Genetic incorporation of unnatural amino acids into proteins in mammalian cells

Wenshe Liu[1], Ansgar Brock[2], Shuo Chen[1,3], Shuibing Chen[1] & Peter G Schultz[1,2]

We developed a general approach that allows unnatural amino acids with diverse physicochemical and biological properties to be genetically encoded in mammalian cells. A mutant *Escherichia coli* aminoacyl-tRNA synthetase (aaRS) is first evolved in yeast to selectively aminoacylate its tRNA with the unnatural amino acid of interest. This mutant aaRS together with an amber suppressor tRNA from *Bacillus stearothermophilus* is then used to site-specifically incorporate the unnatural amino acid into a protein in mammalian cells in response to an amber nonsense codon. We independently incorporated six unnatural amino acids into GFP expressed in CHO cells with efficiencies up to 1 μg protein per 2×10^7 cells; mass spectrometry confirmed a high translational fidelity for the unnatural amino acid. This methodology should facilitate the introduction of biological probes into proteins for cellular studies and may ultimately facilitate the synthesis of therapeutic proteins containing unnatural amino acids in mammalian cells.

Several methods have been developed to incorporate unnatural amino acids site-specifically into proteins in mammalian cells. Chemically aminoacylated suppressor tRNAs have been microinjected or electroporated into CHO cells and neurons, respectively, and used to suppress nonsense amber mutations with a series of unnatural amino acids[1]. The use of the aminoacylated tRNA as a stoichiometric reagent, however, severely limits the amount of protein that can be produced. Alternatively, heterologous suppressor tRNA-aaRS pairs that do not cross-react with host tRNAs, aaRSs or amino acids (orthogonal tRNA-aaRSs) have been engineered to incorporate unnatural amino acids selectively into proteins. For example, a *B. stearothermophilus* amber suppressor $tRNA^{Tyr}_{CUA}$ (*Bst*$tRNA^{Tyr}_{CUA}$) and a mutant *E. coli* tyrosyl-tRNA synthetase (*Ec*TyrRS) pair have been used to incorporate 3-iodo-L-tyrosine into proteins in CHO cells[2]. Similarly, an orthogonal *B. subtilis* suppressor tRNA and tryptophanyl-tRNA synthetase pair has been engineered to incorporate 5-hydroxy-tryptophan into proteins in mammalian cells[3]. However, the use of structure-based mutagenesis to generate aaRS variants that aminoacylate an amino acid whose side chain differs substantially from that of the wild-type substrate likely requires mutations of multiple active site residues that are difficult to predict *a priori*[4–6]; moreover the mutant may still recognize host amino acids as is the case with a mutant aaRS that charges its cognate $tRNA^{Tyr}_{CUA}$ with 3-iodo-L-tyrosine[2,7]. Alternatively, one can attempt to evolve aaRSs with altered specificities directly in mammalian cells. For example, somatic hypermutation in a human B cell line recently was used to directly evolve a monomeric red fluorescent protein with enhanced photostability and far-red emissions[8]. Somatic hypermutation, however, introduces random mutations in the entire protein, which may be less effective than genetic diversity created by targeted mutagenesis of the active site when evolving variants with altered substrate specificity. The latter, however, is limited by difficulties in generating large stable libraries in mammalian cells.

We recently developed a general approach that makes it possible to genetically encode unnatural amino acids with diverse physical, chemical or biological properties in bacteria and yeast[9]. In this method, an orthogonal suppressor tRNA-aaRS pair is used to introduce the unnatural amino acid in response to a unique nonsense or frameshift codon in the gene of interest. In *E. coli*, functional amber and frameshift suppressor tRNA–aaRS pairs have been derived from a *Methanococcus jannaschii* $tRNA^{Tyr}$–TyrRS pair, a *Pyrococcus horikoshii* glutamyl-tRNA synthetase–$tRNA_{CUA}$ pair and a *Pyrococcus horikoshii* lysyl-tRNA synthetase–$tRNA_{CUA}$ pair[10–12]. In *Saccharomyces cerevisiae*, functional $tRNA_{CUA}$-aaRS pairs have been derived from the corresponding *E. coli* $tRNA^{Tyr}$-TyrRS and $tRNA^{Leu}$–leucyl-tRNA synthetase pairs[13–15]. Directed evolution of these suppressor tRNA-aaRS pairs using a combination of positive and negative selections has allowed the efficient, highly selective *in vivo* incorporation of a large number of diverse unnatural amino acids in *E. coli* and *S. cerevisiae*. These include fluorescent[16,17], glycosylated[18], sulfated[19], metal-ion-binding and redox-active amino acids[20], as well as amino acids with new chemical and photochemical reactivity[9]. This methodology provides a powerful tool for exploring protein structure and function *in vitro* and *in vivo*, and for generating proteins with new or enhanced properties. The extension of this methodology to mammalian cells would considerably enhance the utility of this technology.

[1]Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 Torrey Pines Road, La Jolla, California 92037, USA. [2]Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, California 92121, USA. [3]Present address: Department of Biological Sciences, Ohio University, Athens, Ohio 45701, USA. Correspondence should be addressed to P.G.S. (schultz@scripps.edu).

