

# Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion

Kaihang Wang<sup>1,2</sup>, Heinz Neumann<sup>1,2</sup>, Sew Y Peak-Chew<sup>1</sup> & Jason W Chin<sup>1</sup>

***In vivo* incorporation of unnatural amino acids by amber codon suppression is limited by release factor-1-mediated peptide chain termination. Orthogonal ribosome-mRNA pairs function in parallel with, but independent of, natural ribosomes and mRNAs. Here we show that an evolved orthogonal ribosome (ribo-X) improves tRNA<sub>CUA</sub>-dependent decoding of amber codons placed in orthogonal mRNA. By combining ribo-X, orthogonal mRNAs and orthogonal aminoacyl-tRNA synthetase/tRNA pairs in *Escherichia coli*, we increase the efficiency of site-specific unnatural amino acid incorporation from ~ 20% to >60% on a single amber codon and from <1% to >20% on two amber codons. We hypothesize that these increases result from a decreased functional interaction of the orthogonal ribosome with release factor-1. This technology should minimize the functional and phenotypic effects of truncated proteins in experiments that use unnatural amino acid incorporation to probe protein function *in vivo*.**

The genetic code of prokaryotic and eukaryotic organisms has been expanded to allow *in vivo*, site-specific incorporation of over 20 designed unnatural amino acids, including post-translational modifications, chemical and photochemical handles, and fluorophores<sup>1</sup>. This is accomplished by endowing organisms with evolved orthogonal aminoacyl-tRNA synthetase/tRNA<sub>CUA</sub> pairs that direct site-specific incorporation of an unnatural amino acid in response to an amber codon<sup>2–5</sup>. The orthogonal aminoacyl-tRNA synthetase aminoacylates a cognate orthogonal tRNA, but no other cellular tRNAs, with an unnatural amino acid, and the orthogonal tRNA is a substrate for the orthogonal synthetase but is not substantially aminoacylated by any endogenous aminoacyl-tRNA synthetase. Genetic code expansion in *E. coli* using evolved variants of the orthogonal *Methanococcus jannaschii* tyrosyl-tRNA synthetase/tRNA<sub>CUA</sub> pair greatly increases the yield of proteins containing unnatural amino acids as compared with methods that rely on the addition of stoichiometrically pre-aminoacylated suppressor tRNAs to cells<sup>6–9</sup> or on *in vitro* translation reactions<sup>10–14</sup>. However, because release factor-1 (RF-1)-mediated peptide chain termination<sup>15–17</sup> competes with tRNA<sub>CUA</sub>-mediated chain elongation, the efficiency of unnatural amino acid incorporation in proteins synthesized from genes containing single amber stop codons is limited to 20–30%. Moreover, as the number of amber stop codons in a gene increases, the efficiency of unnatural amino acid incorporation decreases multiplicatively: for example, less than one-tenth of protein synthesis initiated on a gene containing two amber codons typically reaches completion.

Many potential applications of unnatural amino acid mutagenesis would benefit from more efficient unnatural amino acid incorporation. For example, more efficient single-site incorporation of photo-crosslinkable amino acids would minimize the production of

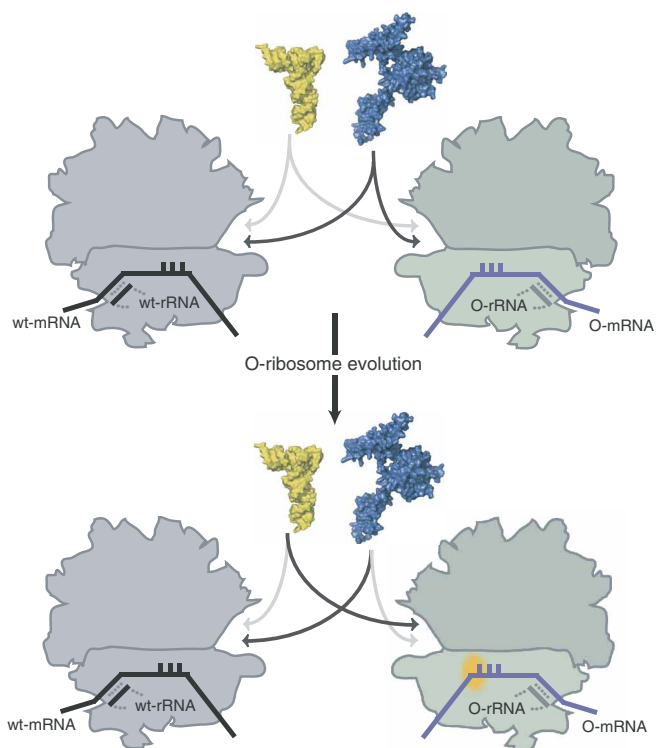
truncated proteins that may interfere with experiments directed at mapping protein-protein interactions *in vivo*<sup>18–20</sup>. In addition, the ability to incorporate multiple amino acids with post-translational modifications (e.g., methylation, acetylation and phosphorylation) would be useful for deciphering post-translational signaling codes.

Unnatural amino acid incorporation in *in vitro* translation reactions can be increased by using S30 extracts that contain a thermally inactivated mutant of RF-1 (ref. 21). Temperature-sensitive mutants of RF-1 allow transient increases in global amber suppression *in vivo*<sup>22</sup>, but RF-1 knockouts are lethal and therefore cannot be used for this purpose. Increases in tRNA<sub>CUA</sub> gene copy number and a transition from minimal to rich media<sup>4,23</sup> improve the yield of unnatural amino acid incorporation in *E. coli*, although the efficiency (defined as the ratio of full-length protein to truncated protein) does not exceed 20–30%. Moreover, this strategy indiscriminately increases suppression of all amber codons in the cell and therefore enhances the read-through of stop codons on chromosomal genes (320 genes in *E. coli* terminate in UAG<sup>24</sup>, including 44 essential genes); by interfering with cellular protein synthesis, this approach potentially disturbs cellular physiology. Ideally, efficiency could be maximized by minimizing the effects of RF-1-mediated chain termination while leaving the decoding of chromosomal amber stop codons unaltered. We conceived of using recently described orthogonal ribosome-mRNA pairs<sup>25</sup>, that extend the concept of specialized ribosomes<sup>26</sup>, as previously described<sup>25</sup>, to address this challenge (Fig. 1).

Orthogonal ribosome (O-ribosome)orthogonal mRNA (O-mRNA) pairs are composed of an mRNA containing a ribosome-binding site that does not direct translation by the endogenous ribosome, and an orthogonal ribosome that efficiently and specifically translates the orthogonal mRNA, but does not appreciably translate cellular

<sup>1</sup>Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England, UK. <sup>2</sup>These authors contributed equally to this work. Correspondence should be addressed to J.W.C. (chin@mrc-lmb.cam.ac.uk).

Received 20 February; accepted 23 May; published online 24 June 2007; doi:10.1038/nbt1314



**Figure 1** Diverging the decoding properties of natural and orthogonal ribosomes. The natural ribosome (gray) and the progenitor orthogonal ribosome (green) decode wt- (black) and orthogonal- (purple) mRNAs, respectively. Because RF-1 (blue) competes efficiently (dark gray arrows) for UAG codons in the A-site of both ribosomes, amber suppressor tRNAs (yellow), that may be uniquely aminoacylated with an unnatural amino acid, are decoded with equal and low efficiency (light gray arrows) on both ribosomes. Synthetic evolution of the orthogonal ribosome leads to an evolved scenario in which a mutant (orange patch) orthogonal ribosome more efficiently decodes amber suppressor tRNAs within the context of orthogonal mRNAs. Decoding of natural mRNAs is unaffected because the orthogonal ribosome does not read natural mRNAs and the natural ribosome is unaltered. Surface structure figures were created using Pymol v0.99 (<http://www.pymol.org/>) and PDB IDs 2B64 and 1J1U.

mRNAs<sup>25</sup>. Unlike the natural ribosome, the orthogonal ribosome does not appreciably synthesize the proteome and is therefore tolerant to mutations in the highly conserved rRNA that would cause lethal or dominant negative effects in the natural ribosome<sup>27</sup>. To increase the *in vivo* efficiency of unnatural amino acid incorporation, we evolved orthogonal ribosomes that show an increased tRNA<sub>CUA</sub>-dependent amber suppression in the context of orthogonal mRNAs.

