:14306-14307.

11. Kowal AK, Kohrer C, RajBhandary UL: **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* 2001, **98**:2268-2273.

12. Santoro SW, Wang L, Herberich B, King DS, Schultz PG: **An efficient system for the evolution of aminoacyl-tRNA synthetase specificity**. *Nat Biotechnol* 2002, **20**:1044-1048.

13. Chin JW, Cropp TA, Chu S, Meggers E, Schultz PG: **Progress toward an expanded eukaryotic genetic code**. *Chem Biol* 2003, **10**:511-519.

14. • Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, Kobayashi T, Kitabatake M, Takio K, Saito K *et al.*: **Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells**. *Nucleic Acids Res* 2002, **30**:4692-4699.
An early example of unnatural amino acid incorporation in mammalian cells. A heterologous pair of *E. coli* TyrRS mutant and *B. stearothermophilus* amber suppressor $tRNA^{Tyr}$ were used to incorporate 3-iodotyrosine into proteins in CHO cells.

15. Hino N, Okazaki Y, Kobayashi T, Hayashi A, Sakamoto K, Yokoyama S: **Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid**. *Nat Methods* 2005, **2**:201-206.

16. • Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, Schultz PG: **Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells**. *Proc Natl Acad Sci USA* 2004, **101**:8882-8887.
An orthogonal synthetase–tRNA pair was generated from the *Bacillus subtilis* tryptophanyl tRNA-synthetase pair to incorporate 5-hydroxytryptophan into proteins in mammalian 293T cells.

17. • Mehl RA, Anderson JC, Santoro SW, Wang L, Martin AB, King DS, Horn DM, Schultz PG: **Generation of a bacterium with a 21 amino acid genetic code**. *J Am Chem Soc* 2003, **125**:935-939.
This article describes the generation of a completely autonomous bacterium with a 21-amino-acid genetic code. This bacterium can biosynthesize a non-standard amino acid (*p*-amino-L-phenylalanine) from basic carbon sources using an imported biosynthetic pathway, and incorporate this synthesized amino acid into proteins in response to the amber nonsense codon.

18. Wang L, Zhang Z, Brock A, Schultz PG: **Addition of the keto functional group to the genetic code of *Escherichia coli***. *Proc Natl Acad Sci USA* 2003, **100**:56-61.

19. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG: **Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli***. *J Am Chem Soc* 2002, **124**:9026-9027.

20. Zhang Z, Smith BA, Wang L, Brock A, Cho C, Schultz PG: **A new strategy for the site-specific modification of proteins *in vivo***. *Biochemistry* 2003, **42**:6735-6746.

21. Deiters A, Cropp TA, Summerer D, Mukherji M, Schultz PG: **Site-specific PEGylation of proteins containing unnatural amino acids**. *Bioorg Med Chem Lett* 2004, **14**:5743-5745.

22. Liu H, Wang L, Brock A, Wong CH, Schultz PG: **A method for the generation of glycoprotein mimetics**. *J Am Chem Soc* 2003, **125**:1702-1703.

23. Tsao M, Tian F, Schultz PG: **The Staudinger modification of proteins containing azido amino acids**. *ChemBioChem* 2005, in press.

24. Chin JW, Schultz PG: ***In vivo* photocrosslinking with unnatural amino acid mutagenesis**. *ChemBioChem* 2002, **3**:1135-1137.

25. Chin JW, Martin AB, King DS, Wang L, Schultz PG: **Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli***. *Proc Natl Acad Sci USA* 2002, **99**:11020-11024.

26. Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW: **Photo-cross-linking interacting proteins with a genetically encoded benzophenone**. *Nat Methods* 2005, **2**:377-384.

27. Bose M, Groff D, Xie J, Brustad E, Schultz PG: **The incorporation of a photoisomerizable amino acid into proteins in *E. coli***. *J Am Chem* Soc 2005, in press.

28. • Xie J, Wang L, Wu N, Brock A, Spraggon G, Schultz PG: **The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination**. *Nat Biotechnol* 2004, **22**:1297-1301.
A heavy-atom-containing amino acid, iodophenylalanine, was introduced into proteins with high fidelity and efficiency, affording a reliable method to prepare iodinated proteins to facilitate SAD phasing in crystallography.

29. Deiters A, Geierstanger BH, Schultz PG: **Site-specific *in vivo* labeling of proteins for NMR studies**. *ChemBioChem* 2005, **6**:55-58.

30. Xu R, Hanson SR, Zhang Z, Yang YY, Schultz PG, Wong CH: **Site-specific incorporation of the mucin-type N-acetylgalactosamine-alpha-O-threonine into protein in *Escherichia coli***. *J Am Chem Soc* 2004, **126**:15654-15655.

31. • Zhang Z, Gildersleeve J, Yang YY, Xu R, Loo JA, Uryu S, Wong CH, Schultz PG: **A new strategy for the synthesis of glycoproteins**. *Science* 2004, **303**:371-373.
The incorporation of a glycosylated amino acid directly into proteins provides a promising tool for the synthesis of homogenous glycoproteins.

32. Alfonta L, Zhang Z, Uryu S, Loo JA, Schultz PG: **Site-specific incorporation of a redox-active amino acid into proteins**. *J Am Chem Soc* 2003, **125**:14662-14663.

33. Hooker JM, Kovacs EW, Francis MB: **Interior surface modification of bacteriophage MS2**. *J Am Chem Soc* 2004, **126**:3718-3719.

34. Wang L, Brock A, Schultz PG: **Adding L-3-(2-naphthyl)alanine to the genetic code of *E. coli***. *J Am Chem Soc* 2002, **124**:1836-1837.

35. Turner JM, Graziano J, Spraggon G, Schultz PG: **Structural characterization of a *p*-acetylphenylalanyl aminoacyl-tRNA synthetase**. *J Am Chem Soc* 2005, in press.

36. • Tian F, Tsao ML, Schultz PG: **A phage display system with unnatural amino acids**. *J Am Chem Soc* 2004, **126**:15962-15963.
Unnatural amino acid mutagenesis was used to expand the scope of phage display technology.