RECEIVED 9 NOVEMBER 2006; ACCEPTED 22 JANUARY 2007; PUBLISHED ONLINE 25 FEBRUARY 2007; DOI:10.1038/NMETH1016

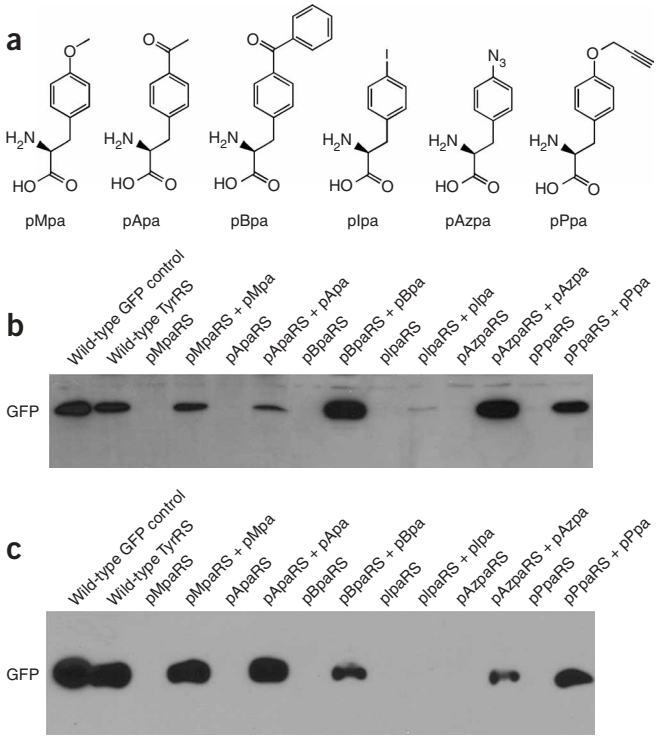

Figure 1 | Amber suppression of six *Bst*RNA$^{Tyr}_{CUA}$-*Ec*TyrRS pairs in T-REx CHO and T-REx 293 cells. (**a**) Structures of the six unnatural amino acids. (**b**) Western blot analysis of full-length GFP expression in T-REx CHO cells that were cotransfected with plasmids pcDNA4-GFP37TAG, pcDNA4-*Ec*TyrRS variant and pUC18-3*Bst*RNA$^{Tyr}_{CUA}$ and grown in the presence or absence of unnatural amino acids. The first lane is wild-type GFP expression in T-REx CHO cells that were transiently transfected with pcDNA4-GFP. The second lane is full-length expression of GFP37TAG suppressed by wild-type *Ec*TyrRS together with *Bst*RNA$^{Tyr}_{CUA}$. The following lanes are full-length expression of GFP37TAG suppressed by *Ec*TyrRS variants together with *Bst*RNA$^{Tyr}_{CUA}$ in the presence or absence of unnatural amino acids, as indicated. A 40-μg aliquot of the cell lysate for each reaction (10 μg cell lysate was loaded into the first lane) was analyzed with anti–c-myc. (**c**) Western blot analysis of full-length GFP expression in T-REx 293 cells. A 20-μg aliquot of the cell lysate for each reaction (5 μg for the control reaction) was analyzed with anti-His–HRP.

Here we describe a general approach in which mutant aaRSs (originally evolved in *S. cerevisiae*) together with an amber suppressor from *B. stearothermophilus* were used to incorporate six unnatural amino acids, including *p*-methoxyphenylalanine (pMpa), *p*-acetylphenylalanine (pApa), *p*-benzoylphenylalanine (pBpa), *p*-iodophenylalanine (pIpa), *p*-azidophenylalanine (pAzpa) and *p*-propargyloxyphenylalanine (pPpa; **Fig. 1a**), into proteins in CHO and human 293T cells. These unnatural amino acids are incorporated site-specifically into GFP in response to an amber nonsense codon with excellent fidelity and good efficiency.

RESULTS

Nonsense suppression by *Ec*TyrRS-*Bst*RNA$^{Tyr}_{CUA}$ pair

The most straightforward strategy to genetically encode unnatural amino acids in mammalian cells would be to adapt the existing mutant tRNA-aaRS pairs that have been generated in *E. coli* or *S. cerevisiae*. Because the tRNA^{Tyr} identity elements in *E. coli* (which include the variable arm and G1·C72 base pair in the acceptor stem) are distinct from those in mammalian cells[21,22], the mutant tRNA-aaRS pairs that have been used in bacteria are likely not orthogonal in eukaryotic cells. In contrast, tRNAs in *S. cerevisiae* and mammalian cells are processed similarly and have the same identity elements[22]. Moreover, because the translational machinery of *S. cerevisiae* is also homologous to that of higher eukaryotes, it is likely that one can directly transfer an orthogonal tRNA-aaRS pair evolved in *S. cerevisiae* into mammalian cells. Indeed, an *Ec*TyrRS variant that had been evolved in *S. cerevisiae* to accept pBpa has been previously used to incorporate this photoreactive amino acid into human Grb2 protein in CHO cells[23]. Unfortunately, functionally active *E. coli* tRNA$^{Tyr}_{CUA}$ does not express well in mammalian cells, severely limiting the yields of mutant protein[2]. To overcome this limitation, we created a general system that allows one to rapidly transfer mutant suppressor tRNA-aaRS pairs evolved in yeast to mammalian cells and used this method to efficiently introduce several unnatural amino acids into GFP in both CHO and 293T cells.