## RESULTS

### Design of a ribosome A-site library

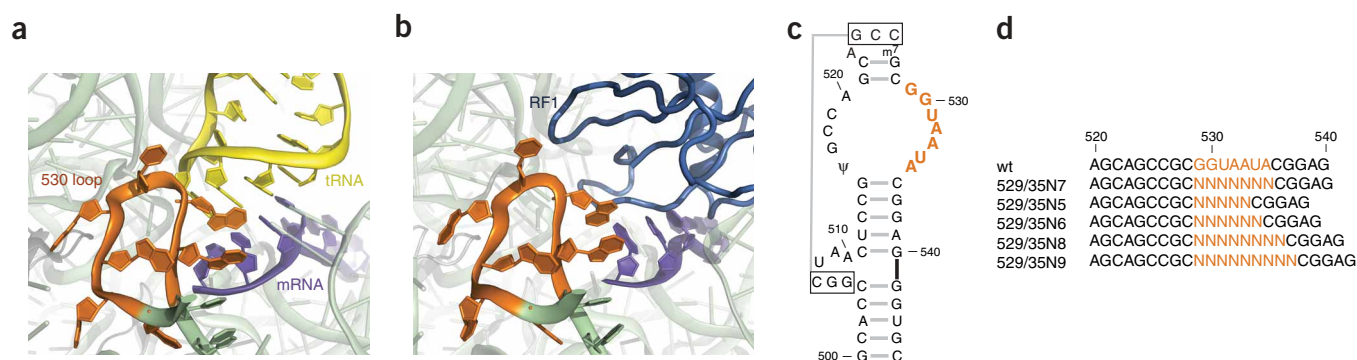
The A-site of the ribosome is the gateway to the tRNA translocation corridor composed of the A, P and E sites<sup>28–30</sup>. In response to an amber codon in the A-site of the ribosome, RF-1 can bind to the A-site and compete with decoding of amber suppressor tRNAs. We reasoned

that combinations of mutations in 16S rRNA, which forms the A-site of the ribosome, might yield a variant A-site that allows amber suppressor tRNAs to compete more effectively with RF-1 for A-site binding, and thus favor amber suppression and elongation over polypeptide chain termination.

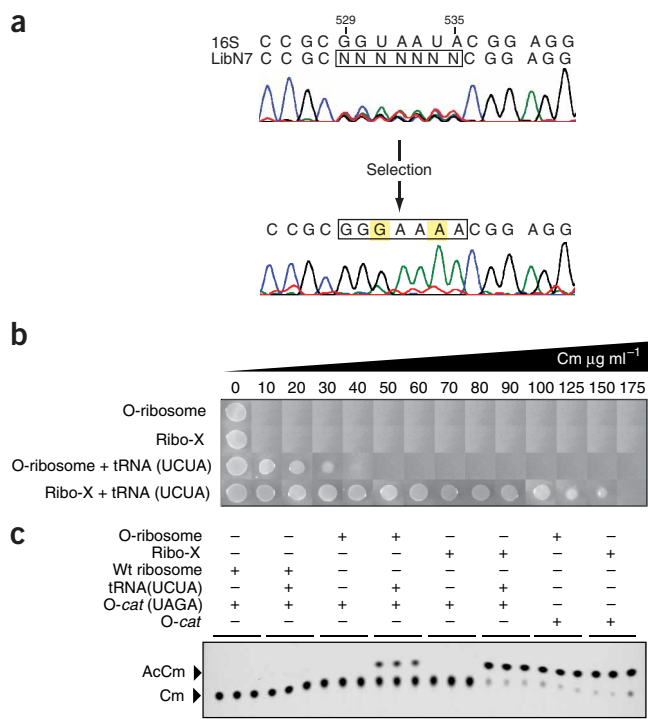
To design an A-site library (Fig. 2), we examined the structures of tRNAs or RF-1 bound to the ribosomal A-site<sup>29–32</sup>. These structures show that the 530 loop in 16S rRNA is proximal to both substrates (Fig. 2). We reasoned that combinations of mutations in the 530 loop might maintain tRNA binding, but decrease functional interaction with RF-1 in the presence of a UAG codon. We therefore randomized seven nucleotides (529–535) in the 530 loop to all combinations of nucleotides to create an N7 library. Moreover, we created longer and shorter 530 loop–sequence libraries (N5, N6, N8 and N9 (Fig. 2)) to expand the functional space sampled. All libraries were >99% complete (Supplementary Table 1 online).

### Selection of evolved decoding in orthogonal ribosomes

To create a selection system for orthogonal ribosomes that more efficiently read amber codons, we required a reporter of orthogonal ribosome activity that contains selector codons. We decided to work initially with a UAGA-containing reporter and tRNA<sup>ser2</sup> (UCUA) (Supplementary Fig. 1 online), which is aminoacylated by seryl-tRNA synthetase<sup>33</sup>, rather than a simple UAG suppressor, because it allows selection for improved ribosome activity over a larger dynamic range. Cells containing O-*cat* (UAGA103, UAGA146)/tRNA<sup>ser2</sup> (UCUA) and the O-ribosome had an IC<sub>50</sub> on chloramphenicol of 25 μg ml<sup>-1</sup>. For comparison, the O-*cat* reporter devoid of UAGA



**Figure 2** Design of ribosome-decoding libraries. (a) Structure of a tRNA anticodon stem loop (yellow) bound to mRNA (purple) in the A-site of the ribosome (green). The 530 loop is shown in orange. (b) Structural model of RF-1 (blue) bound in the A-site of the ribosome. (c) Secondary structure of the 530 loop. The region targeted for mutation is colored orange. The boxed sequences form a pseudo knot,  $\psi$  is pseudouridine and m<sup>7</sup>G is 7-methyl guanosine. (d) The sequence of ribosome-decoding libraries. 3D structure figures were created using Pymol v0.99 (<http://www.pymol.org/>) and PDB IDs 11BM and 2B64. 2D structure figures were adapted from <http://www.rna.cccb.utexas.edu/>.



codons supports growth on 500  $\mu\text{g ml}^{-1}$  of chloramphenicol in the presence of the O-ribosome.

To select mutant ribosomes that more efficiently decode the UAGA codon we combined each orthogonal-ribosome library with O-*cat* (UAGA103, UAGA146)/tRNA<sup>ser2</sup>(UCUA) (in which *cat* containing UAGA codons is translated from an orthogonal-ribosome binding site, **Supplementary Fig. 2** online) and challenged the cells to grow on chloramphenicol concentrations at which the O-ribosome does not support growth. No clones containing insertions or deletions survived for libraries N5, N6, N8 and N9, suggesting that the 530 loop is intolerant to longer or shorter sequences of any composition. However, clones from the N7 library did survive on 100  $\mu\text{g ml}^{-1}$  chloramphenicol. All ten clones sequenced from the N7 selection were identical, and contained the mutations U531G and U534A in 16S rRNA (**Fig. 3**). The U531G mutation is present in only two sequenced vertebrate mitochondrial rRNAs, whereas the U534A mutation is present in 0.2% of bacterial rRNAs. No sequenced natural ribosome contains this combination of mutations<sup>34</sup>. Because it remained a formal possibility that more efficient ribosomes for UAG decoding existed in the libraries but were not captured by the UAGA selection, we repeated the selection with reporters containing the UAG codon and an amber suppressor derived from tRNA<sup>ser2</sup> (**Supplementary Fig. 1**). We found that the same sequence was uniquely selected. We therefore decided to characterize the U531G, U534A mutant ribosome, which we refer to as ribo-X, in more detail.