tRNAs in eukaryotes are transcribed by RNA polymerase III, which recognizes two conserved intragenic transcriptional control elements, the A and B boxes[24]. *E. coli* tRNA^{Tyr} only has a B box element, and it has been shown that the introduction of a pseudo-A box results in a nonfunctional tRNA that is not recognized by *Ec*TyrRS[2]. Unlike *E. coli* tRNA^{Tyr}, the tRNA^{Tyr} from *B. stearothermophilus* (which has similar identity elements and is still charged by *Ec*TyrRS[25]) has naturally occurring internal A and B boxes. Thus, this tRNA together with *Ec*TyrRS functions as an orthogonal tRNA-aaRS pair in mammalian cells[2]. To afford an amber suppressor *Bst*RNA$^{Tyr}_{CUA}$, we changed the trinucleotide anticodon of *Bst*RNA^{Tyr} to CUA (positions 34–36). Because guanine 34 of prokaryotic tRNA^{Tyr} is only a weak identity element of TyrRS[26], mutation of this base to a cytosine should not significantly affect the binding of *Ec*TyrRS to *Bst*RNA$^{Tyr}_{CUA}$. Furthermore, nonsense amber suppression should be better tolerated in mammalian cells than opal or ochre suppression owing to the lower frequency of the TAG stop codon in mammalian genomes (TAG, 23%; TAA, 30%; TGA, 47% in *Homo sapiens*). Because expression of the *Bst*RNA$^{Tyr}_{CUA}$ gene in eukaryotes also depends on the 5′ flanking sequence, we added the 5′ flanking sequence of human tRNA^{Tyr} to *Bst*RNA$^{Tyr}_{CUA}$ to enhance its transcription in mammalian cells. To further increase transcription of *Bst*RNA$^{Tyr}_{CUA}$, we constructed a gene cluster containing three tandem repeats of the *Bst*RNA$^{Tyr}_{CUA}$ gene and inserted it into pUC18 plasmid to create pUC18-3*Bst*RNA$^{Tyr}_{CUA}$. Northern blot analysis of isolated total tRNAs from CHO cells transfected with pUC18-3*Bst*RNA$^{Tyr}_{CUA}$ showed a twofold higher level of *Bst*RNA$^{Tyr}_{CUA}$ than cells transfected with a pUC18 plasmid containing only one copy of *Bst*RNA$^{Tyr}_{CUA}$ (data not shown).

Next we inserted the wild-type *Ec*TyrRS gene into the mammalian expression vector pcDNA4/TO/myc-HisA to afford pcDNA4-*Ec*TyrRS in which expression is controlled by a tetracycline-regulated CMV promoter. The use of an inducible expression system was intended to lower possible toxicity that results from heterologous expression of *Ec*TyrRS in mammalian cells. We then assayed for the ability of the resulting suppressor *Bst*RNA$^{Tyr}_{CUA}$-*Ec*TyrRS pair to efficiently suppress an amber codon mutation at position Tyr37 of GFP (GFP37TAG) in mammalian cells. Because Tyr37 is located at the surface of GFP and distal to the fluorophore, the introduction of an unnatural amino acid at this position is not expected to affect the folding and fluorescent

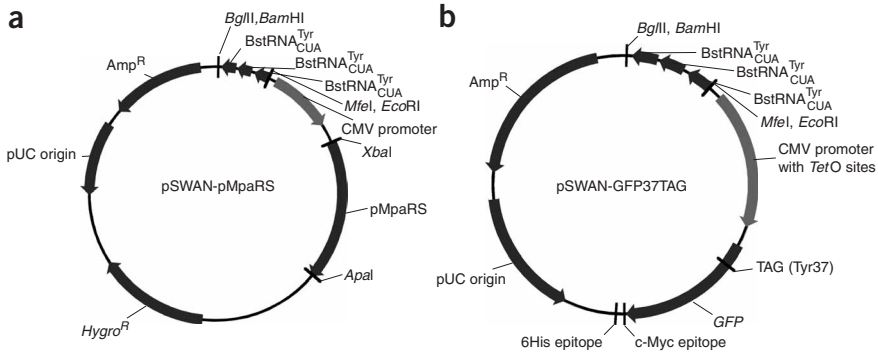

Figure 2 | Plasmid maps. (**a**) pSWAN-pMpaRS. (**b**) pSWAN-GFP37TAG.

properties of the protein[27]. Moreover, amber suppression of GFP37TAG results in expression of full-length GFP, providing a rapid qualitative assay of amber-suppression efficiency. We inserted the mutant *GFP37TAG* gene into pcDNA4/TO/myc-HisA to afford pcDNA4-GFP37TAG, which contains a gene encoding GFP fused to myc and 6His epitopes at the C terminus, and is under the control of a tetracycline-regulated promoter.

We then examined the ability of the *Bst*RNA$^{Tyr}_{CUA}$-*Ec*TyrRS pair to suppress the nonsense amber codon of GFP37TAG in both T-REx CHO and 293 cells. Both cell lines constitutively express the tetracycline repressor and are suitable for tetracycline-regulated expression of proteins using the pcDNA4/TO/myc-HisA plasmid. We grew cells to 80–90% confluency and transiently transfected them with 3 μg of plasmids per 2×10^6 cells using Fugene 6 (0.5 μg of pUC18-3*Bst*RNA$^{Tyr}_{CUA}$, 0.5 μg of pcDNA4-*Ec*TyrRS and 2 μg of pcDNA4-GFP37TAG; we used the same amount of pUC18 in the absence of pUC18-3*Bst*RNA$^{Tyr}_{CUA}$or pcDNA4-*Ec*TyrRS). We induced protein expression by the addition of 1 μg/ml tetracycline 6 h after transfection, and grew cells at 37 °C for 2 d. In both cell lines, the expression of full-length GFP was dependent on the presence of both *Bst*RNA$^{Tyr}_{CUA}$ and *Ec*TyrRS genes (**Supplementary Fig. 1** online). When transfected with pcDNA4-GFP37TAG and pcDNA4-*Ec*TyrRS or pUC18-3*Bst*RNA$^{Tyr}_{CUA}$ alone, cells did not exhibit green fluorescence. These experiments suggest that *Bst*RNA$^{Tyr}_{CUA}$ is charged only by *Ec*TyrRS and that the *Bst*RNA$^{Tyr}_{CUA}$-*Ec*TyrRS pair can function efficiently to suppress a nonsense amber codon in mammalian cells.

Genetic incorporation of six unnatural amino acids

To test the generality of this system, we examined the ability of *Bst*RNA$^{Tyr}_{CUA}$ and six *Ec*TyrRS variants (previously evolved in *S. cerevisiae* to encode pMpa, pApa, pBpa, pIpa, pAzpa and pPpa)[13,28] to incorporate unnatural amino acids into GFP37TAG in both T-REx CHO and T-REx 293 cell lines. We inserted each of the six *Ec*TyrRS genes into pcDNA4/TO/myc-HisA to afford a pcDNA4-*Ec*TyrRS derivative, and verified expression of the aaRS in T-REx CHO and 293 cells by western blot analysis (**Supplementary Fig. 2** online). All six *Ec*TyrRS variants showed similar expression levels in both cell lines. We then tested the ability of the six *Ec*TyrRS variants together with *Bst*RNA$^{Tyr}_{CUA}$ to suppress the amber codon in GFP37TAG in the presence and absence of unnatural amino acids in the growth medium. We transiently transfected cells with plasmids as described for wild-type *Ec*TyrRS. Six hours after transfection, we replaced the medium with fresh medium containing 1 μg/ml tetracycline and supplemented with the corresponding unnatural amino acid (10 mM pMpa, 10 mM pApa, 1 mM pBpa, 8 mM pIpa, 5 mM pAzpa or 1 mM pPpa). We then grew cells for an additional 24 h for T-REx CHO cell lines and 48 h for T-REx 293 cell lines before collecting the cells. In both T-REx CHO and 293 cells, the expression of full length GFP37TAG was dependent upon the presence of unnatural amino acids in the growth medium (**Fig. 1b**,**c**). In the absence of unnatural amino acids, we detected no GFP expression (<1%), indicating that *Ec*TyrRS variants specifically charge *Bst*RNA$^{Tyr}_{CUA}$ with their cognate unnatural amino acids with high fidelity. We detected little or no expression of GFP containing pIpa in T-REx or CHO 293 cells, possibly because of the cellular toxicity of pIpa (in the presence of 8 mM pIpa, T-REx 293 cells die in 6 h).

Because heterologous expression of pMpa-tRNA synthetase (pMpaRS) did not show any apparent cellular toxicity under the growth conditions, we modified the plasmid containing the pMpaRS gene to encode both three tandem repeats of the *Bst*RNA$^{Tyr}_{CUA}$ gene and the pMpaRS gene (pSWAN-pMpaRS; **Fig. 2a**). We inserted the pMpaRS gene directly after the non-regulated CMV promoter for efficient and continuous expression of pMpaRS. We also constructed another plasmid (pSWAN-GFP37TAG; **Fig. 2b**) containing both three tandem repeats of the *Bst*RNA$^{Tyr}_{CUA}$ gene and the GFP37TAG gene. We inserted the gene encoding GFP37TAG after a tetracycline-regulated CMV promoter to minimize potential readthrough of the nonsense amber codon caused by endogenous amber suppression in mammalian cells. We then assayed these two plasmids for their suppression efficiency. Because both plasmids contain three tandem repeats of the *Bst*RNA$^{Tyr}_{CUA}$ gene, varying the ratio of the two plasmids to increase the suppression level was not expected to change the total amount of the *Bst*RNA$^{Tyr}_{CUA}$ gene transfected into cells. Under optimized transfection conditions (0.5 μg of pSWAN-pMpaRS and 2.5 μg pSWAN-GFP37TAG per 2×10^6 cells), the suppression level was roughly twofold higher than that using the three plasmids pUC18-3*Bst*RNA$^{Tyr}_{CUA}$, pcDNA4-pMpaRS, and pcDNA4-GFP37TAG (the suppression level was determined by the number of fluorescent cells; data not shown). In the presence of 10 mM pMpa in the growth medium, 1 μg of mutant GFP can be obtained from 2×10^7 adhesive T-REx CHO cells grown for 1–2 d after induction.

Mass spectroscopy and fidelity analysis

To further characterize GFP-pMpa, we expressed the mutant protein in T-REx CHO cells transiently transfected with both pSWAN-pMpaRS and pSWAN-GFP37TAG. We purified the protein using an anti-myc antibody agarose column and then separated it by SDS-PAGE. We digested the GFP-pMpa band from the gel with trypsin for analysis by nanoscale reversed-phase liquid chromatography/tandem mass spectrometry (nano-RP LC/MS/MS). In parallel, we purified wild-type GFP from cells transfected with pcDNA4-GFP and subjected it to the same analysis. The tandem mass spectra of the tryptic Tyr37 containing fragments of FSVSGEGEGDATY*GK (Y* denotes tyrosine or pMpa; **Fig. 3a**) from the mutant