### Ribo-X enhances tRNA-dependent UAGA decoding

To measure the extent to which ribo-X enhances tRNA<sup>ser2</sup>(UCUA)-dependent decoding of the UAGA codon, we compared the chloramphenicol resistance of cells containing O-*cat* (UAGA103, UAGA146)/tRNA<sup>ser2</sup>(UCUA) and either ribo-X or the progenitor O-ribosome (**Fig. 3**). Ribo-X-containing cells survive on concentrations of chloramphenicol five times higher than cells containing the progenitor O-ribosome. The enhanced, tRNA<sup>ser2</sup>(UCUA)-dependent, quadruplet-

**Figure 3** Selection and phenotypic characterization of ribo-X. **(a)** The selection of ribosomes that decode UAGA and UAG codons, using cognate tRNAs derived from tRNA<sup>ser2</sup>. The top sequence trace shows the 16S rDNA before selection, whereas the lower traces show the convergence of the pool (sequencing from all colonies surviving a selection on chloramphenicol plates). **(b)** The ribo-X, tRNA<sup>ser2</sup>(UCUA)-dependent enhancement in decoding UAGA codons in the O-*cat* (UAGA103, UAGA146) gene measured by survival on chloramphenicol. **(c)** As in **b**, but measuring CAT activity directly. Thin-layer chromatography showing the acetylation of chloramphenicol (Cm) to acetylated chloramphenicol (AcCm) by CAT produced from either O-*cat* or O-*cat* (UAGA103, UAGA146).

decoding was further confirmed by *in vitro* chloramphenicol acetyl transferase (CAT) assays<sup>35</sup> (**Fig. 3**). Similar ribo-X-mediated enhancements were observed using tRNA<sup>ser2</sup>(CUA) and a cognate O-*cat* reporter (data not shown). The chloramphenicol resistance conferred on cells by ribo-X and O-*cat* and the progenitor O-ribosome and O-*cat* is identical (500  $\mu\text{g ml}^{-1}$ ), indicating that ribo-X is efficient and processive in translation of sense codons.

### Ribo-X and natural ribosomes have comparable fidelity

To demonstrate that ribo-X synthesizes proteins with a fidelity comparable to the natural ribosome, we compared the mass spectra and amino acid misincorporation frequency of proteins synthesized by wild-type ribosomes, the progenitor orthogonal ribosome and ribo-X.

Expression of O-*gst-male* (a genetic fusion between the genes encoding glutathione-S-transferase (GST) and maltose-binding protein (MBP) driven by an orthogonal-ribosome binding site) in the presence of ribo-X or the progenitor O-ribosome produced GST-MBP with a purified yield of 30–40 mg/l, comparable to the yield of GST-MBP produced from a *gst-male* fusion by wild-type ribosomes. As expected, no GST-MBP can be purified from O-*gst-male* in the absence of orthogonal ribosomes (**Fig. 4a**). Thrombin cleavage of GST-MBP, at a site between GST and MBP in the protein fusion (**Supplementary Fig. 3** online), produced two proteins that were amenable to mass determination by electrospray ionization mass-spectrometry. Proteins produced from each ribosome had the same mass (**Fig. 4b**).

To explicitly compare the translational fidelity of ribo-X to that of progenitor ribosomes, we measured the frequency of <sup>35</sup>S-cysteine misincorporation<sup>36</sup> into MBP, which contains no cysteine codons (**Fig. 4c** and **Supplementary Fig. 4** online). The error frequency per codon translated by ribo-X was less than  $1 \times 10^{-3}$ . Control experiments with the progenitor orthogonal ribosome and the wild-type ribosome allowed us to put the same limit on their error frequency. This limit compares well with a previous measurement for amino acid misincorporation frequency, as measured by <sup>35</sup>S-cysteine misincorporation<sup>36</sup>, of  $4 \times 10^{-3}$  errors per codon. To further probe the translational fidelity of ribo-X with respect to defined perturbations in the codon-anticodon interaction, we took advantage of a dual luciferase reporter system (DLR) that has previously been used to measure the fidelity of natural and error-prone ribosomes in decoding near-cognate and noncognate codons<sup>37</sup>. We created a DLR with an orthogonal-ribosome binding site (O-DLR), and demonstrated that its translation is dependent on the presence of a cognate orthogonal ribosome (**Supplementary Fig. 5** online). We translated O-DLR variants, for which K529(AAA) was mutated at each position of the codon-anticodon interaction (**Fig. 4d**), using ribo-X or the progenitor orthogonal ribosome, and compared the resulting luciferase activities as a measure of translational misreading. We found that the fidelity of ribo-X was at least as good as that measured for the progenitor orthogonal ribosome and the natural ribosome across all four codon-anticodon

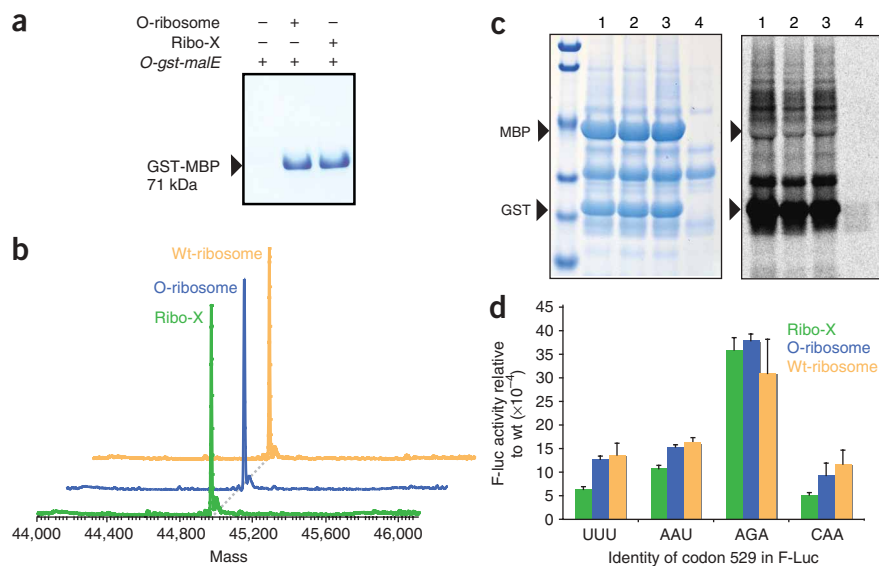
**Figure 4** The translational fidelity of ribo-X is comparable to that of the natural ribosome.

(a) Translation from *O-gst-malE* is dependent on ribo-X or the O-ribosome. (b) Ribo-X synthesizes proteins of identical composition to those synthesized by the wild-type ribosome, as judged by electrospray ionization mass spectrometry.

The electrospray ionization spectra of MBP synthesized by ribo-X, the progenitor O-ribosome or the wild-type ribosome is shown. Each ribosome was used to synthesize the GST-MBP protein, which was purified on glutathione sepharose and subject to thrombin cleavage at a site in the linker (Supplementary Fig. 3). The resulting pairs of fragments have identical electrospray ionization spectra (Found: O-ribosome 44,984 Da, ribo-X 44,984 Da, wt ribosome 44,984 Da, expected 44,981 Da).