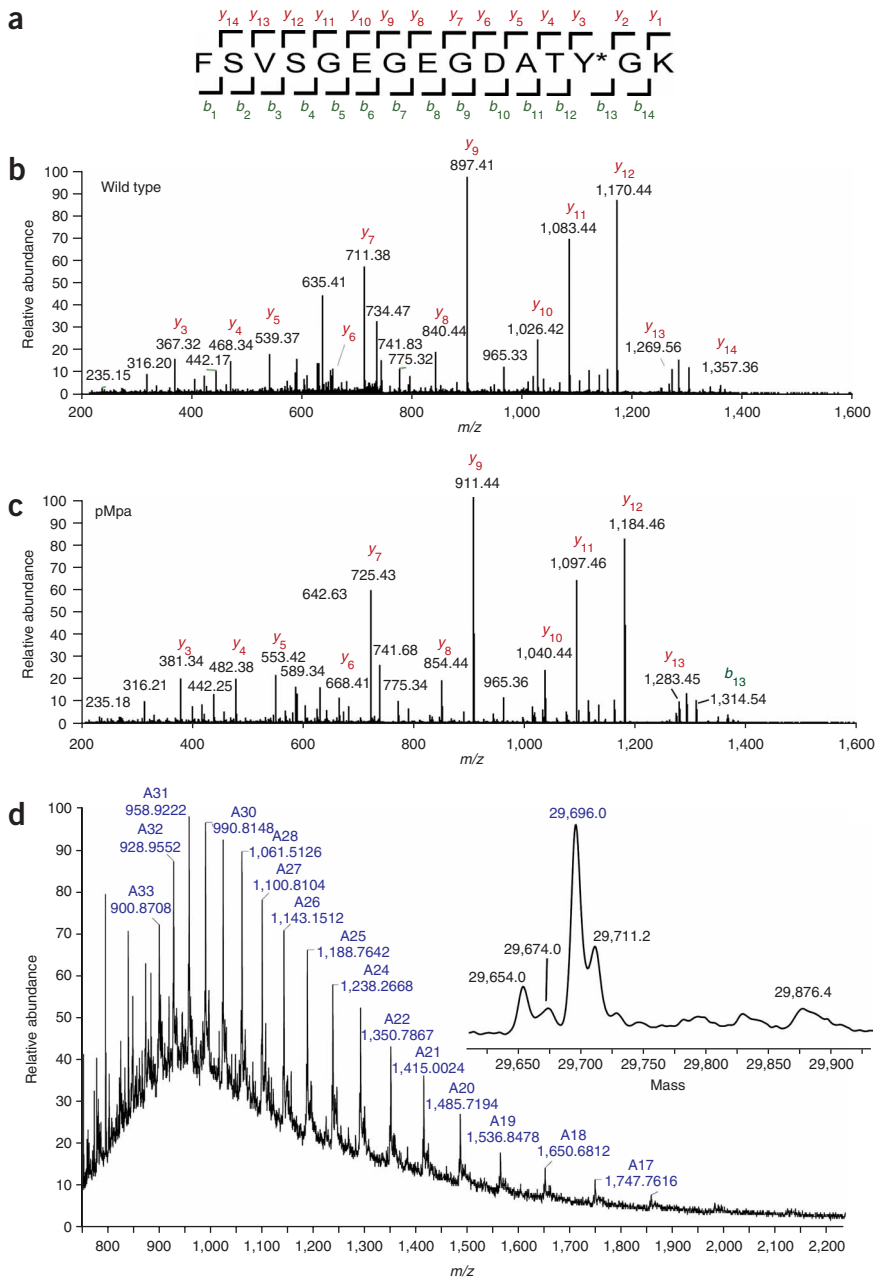

Figure 3 | Annotated tandem MS spectra of the peptide FSVSGEGEGDATY*GK from wild-type GFP and GFP-pMpa. (**a**) *y*- and *b*-type ions generated during fragmentation of the peptide FSVSGEGEGDATY*GK. Y* denotes tyrosine in wild type GFP or pMpa in GFP-pMpa. For comparison and clarity, only the abundant *y*-ion series are annotated as well as the b_{13} ion that locates pMpa unambiguously at position 37. Proteins were purified by anti-myc affinity column and analyzed by SDS-PAGE; the GFP bands were excised for MS analysis. (**b**) Tandem MS spectra of FSVSGEGEGDATYGK from wild-type GFP. (**c**) Tandem MS spectra of FSVSGEGEGDATY*GK from GFP-pMpa. (**d**) Intact protein ESI-TOF MS spectrum of affinity purified GFP-pMpa. Inset, deconvoluted charge-state envelope shows one major component of 29,696.0 Da.

protein revealed intense pMpa peaks, indicating efficient incorporation of pMpa (**Fig. 3b**,**c**). The Y*–containing ions (y_3 to y_{14}) all had an expected mass shift of 14 Da in comparison to wild-type GFP, which matches exactly the mass difference between tyrosine and pMpa. The mass shift of *y* ion series together with the observation of an identical mass shift of the b_{13} ion in **Figure 3c** led us to unambiguously assign the site of pMpa incorporation to position 37 of GFP. We also obtained the ratios of the mass spectrometry (MS) peaks of pMpa containing peptides to those of tyrosine containing peptides. Integration of the single ion chromatograms of the precursor ions of FSVSGEGEGDATY*GK as well as analysis of the three most abundant fragment ions (y_9, y_{11} and y_{12}) in the tandem MS data suggests a high fidelity (>95%) for incorporation of pMpa. To acquire the masses of the parent protein, we subjected the mutant protein purified by anti-myc antibody column to analysis with electrospray ionization–time-of-flight MS (ESI-TOF MS). The theoretical mass of the acetylated mutant protein missing the N-terminal methionine is 29,696.2 Da, which is in good agreement with the major component of 29,696.0 Da observed in the charged-state deconvoluted ESI-TOF MS spectrum (**Fig. 3d**). A smaller feature at 29,654.0 Da was assigned to the mass of GFP-pMpa lacking N-terminal acetylation. We detected no wild-type GFP signals or signals indicating incorporation of multiple pMpa residues. This result further confirmed the selective incorporation of pMpa at position 37 of GFP. (The ESI-TOF MS of wild-type GFP is shown in **Supplementary Fig. 3** online.)