(c) The translational error frequency measured by <sup>35</sup>S-cysteine misincorporation is indistinguishable for ribo-X and the natural ribosome. GST-MBP was synthesized by each ribosome in the presence of <sup>35</sup>S-cysteine, purified on glutathione sepharose and digested with thrombin. The left panel shows a Coomassie stain of the thrombin digest. The unannotated bands result primarily from the thrombin preparation. The right panel shows <sup>35</sup>S labeling of proteins in a similar gel, imaged using a Storm Phosphorimager; a Coomassie stain of the <sup>35</sup>S gel is shown in Supplementary Figure 4.

Lanes 1–3 show thrombin cleavage reactions of purified protein derived from cells containing pSC101\*-ribo-X & pO-*gst-malE*, pSC101\*-O-ribosome and pO-*gst-malE*, and pSC101\*-BD and p*gst-malE*. Lane 4 is a negative control in which cells lacking a *gst-malE* gene fusion were treated identically to the other samples. The size markers are pre-stained standards (Bio-Rad 161-0305). (d) The translational fidelity of ribo-X is comparable to that of the natural ribosome as measured by a dual luciferase assay. In this system a C-terminal firefly luciferase is mutated at codon K529(AAA), which codes for an essential lysine residue. The extent to which the mutant codon is misread by tRNA<sup>Lys</sup>(UUU) is determined by comparing the firefly luciferase activity resulting from the expression of the mutant gene to the wild-type firefly luciferase, and normalizing any variability in expression using the activity of the cotranslated N-terminal *Renilla* luciferase. Previous work has demonstrated that measured firefly luciferase activities in this system result primarily from the synthesis of a small amount of protein that misincorporates lysine in response to the mutant codon, rather than a low activity resulting from the more abundant protein containing encoded mutations<sup>37</sup>. In experiments examining the fidelity of ribo-X, lysate from cells containing pSC101\*-ribo-X and pO-DLR and its codon 529 variants were assayed. Control experiments used lysates from cells containing pSC101\*-O-ribosome and pO-DLR and its codon 529 variants or pwt-DLR and its variants. Assays on pO-DLR in the presence and absence of the orthogonal ribosome or ribo-X indicate that >98% of translation on pO-DLR is derived from ribo-X or the orthogonal ribosome (Supplementary Fig. 5), confirming that the fidelity measurements on pO-DLR reflect the activity of ribo-X.



interactions tested. Overall, the mass spectra, <sup>35</sup>S misincorporation assay and dual luciferase assays demonstrated that ribo-X has a translational fidelity comparable to that of the natural ribosome.

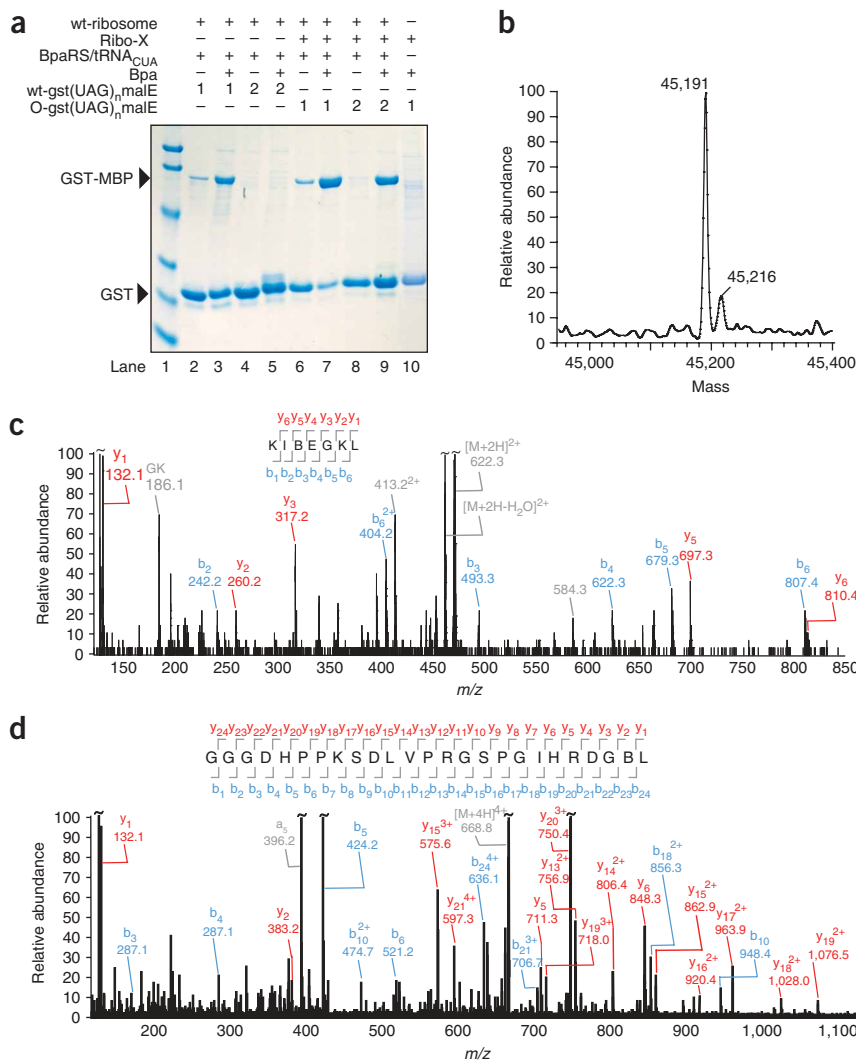
### Increased efficiency unnatural amino acid incorporation

To demonstrate the substantial increase in efficiency of site-specific unnatural amino acid incorporation with ribo-X, we chose to work with the photocrosslinking amino acid *p*-benzoyl-L-phenylalanine (Bpa)<sup>38</sup>. This amino acid has been added to the genetic code of *E. coli*, yeast and mammalian cells and used extensively to map the topology of protein-protein interactions *in vitro* and *in vivo*<sup>5,20,39–42</sup>.

We expressed *gst(UAG)malE* in the presence of a *p*-benzoyl-L-phenylalanyl-tRNA synthetase/tRNA<sub>CUA</sub> (BpaRS/tRNA<sub>CUA</sub>) pair<sup>39</sup> (evolved from the *MjTyrRS*/tRNA<sub>CUA</sub> pair), Bpa and wild-type ribosomes (Fig. 5a). As expected this produced GST-MBP, incorporating Bpa, with low efficiency (24%). However, when we synthesized GST-MBP containing Bpa from *O-gst(UAG)malE*, using the BpaRS/tRNA<sub>CUA</sub> pair, Bpa and ribo-X, the efficiency increased to 62%. As expected, based on the previously reported specificity of BpaRS, full-length protein synthesis is Bpa-dependent<sup>39</sup>. In our experiments, performed in Luria-Bertani (LB) medium, as previously described<sup>23</sup>, we see a small amount of full-length protein synthesis that is BpaRS- and tRNA<sub>CUA</sub>-dependent, but not amino acid-dependent (compare Fig. 5a, lanes 2, 6 and 10). This effect is minimized in minimal medium, where the total yields of overexpressed proteins are also approximately five times lower (Supplementary Fig. 6 online). In the

presence of Bpa, the aminoacylation of natural amino acids onto tRNA<sub>CUA</sub> by noncognate aminoacyl-tRNA synthetases observed in rich media is outcompeted, and incorporation of Bpa is quantitative<sup>23</sup> (Fig. 5b–d). Mass spectrometry shows that BpaRS expression from pSupBpa does not lead to detectable levels of unnatural amino acid incorporation in response to sense codons (by misacylation of endogenous tRNAs), as expected from the observation that *MjTyrRS* does not aminoacylate any *E. coli* tRNAs with tyrosine, even in the absence of competing endogenous aminoacyl-tRNA synthetase enzymes<sup>43</sup>. In the absence of a functional aminoacyl-tRNA synthetase/tRNA<sub>CUA</sub> pair, ribo-X terminates translation on the amber codon, and no full-length GST-MBP fusion is purified (Fig. 5a, lane 10). Similarly, ribo-X does not measurably enhance read-through of a UAA or UGA codon (data not shown).

The ribo-X-mediated enhancement of efficiency was even more dramatic for a gene containing two amber stop codons (Fig. 5a). Wild-type ribosomes produced GST-MBP containing two Bpas from *gst(UAG)<sub>2</sub>malE* with an efficiency of <1%, whereas ribo-X produced GST-MBP containing two Bpas with an efficiency at least 20-fold higher (22%) from *O-gst(UAG)<sub>2</sub>malE*. Extrapolation of the single UAG efficiencies to two sites predicts efficiencies of 38% and 6% for ribo-X and the wild-type ribosome, respectively. Comparison of the ratio of predicted-to-observed efficiencies for each ribosome suggests that ribo-X may be more robust than the wild-type ribosome to context effects that decrease the efficiency of UAG suppression<sup>44</sup>. Electrospray ionization mass spectrometry of MBP produced by



**Figure 5** Ribo-X enhances the efficiency of BpaRS/tRNA<sub>CUA</sub>-dependent unnatural amino acid incorporation in response to single and double UAG codons. **(a)** In each lane an equal volume of protein purified from glutathione sepharose under identical conditions is loaded. Ribo-X is produced from pSC101<sup>\*</sup>-ribo-X derived rRNA. Bpa, *p*-benzoyl-L-phenylalanine. BpaRS, *p*-benzoyl-L-phenylalanyl-tRNA synthetase. BpaRS/tRNA<sub>CUA</sub> are produced from pSUPBpa<sup>23</sup> that contains six copies of *MjtRNA*<sub>CUA</sub> and is the most efficient unnatural amino incorporation vector reported to date. (UAG)<sub>n</sub> describes the number of stop codons (n) between *gst* and *malE* in *O-gst(UAG)<sub>n</sub>malE* or *gst(UAG)<sub>n</sub>malE*. Lane 10 is from a different gel. The markers are as described in **Figure 3**. **(b)** The mass of protein expressed from *O-gst(UAG)<sub>2</sub>malE* by *ribo-X* is as expected for the incorporation of 2 Bpas. Purified full-length protein was cleaved with thrombin to produce an MBP fragment amenable to accurate mass determination. The found mass (45,191) is identical to the expected mass for incorporation of two Bpas into MBP (45,191.6). The small peak at 45,216 Da is the Na<sup>+</sup> adduct. **(c,d)** MS/MS fragmentation of chymotryptic peptides derived from GST-MBP synthesized by *ribo-X* and incorporating two Bpas. The spectra confirm Bpa incorporation at both the expected sites. The fragmentation sites for each fragment ion are illustrated above the spectra. B denotes Bpa.

efficiency makes it possible to synthesize proteins incorporating unnatural amino acids at multiple sites and should minimize the functional and phenotypic effects of truncated proteins *in vivo*.

Because *ribo-X* increases suppression of amber stop codons by suppressor tRNAs of distinct sequence and structure, we suggest

*ribo-X* in the presence of the BpaRS/tRNA<sub>CUA</sub> pair and Bpa confirmed the incorporation of two Bpas; no peaks were detected corresponding to the incorporation of natural amino acids (**Fig. 5b**). The sites of Bpa incorporation were further confirmed by analysis of the tandem mass spectrometry (MS/MS) fragmentation series of the relevant chymotryptic peptides (**Fig. 5c,d**). We observe that the *ribo-X*-mediated improvement in efficiency for one and two amber codons is conserved in minimal medium (**Supplementary Fig. 6**), demonstrating that the effect mediated by *ribo-X* is robust under different expression conditions. Overall, the protein expression data and mass spectrometry data clearly demonstrate that the modular combination of *ribo-X*, BpaRS/tRNA<sub>CUA</sub> and an orthogonal mRNA containing multiple UAG codons allows the site-specific incorporation of Bpa with high fidelity and efficiency at multiple sites in GST-MBP.

## DISCUSSION

We have demonstrated the synthetic evolution of an orthogonal ribosome (*ribo-X*) for the efficient, high-fidelity, suppressor tRNA-dependent decoding of amber stop codons placed within the context of an orthogonal mRNA in living cells. We have combined *ribo-X* with orthogonal mRNAs and orthogonal aminoacyl-tRNA synthetase/tRNA<sub>CUA</sub> pairs to substantially increase the efficiency of site-specific unnatural amino acid incorporation in *E. coli*. This increase in

that *ribo-X* operates by decreasing its functional interaction with RF-1, allowing the suppressor tRNAs to more efficiently compete for A-site binding in the presence of a UAG codon on the mRNA. It will be interesting to see whether extensions of the strategy reported here will yield further increases in amber suppression while maintaining translational fidelity. In this regard it is encouraging both that biochemical evidence suggests distinct conformations of the decoding center are recognized by tRNAs and RF-1 (ref. 45) and that we have been able to select combinations of mutations for which RF-1-mediated termination is decreased without decreasing the fidelity of tRNA decoding. These observations indicate that the molecular determinants for the fidelity of tRNA decoding and RF-1 binding need not be tightly coupled and further independent modulation may be possible within the orthogonal system.

By improving amber suppression efficiency on the orthogonal ribosome we have diverged the decoding properties of the orthogonal ribosome from those of the cellular ribosome such that the same insertion signal is read with a different efficiency on cellular and orthogonal mRNAs within the same cell. A conceptually similar, but mechanistically distinct, strategy involving localization of specialized translational components is used by nature to direct the incorporation of selenocysteine in response to a subset of UGA codons<sup>46</sup>. It will be interesting to see whether similar strategies can be used to

enhance the efficiency of synthetic eukaryotic genetic code expansion. Because the meaning of a codon on any mRNA is set by the translational machinery that decodes that mRNA, it may be possible to use extensions of our approach to write entirely new genetic codes on orthogonal mRNAs and to undo the “frozen accident” of the existing genetic code<sup>47</sup>. For example, it may be possible to use tRNAs that are poor substrates for the cellular ribosome but are efficiently decoded by evolved orthogonal ribosomes to write independent and parallel codes (orthogonal genetic codes). Orthogonal genetic codes may form a basis for a biological ‘virtual operating system’ that further expands the information storage and genetic encoding capacity of the cell.

## METHODS

**Construction of ribosome libraries and reporters.** 16S rDNA libraries were constructed by enzymatic inverse PCR<sup>25</sup> on pRSF vectors containing a previously described O-rDNA (pRSF-O-rDNA)<sup>48</sup>. To create the UAGA reporter plasmid, we introduced the amber-derived UAGA codon at two sites in the chloramphenicol acetyl transferase (*cat*) gene (Ser103 and Ser146, an essential and conserved catalytic serine residue<sup>49</sup> that ensures the fidelity of incorporation, **Supplementary Fig. 2**), downstream of an orthogonal ribosome-binding site, producing *O-cat* (UAGA103, UAGA146). This construct was created by multiple rounds of Quik Change mutagenesis (Stratagene) on an *O-cat* reporter derived from p21 (ref. 25) by replacement of the *cat-upp* fusion with the *cat* gene alone. tRNA genes were introduced into the *O-cat* (UAGA103, UAGA146) plasmid at a unique *Bst* Z171 restriction site, via a cassette containing a 5' synthetic *lpp* promoter and a 3' *rrnC* transcriptional terminator, to create the vector *O-cat* (UAGA103, UAGA146)/tRNA(UCUA); the sequence of the extended anticodon is written 5' to 3'. UAG codon reporters and CUA anticodon tRNAs were derived by Quik Change mutagenesis from the UAGA or UCUA constructs. All final plasmids were confirmed by DNA sequencing. For a complete description of oligonucleotides used for vector construction see **Supplementary Table 2** online.