We then constructed plasmids containing pApa-tRNA synthetase (pApaRS), pBpa-tRNA synthetase (pBpaRS), pIpa-tRNA synthetase (pIpaRS), pAzpa-tRNA synthetase (pAzpaRS) and pPpa-tRNA synthetase (pPpaRS) genes to afford pSWAN-*Ec*TyrRS variants. Transient transfection of T-REx CHO and 293 cells with the pSWAN-*Ec*TyrRS variants and pSWAN-GFP37TAG (0.5 μg of pSWAN-*Ec*TyrRS variant and 2.5 μg of pSWAN-GFP37TAG per 2×10^6 cells) all afforded roughly twofold higher suppression levels than for the three plasmids pUC18-3*Bs*tRNA$^{Tyr}_{CUA}$, pcDNA4-*Ec*TyrRS variant and pcDNA4-GFP37TAG (with the exception of the pIpa specific variant). We also expressed mutant GFPs containing pApa, pBpa, pAzpa or pPpa in T-REx CHO cells transfected with the pSWAN-*Ec*TyrRS variants and pSWAN-GFP37TAG, grew the cells in medium containing the corresponding unnatural amino acid (10 mM pApa, 1 mM pBpa, 5 mM pAzpa or 1 mM pPpa) for 1–2 d, and purified the proteins using an anti-myc antibody column. The yields of the mutant proteins containing pApa and pAzpa were close to that for pMpa (~1 μg per 2×10^7 cells), and the yields of mutant proteins containing pBpa and pPpa were somewhat lower (~0.7 μg per 2×10^7 cells). Because the expression levels of *Ec*TyrRS variants in T-REx CHO cells are similar, the

lower yield of the mutant proteins containing pBpa and pPpa may be due to the lower solubility of these unnatural amino acids in the medium. We also subjected the mutant proteins to analysis with nano-RP LC/MS/MS to obtain tandem mass spectra of the tryptic fragments of FSVSGEGEGDATY*GK. Tandem MS data (**Fig. 4**) showed clearly the incorporation of the unnatural amino acids at position 37 of GFP. The mutant GFP containing pAzpa was only weakly detectable. A closer look at the data revealed the presence of the *p*-aminophenylalanine (pAmpa)-containing peptide instead (**Supplementary Fig. 4** online), which is not surprising considering the chemical reactivity and photoinstability of the azido group (pAmpa has previously been observed from MS analysis of pAzpa-containing peptides[14]). We identified a trace amount of the wild-type peptide in the pBpa sample, but the signals were too weak for accurate quantitation of the pBpa/tyrosine ratio. We detected no wild-type signals in the pApa, pAzpa and pPpa samples. The data from all five mutant proteins indicate the high selectivity and fidelity of the incorporation of unnatural amino acids.