**Selection of evolved O-ribosomes.** To select O-ribosomes with improved UAGA decoding, each pRSF-O-rDNA library was transformed by electroporation into GeneHog *E. coli* (Invitrogen) containing *O-cat* (UAGA103, UAGA146)/tRNA(UCUA). Transformed cells were recovered for 1 h in SOB medium containing 2% glucose and used to inoculate 200 ml of LB-GKT (LB medium with 2% glucose, 25 µg ml<sup>-1</sup> kanamycin and 12.5 µg ml<sup>-1</sup> tetracycline). After overnight growth (37 °C, 250 r.p.m., 16 h), 2 ml of the cells were pelleted by centrifugation (3,000g), and washed three times with an equal volume of LB-KT (LB medium with 12.5 µg ml<sup>-1</sup> kanamycin and 6.25 µg ml<sup>-1</sup> tetracycline). The resuspended pellet was used to inoculate 18 ml of LB-KT, and the resulting culture incubated (37 °C, 250 r.p.m. shaking, 90 min). To induce expression of plasmid encoded O-rRNA, 2 ml of the culture was added to 18 ml LB-IKT (LB medium with 1.1 mM isopropyl- $\beta$ -thiogalactopyranoside (IPTG), 12.5 µg ml<sup>-1</sup> kanamycin and 6.25 µg ml<sup>-1</sup> tetracycline) and incubated for 4 h (37 °C, 250 r.p.m.). Aliquots (250 µl optical density at 600 nm (OD<sub>600</sub>) = 1.5) were plated on LB-IKT agar (LB agar with 1 mM IPTG, 12.5 µg ml<sup>-1</sup> kanamycin and 6.25 µg ml<sup>-1</sup> tetracycline) supplemented with 50 µg ml<sup>-1</sup> chloramphenicol and incubated (37 °C, 40 h).

**Characterization of evolved O-ribosomes.** To separate selected pRSF-O-rDNA plasmids from the *O-cat* (UAGA103, UAGA146)/tRNA<sup>ser2</sup>(UCUA) reporter plasmids, total plasmid DNA from selected clones was purified and digested with *NotI* restriction endonuclease, and transformed into DH10B *E. coli*. Individual transformants were replica plated onto kanamycin agar and tetracycline agar and plasmid separation of pRSF-O-rDNA from the reporter confirmed by restriction digest and agarose gel analysis.

To quantify the UAGA-decoding activity of selected 16S rDNA clones, the selected pRSF-O-rDNA plasmids were cotransformed with *O-cat* (UAGA103, UAGA146) or *O-cat* (UAGA103, UAGA146)/tRNA<sup>ser2</sup>(UCUA). Cells were recovered (SOB, 2% glucose, 1 h) and used to inoculate 10 ml of LB-GKT, which was incubated (16 h, 37 °C, 250 r.p.m.). We used 1 ml of the resulting

culture to inoculate 9 ml of LB-KT, which was incubated (90 min, 37 °C, 250 r.p.m.). We used 1 ml of the LB-KT culture to inoculate 9 ml of LB-IKT medium, which was incubated (37 °C, 250 r.p.m., 4 h). Individual clones were transferred to a 96-well block and arrayed, using a 96-well pin tool, onto LB-IKT agar plates containing chloramphenicol at concentrations from 0 to 250 µg ml<sup>-1</sup>. The plates were incubated (37 °C, 16 h). We performed analogous experiments for other tRNA codon pairs.

To extract soluble cell lysates for *in vitro* CAT assays, 1 ml of each induced LB-IKT culture was pelleted by centrifugation at 3,000g. The cell pellets were washed three times with 500 µl Washing Buffer (40 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 7.5) and once with 500 µl lysis buffer (250 mM Tris-HCl, pH 7.8). Cells were lysed in 200 µl Lysis Buffer by five cycles of flash-freezing in dry ice/ethanol, followed by rapid thawing in a 50 °C water bath. Cell debris was removed from the lysate by centrifugation (12,000g, 5 min) and the top 150 µl of supernatant frozen at -20 °C. To assay CAT activity in the lysates, 10 µl of soluble cell extract was mixed with 2.5 µl of FAST CAT Green (deoxy) substrate (Invitrogen) and preincubated (37 °C, 5 min). We added 2.5 µl of 9 mM acetyl-CoA (Sigma), and incubated (37 °C, 1 h). The reaction was stopped by the addition of ice-cold ethyl acetate (200 µl, vortex 20 s). The aqueous and organic phases were separated by centrifugation (12,000g, 10 min) and the top 100 µl of the ethyl acetate layer collected. We spotted 1 µl of the collected solution onto a silica gel Thin-layer chromatography plate (Merck) for thin-layer chromatography in chloroform:methanol (85:15 vol/vol). The fluorescence of the spatially resolved substrate and product was visualized and quantified using a phosphorimager (Storm 860, Amersham Biosciences) with excitation and emission wavelengths of 450 nm and 520 nm, respectively.

**Construction of GST-MBP protein expression vectors.** *gst* was amplified from pGEX-2T (GE Healthcare) with the primers: 5'-GAACTCGAGACAATTTTCA TATCCCTCCGCAAATGTCCCTATACTAGGTTATTGGAAAATTAAG-3' and 5'-GAAGAGGTACCCGTCACGATGAATTCGCCGGGATCCACGCGGAAC-3', and digested with *XhoI* and *KpnI*. *male* was amplified from pMAL (NEB) with PCR primers 5'-GAAGGTACCTCAAATCGAAGAAGGTAACCTGGTAA TC-3' and 5'-CCAAAGCTTAGCTTGCCTGCAGGTCGACTC-3' and digested with *HindIII* and *KpnI*. pO-*gst-male* was generated from pGFPmut3.1 (Promega), by replacing A<sup>191</sup>T<sup>192</sup>A<sup>193</sup> in the vector (vector map available at [http://www.clontech.com/images/pt/dis\\_vectors/PT3209-5.pdf](http://www.clontech.com/images/pt/dis_vectors/PT3209-5.pdf)) with CTCGAG (*XhoI* site). This mutates the lac operator and renders expression of the downstream gene constitutively active. *Gfp* was excised from between the *HindIII* site and the newly introduced *XhoI* site, and the *gst-male* fusion introduced with the same sites via a three-fragment ligation. The vector p *gst-male* was created by changing the orthogonal ribosome binding site to a single wild-type ribosome binding site with the enzymatic inverse PCR primers: 5'-GTAGGCTCTCGGATCCCCGGGTACCTAGAATTAAGAGGAGAAATTAAG CATGTCCCTATACTAGGTTATTG-3' and 5'-GTAGGCTCTCGGATCTCTAG AGTGCACCTGCAGGAATGCAAGCTTGGCGTAACTCGAGCCGCTCACAAAT CCACAC-3'. To create vectors containing a single amber codon between *gst* and *male* (*pgst(UAG)male* and pO-*gst(UAG)male*) the Tyr codon, TAC, in the linker between *gst* and *male* was changed to TAG by Quik Change mutagenesis (Stratagene), using the primers 5'-GAATTCATCGTGACGGGTAGCTCAAA ATCGAAGAAGGTAACCTGGTAATCTG-3' and 5'-CTTCGATTTTGAGCTAC CCGTACGATGAATTCGCCGGGATCCACGCGGA AC-3'. For double UAG mutants we additionally mutated the fourth codon in *male* from GAA to TAG by Quik Change, with the primers 5'-CTCAAAATCTAGGAAGGTAACCTG GTAATCTGGATTAACGGCGATAAAG-3' and 5'-CAGTTTACCTTCCTAGAT TTTGAGCTACCCGTCACGATG-3' to create the vectors *pgst(UAG)<sub>2</sub>male* and pO-*gst(UAG)<sub>2</sub>male*.

**Construction of PIP2 ribosomes for protein expression.** The kanamycin-resistance gene and the SC101\* origin of plasmid pZS\*24-MCS1 (ref. 50) were amplified using the following primers: KanSC101fw, 5'-ACTGGA TCC TGC TAG AGG CAT CAA ATA AAA C-3', and KanSC101rv, 5'-AGT ACC GGT TAG ACG TCG GAA TTG CCA GC-3'. The resulting PCR product was digested with *Bam*HI and *Age*I. The *rrnB* operon, including the PIP2 promoter and *rrnC* terminator, was excised from plasmid pTrc PIP2 *rrnB* by digestion with *Ngo*M IV and *Bam*HI, and the amplified SC101\* fragment

and P1P2rrmB fragment were ligated to create plasmid pSC101\*-BD. The *XhoI*/*XbaI* fragment of this plasmid was replaced with a corresponding fragment from pTrcRSF-O-rRNA or pTrcRSF-ribo-X, yielding pSC101\*-O-ribosome and pSC101\*-ribo-X, respectively.

**Expression and purification of GST-MBP fusions.** *E. coli* containing the appropriate plasmid combinations were pelleted (3,000g, 10 min) from 50 ml overnight cultures, resuspended in 1 ml lysis buffer (PBS supplemented with 1× protease inhibitor cocktail (Roche), 1 mM PMSF and 1 mg ml<sup>-1</sup> lysozyme (Sigma)), and incubated (15 min, 37 °C, 1,000 r.p.m.). Cells were chilled on ice before lysis by sonication (30 s, 30 W). The lysate was clarified by centrifugation (6 min, 25,000g, 2 °C). GST containing proteins from the lysate (875 µg, 400 µl) were bound in batch (1 h, 4 °C) to 50 µl of glutathione sepharose beads (GE Healthcare). Beads were washed 3 times with 1 ml PBS, before elution by heating for 10 min at 80 °C in 60 µl 1× SDS gel-loading buffer. All samples were analyzed on 4–20% Tris-glycine gels (Invitrogen).

The densities of the bands for GST-MBP and GST were quantified from Coomassie-stained gels with NIH image 1.63. We divided the background-corrected values by the molecular mass of the corresponding proteins (GST-MBP, 71 kDa and GST, 27 kDa) and used these values to calculate the percentage of amber codon suppression, by dividing the amount of GST-MBP by total amount of protein.

**<sup>35</sup>S-cysteine misincorporation.** GeneHog *E. coli* containing either pO-*gst-malE* and pSC101\*-O-ribosome, pO-*gst-malE* and pSC101\*-ribo-X or p*gst-malE* were resuspended in LB media (supplemented with <sup>35</sup>S-cysteine (1,000 Ci mmol<sup>-1</sup>) to a final concentration of 3 nM, 750 µM methionine, 25 µg ml<sup>-1</sup> ampicillin and 12.5 µg ml<sup>-1</sup> kanamycin) to an OD<sub>600</sub> of 0.1, and cells were incubated (3.5 h, 37 °C, 250 r.p.m.). 10 ml of the resulting culture was pelleted (5,000g, 5 min), washed twice (1 ml PBS per wash), resuspended in 1 ml lysis buffer containing 1% Triton-X, incubated (30 min, 37 °C, 1,000 r.p.m.) and lysed on ice by pipetting up and down. The clarified cell extract was bound to 100 µl of glutathione sepharose beads (1 h, 4 °C) and the beads were pelleted (5,000g, 10 s) and washed twice in 1 ml PBS. The beads were added to 10 ml polypropylene column (Biorad) and washed (30 ml of PBS; 10 ml 0.5 M NaCl, 0.5× PBS; 30 ml PBS) before elution in 1 ml of PBS supplemented with 10 mM glutathione. Purified GST-MBP was digested with 12.5 units of thrombin, to yield a GST fragment and an MBP fragment. The reaction was precipitated with 15% trichloroacetic acid and loaded onto an SDS-PAGE gel to resolve the GST, MBP and thrombin, and stained with GelCode blue (Invitrogen). The <sup>35</sup>S activity in the GST and MBP protein bands were quantified by densitometry, using a Storm Phosphorimager (Molecular Dynamics) and ImageQuant (GE Healthcare). The error frequency per codon for each ribosome examined was determined as follows: GST contains four cysteine codons, so the number of counts per second (c.p.s.) resulting from GST divided by four gives A, the cps per quantitative incorporation of cysteine. MBP contains no cysteine codons, but misincorporation at noncysteine codons gives B c.p.s. Because GST and MBP are present in equimolar amounts, (A/B) × 410, where 410 is the number of amino acids in the MBP containing thrombin cleavage fragment, gives the number of amino acids translated for one cysteine misincorporation C. Assuming the misincorporation frequency for all 20 amino acids is the same as that for cysteine the number of codons translated per misincorporation is C/20, and the error frequency per codon is given by (C/20)<sup>-1</sup>.

**Dual luciferase assays.** pO-DLR contains a genetic fusion between a 5' *Renilla* luciferase (R-luc) and a 3' firefly luciferase (F-luc) on an orthogonal ribosome binding site. To create pO-DLR the R-luc open reading frame from the plasmid pGL4.70[hRluc] (Promega) was amplified by PCR using the primers 5'-GAACTCGAGGGCGCGGCTTTCATATCCCTCCGCAAATGGCCCTCAAGGTGTA CGACCCCGAGCAACGCAAACGCATG-3' and 5'-GCTAGATCTCCTAGGGG CCCCCTCGAGATTGCTCGTTCTTCAGCACGCGCTCCACGAAGCTC-3'. The PCR product was digested with *XhoI* and *BglII*. The F-luc ORF was amplified with primer pair 5'-AGGAGATCTAGCGCTGATCCCCCGGGGA GCTCATCGAAGATGCCAAAACATTAAGAAGGGCCAG-3' and 5'-GACA AGCTTACACGGCGATCTTGCCGCCCTTCTTG-3' and digested with *BglII* and *HindIII*. The *gst-malE* gene fusion was excised from pO-*gst-malE* by *XhoI* and *HindIII* digestion and pO-DLR created by a triple ligation of the released

vector backbone with the digested F-luc PCR product and the digested R-luc PCR product. Pwt-DLR was created by a similar strategy, but using the primer pair 5'-GAACCTCGAGTACCTAGATATAAAGAGGAGAAATTAAGCATGGCCCT CCAAGGTGTACGACCCCGAGCAACGCAAACGCATG-3' and 5'-GCTAGAT CTCTAGGGGCCCCCTCGAGATTGCTCGTTCTTCAGCACGCGCTCCA CGAAGCTC-3' to amplify the R-luc ORF. Codon 529 variants were created by Quik Change Mutagenesis (Stratagene).

pO-DLR, and its K529 codon variants, were transformed into GeneHog *E. coli* cells with pSC101\*-O-ribosome or pSC101\*-ribo-X. pwt-DLR, and its K529 codon variants, were transformed into GeneHog cells with pSC101\*-BD. Individual colonies were incubated (37 °C, 250 r.p.m., 20 h) in 2 ml LB supplemented with ampicillin (100 µg ml<sup>-1</sup>) and kanamycin (50 µg ml<sup>-1</sup>), pelleted (5,000g, 5 min) and resuspended in 200 µl (1 mg ml<sup>-1</sup> lysozyme, 10 mM Tris (pH 8.0), 1 mM EDTA). Cells were incubated on ice for 20 min, frozen on dry ice, and thawed on ice. 10 µl samples of this extract were assayed for firefly (F-luc) and *Renilla* (R-luc) luciferase activity using the Dual-Luciferase Reporter Assay System (Promega). Each ribosome reporter combination was assayed from four independent cultures using an Orion microplate luminometer (Berthold Detection Systems) and the data analyzed as previously described. The error reported is the s.d.