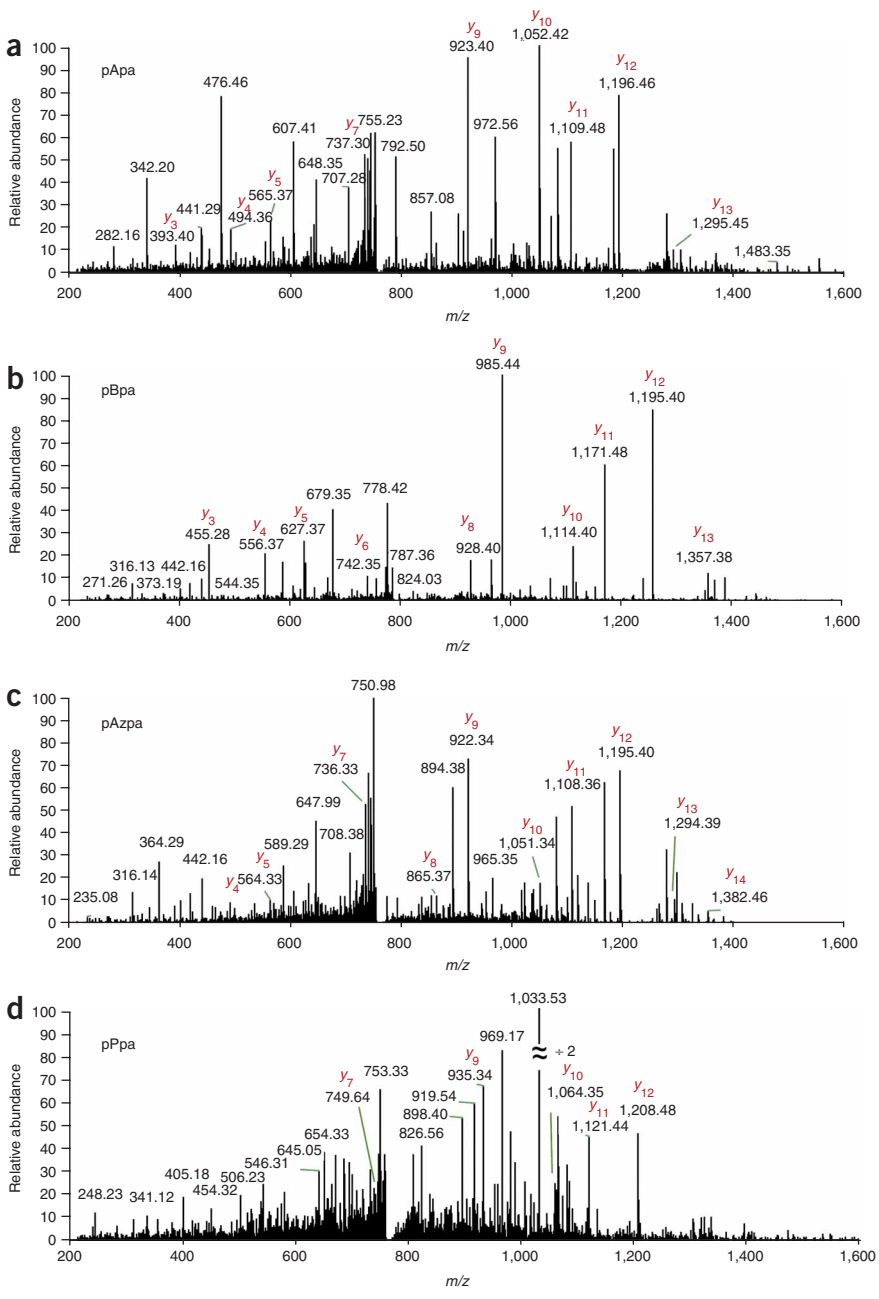

Figure 4 | Annotated tandem MS spectra of the peptide FSUSGEGEGDATY*GK from GFP-pApa, GFP-pBpa, GFP-pAzpa and GFP-pPpa. (**a**) Tandem MS spectra of GFP-pApa. The Y*-ions exhibit a 26-Da mass shift with respect to the same ions in the spectrum of wild-type GFP. (**b**) Tandem MS spectrum of GFP-pBpa. The characteristic mass shift of 86 Da between mutant GFP and wild-type GFP is clearly observed. (**c**) Tandem MS spectrum of GFP-pAzpa. Most of pAzpa decays to pAmpa (see **Supplementary Fig. 4** for tandem MS spectra of tryptic pAmpa containing fragment). However, the characteristic mass shift of 25 Da of Y*-ions in comparison to signals of wild-type GFP is still clearly observed. (**d**) Tandem MS spectrum of GFP-pPpa. The signals of the pPpa fragment are largely swamped by background. Despite this it is still identified with good discrimination by database searching of the spectrum against MSDB. The stronger Y*-ions are observed with a characteristic mass shift of 38 Da.

DISCUSSION

In the mammalian genome, the occurrence of amber stop codons is higher (23% in humans), in comparison to *E. coli* (7%). Therefore, amber suppression might be toxic to cells if essential proteins are not terminated correctly. Yokoyama and co-workers showed that inducible expression of the mutant *Ec*TyrRS that charges its cognate tRNA with 3-iodo-L-tyrosine minimizes possible cellular toxicity resulting from background incorporation of endogenous tyrosine by the mutant aaRS in the absence of 3-iodo-L-tyrosine[2]. The system described here further abrogates this problem. All *Ec*TyrRS variants were evolved from a two-step positive/negative selection scheme in *S. cerevisiae*, which removes aaRS variants that incorporate endogenous amino acids. Cotransfection of cells with a pSWAN-*Ec*TyrRS variant and pSWAN-GFP37TAG did not lead to observable readthrough of the nonsense amber codon in GFP37TAG in the absence of unnatural amino acids, suggesting that the tRNA-aaRS pairs do not efficiently suppress natural TAG stop codons in the absence of unnatural amino acids. Cell lines stably expressing tRNA and aaRS proteins should therefore be viable in the absence of unnatural amino acids. Because we observed no endogenous amber suppression, the expression of the target protein containing unnatural amino acids also should not require induction. Creation of stable cell lines maintaining the tRNA, aaRS and target protein genes should therefore allow efficient production of the target protein containing the unnatural amino acid when cells are supplemented with the unnatural amino acid.

The extension of this methodology to mammalian cells should facilitate cellular studies using biological probes. Ultimately,

it may allow the synthesis of therapeutic proteins containing unnatural amino acids in mammalian systems. Presently we are extending these studies to other unnatural amino acids including photocaged, fluorescent and chemically reactive amino acids, other cell types, and possibly even multicellular organisms.

METHODS

Transfection and western blot analysis. Both T-REx chinese hamster ovary (CHO) and T-REx 293 cells (Invitrogen) constitutively express the tetracycline repressor, which regulates the expression of genes inserted into the pcDNA4/TO/myc-HisA plasmid. We grew T-REx CHO cells in F-12, 10% FBS, 1% pencillin-streptomycin (Pen-Strep), 2 mM L-glutamine and 10 μg/ml blasticidin (Invitrogen) at 37 °C in a humidified atmosphere of 5% CO_2; we grew T-REx 293 cells in Gibco D-MEM medium (Invitrogen), 10% FBS, 1% Pen-Strep, 2 mM L-glutamine and 5 μg/ml blasticidin at 37 °C in a humidified atmosphere of 5% CO_2. We grew cells to 80–90% confluency in Costar 6-well cell-culture clusters and then transfected the cells with plasmids using Fugene 6 (Roche; 9 μl Fugene + 3 μg plasmids). We purchased pMpa, pBpa and pAzpa from Bachem, Inc., and pIpa from Synchem, Inc.

For western blot analysis, we lysed the collected cells in RIPA buffer (Upstate) with a 1:100 dilution of protease inhibitor cocktail (Sigma). We fractionated the supernatant by SDS-PAGE under denaturing conditions and transferred it to a 0.45-μm nitrocellulose membrane (Invitrogen). For T-REx CHO cells, we probed the proteins immobilized on the membrane with anti-c-myc (Invitrogen, 1:5,000 dilution) as the primary antibody and anti-mouse IgG-HRP (Invitrogen) as the secondary antibody. We detected chemiluminescence with the Pierce ECL western blotting substrate. For T-REx 293 cells, we probed the membranes with anti-His-HRP (Invitrogen, 1:5,000 dilution), followed by detection with ECL western blotting substrate.

Additional methods. Descriptions of plasmid construction, protein expression and purification, and mass spectrometric analysis are available in **Supplementary Methods** online.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This work was funded by a US National Institutes of Health grant GM62159. This is manuscript 18455 of the Scripps Research Institute.

AUTHOR CONTRIBUTIONS

W.L. developed and evaluated the method, A.B. carried out the MS analysis, Shou C. and Shuibing C. helped with protein characterization, and W.L. and P.G.S. designed the project, analyzed the data and prepared the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Monahan, S.L., Lester, H.A. & Dougherty, D.A. Site-specific incorporation of unnatural amino acids into receptors expressed in mammalian cells. *Chem. Biol.* **10**, 573–580 (2003).
2. Sakamoto, K. *et al.* Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* **30**, 4692–4699 (2002).
3. Zhang, Z. *et al.* Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proc. Natl. Acad. Sci. USA* **101**, 8882–8887 (2004).
4. Zhang, D., Vaidehi, N., Goddard, W.A., III, Danzer, J.F. & Debe, D. Structure-based design of mutant *Methanococcus jannaschii* tyrosyl-tRNA synthetase for incorporation of O-methyl-L-tyrosine. *Proc. Natl. Acad. Sci. USA* **99**, 6579–6584 (2002).
5. Turner, J.M., Graziano, J., Spraggon, G. & Schultz, P.G. Structural characterization of a p-acetylphenylalanyl aminoacyl-tRNA synthetase. *J. Am. Chem. Soc.* **127**, 14976–14977 (2005).
6. Turner, J.M., Graziano, J., Spraggon, G. & Schultz, P.G. Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. USA* **103**, 6483–6488 (2006).
7. Kiga, D. *et al.* 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 (2002).
8. Wang, L., Jackson, W.C., Steinbach, P.A. & Tsien, R.Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci. USA* **101**, 16745–16749 (2004).
9. Xie, J. & Schultz, P.G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **7**, 775–782 (2006).
10. Wang, L., Brock, A., Herberich, B. & Schultz, P.G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498–500 (2001).
11. Santoro, S.W., Anderson, J.C., Lakshman, V. & Schultz, P.G. An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli*. *Nucleic Acids Res.* **31**, 6700–6709 (2003).
12. Anderson, J.C. *et al.* An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. USA* **101**, 7566–7571 (2004).
13. Chin, J.W. *et al.* An expanded eukaryotic genetic code. *Science* **301**, 964–967 (2003).
14. Chin, J.W., Cropp, T.A., Chu, S., Meggers, E. & Schultz, P.G. Progress toward an expanded eukaryotic genetic code. *Chem. Biol.* **10**, 511–519 (2003).
15. Wu, N., Deiters, A., Cropp, T.A., King, D. & Schultz, P.G. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* **126**, 14306–14307 (2004).
16. Summerer, D. *et al.* A genetically encoded fluorescent amino acid. *Proc. Natl. Acad. Sci. USA* **103**, 9785–9789 (2006).
17. Wang, J., Xie, J. & Schultz, P.G. A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* **128**, 8738–8739 (2006).
18. Zhang, Z. *et al.* A new strategy for the synthesis of glycoproteins. *Science* **303**, 371–373 (2004).
19. Liu, C. & Schultz, P.G. Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat. Biotechnol.* **24**, 1436–1440 (2006).
20. Alfonta, L., Zhang, Z., Uryu, S., Loo, J.A. & Schultz, P.G. Site-specific incorporation of a redox-active amino acid into proteins. *J. Am. Chem. Soc.* **125**, 14662–14663 (2003).
21. Wang, L. & Schultz, P.G. Expanding the genetic code. *Angew. Chem. Int. Edn Engl.* **44**, 34–66 (2004).
22. Bonnefond, L., Giege, R. & Rudinger-Thirion, J. Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems. *Biochimie* **87**, 873–883 (2005).
23. Hino, N. *et al.* Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Methods* **2**, 201–206 (2005).
24. Sprague, K.U. Transcription of eukaryotic tRNA genes (AMS Press, Washington, DC, 1994).
25. Bedouelle, H. Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. *Biochimie* **72**, 589–598 (1990).
26. Hou, Y.M. & Schimmel, P. Modeling with *in vitro* kinetic parameters for the elaboration of transfer RNA identity *in vivo*. *Biochemistry* **28**, 4942–4947 (1989).
27. Ormo, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).
28. Deiters, A. *et al.* Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **125**, 11782–11783 (2003).