**Mass spectrometry.** 25 µM GST-2BPA-MBP (GST-MBP with Bpas incorporated in response to two amber codons) in 22 µl 40 mM (NH<sub>4</sub>)HCO<sub>3</sub> was alkylated and digested with chymotrypsin overnight. To obtain a fragment series for the N-terminal Bpa incorporation, 5 µl of a tenfold dilution of the chymotryptic peptide mixture was desalted and concentrated by using a GELoader tip filled with Poros R3 sorbent (PerSeptive Biosystems). The bound peptides were eluted with 1 µl of 40% acetonitrile/4% formic acid directly into a nanospray capillary and then introduced into an API QSTAR pulsar i hybrid quadrupole-time-of-flight mass spectrometer (MDS Sciex). Product ion scans were carried out in positive ion-mode and MS survey scans for peptides measured. Selected ions (*m/z* = 668.4<sup>4+</sup>) were fragmented by collision-induced dissociation (CID) with nitrogen in the collision cell and spectra of fragment ions produced were recorded in the time-of-flight mass analyzer. To obtain a fragment series for the C-terminal Bpa incorporation, peptides from the chymotryptic digest were separated by nanoscale liquid chromatography (LC Packings) on a reversed-phase C18 column (150 × 0.075 mm internal diameter, flow rate 0.25 ml min<sup>-1</sup>). The eluate was introduced directly into a Q-STAR hybrid tandem mass spectrometer (LC-MS/MS) and peptide with *m/z* = 469.7<sup>4+</sup> fragmented.

Protein total mass was determined on an LCT time-of-flight mass spectrometer with electrospray ionization (ESI). (Micromass). Proteins were rebuffered to 10 mM ammonium bicarbonate pH 7.5 and diluted 1:100 into 50% methanol, 1% formic acid. Samples were infused into the ESI source at 10 ml min<sup>-1</sup>, using a Harvard Model 22 infusion pump (Harvard Apparatus) and calibration performed in positive ion mode using horse heart myoglobin. 60–80 scans were acquired and added to yield the mass spectra. Molecular masses were obtained by deconvoluting multiply charged protein mass spectra using MassLynx version 4.1 (Micromass). Theoretical molecular masses of wild-type proteins were calculated using ProtParam (<http://us.expasy.org/tools/protparam.html>), and theoretical masses for unnatural amino acid containing proteins adjusted manually.

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

#### ACKNOWLEDGMENTS

J.W.C. is an EMBO Young Investigator. K.W. is grateful for a Medical Research Council-Laboratory of Molecular Biology (MRC-LMB) Cambridge Scholarship, an Honorary External Research Studentship from Trinity College, Cambridge, and an Overseas Research Studentship Award. H.N. is an MRC Career Development Fellow. We are grateful to O. Barrett and W. An for sharing unpublished materials and assisting in early stages of this project. We are grateful to M. Babu for extracting *E. coli* amber codon usage, P.G. Schultz (TSRI) for the pSUP Bpa vector. This work was funded by The Medical Research Council.

#### COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

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

1. Xie, J. & Schultz, P.G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **7**, 775–782 (2006).
2. Furter, R. Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in *Escherichia coli*. *Protein Sci.* **7**, 419–426 (1998).
3. Wang, L., Brock, A., Herberich, B. & Schultz, P.G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498–500 (2001).
4. Sakamoto, K. *et al.* Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* **30**, 4692–4699 (2002).
5. Chin, J.W. *et al.* An expanded eukaryotic genetic code. *Science* **301**, 964–967 (2003).
6. Kohrer, C., Xie, L., Kellerer, S., Varshney, U. & RajBhandary, U.L. Import of amber and ochre suppressor tRNAs into mammalian cells: a general approach to site-specific insertion of amino acid analogues into proteins. *Proc. Natl. Acad. Sci. USA* **98**, 14310–14315 (2001).
7. Kohrer, C., Yoo, J.H., Bennett, M., Schaack, J. & RajBhandary, U.L. A possible approach to site-specific insertion of two different unnatural amino acids into proteins in mammalian cells via nonsense suppression. *Chem. Biol.* **10**, 1095–1102 (2003).
8. Lummis, S.C. *et al.* Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248–252 (2005).
9. Nowak, M.W. *et al.* *In vivo* incorporation of unnatural amino acids into ion channels in *Xenopus* oocyte expression system. *Methods Enzymol.* **293**, 504–529 (1998).
10. Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. & Schultz, P.G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182–188 (1989).
11. Bain, J.D., Glabe, C.G., Dix, T.A., Chamberlin, A.R. & Diala, E.S. Biosynthetic site-specific incorporation of non-natural amino acids into a polypeptide. *J. Am. Chem. Soc.* **111**, 8013–8014 (1989).
12. Heckler, T.G. *et al.* T4 RNA ligase mediated preparation of novel “chemically misacylated” tRNA<sup>Phe</sup>S. *Biochemistry* **23**, 1468–1473 (1984).
13. Hohsaka, T., Ashizuka, Y., Murakami, H. & Sisido, M. Incorporation of Non-natural amino acids into streptavidin through *in vitro* frame shift suppression. *J. Am. Chem. Soc.* **118**, 9778–9779 (1996).
14. Murakami, H., Ohta, A., Goto, Y., Sako, Y. & Suga, H. Flexizyme as a versatile tRNA acylation catalyst and the application for translation. *Nucleic Acids Symp. Ser. (Oxf.) no.* **50**, 35–36 (2006).
15. Capecchi, M.R. Polypeptide chain termination *in vitro*: isolation of a release factor. *Proc. Natl. Acad. Sci. USA* **58**, 1144–1151 (1967).
16. Caskey, C.T., Tompkins, R., Scolnick, E., Caryk, T. & Nirenberg, M. Sequential translation of trinucleotide codons for the initiation and termination of protein synthesis. *Science* **162**, 135–138 (1968).
17. Scolnick, E., Tompkins, R., Caskey, T. & Nirenberg, M. Release factors differing in specificity for terminator codons. *Proc. Natl. Acad. Sci. USA* **61**, 768–774 (1968).
18. Dougherty, D.A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **4**, 645–652 (2000).
19. Mori, H. & Ito, K. Different modes of SecY-SecA interactions revealed by site-directed *in vivo* photo-cross-linking. *Proc. Natl. Acad. Sci. USA* (2006).
20. Chin, J.W. & Schultz, P.G. *In vivo* photocrosslinking with unnatural amino acid mutagenesis. *ChemBioChem* **3**, 1135–1137 (2002).
21. Short, G.F., III, Golovine, S.Y. & Hecht, S.M. Effects of release factor 1 on *in vitro* protein translation and the elaboration of proteins containing unnatural amino acids. *Biochemistry* **38**, 8808–8819 (1999).
22. Kleina, L.G., Masson, J.M., Normanly, J., Abelson, J. & Miller, J.H. Construction of *Escherichia coli* amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. *J. Mol. Biol.* **213**, 705–717 (1990).
23. Ryu, Y. & Schultz, P.G. Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods* **3**, 263–265 (2006).
24. Blattner, F.R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
25. Rackham, O. & Chin, J.W. A network of orthogonal ribosome • mRNA pairs. *Nat. Chem. Biol.* **1**, 159–166 (2005).
26. Hui, A. & de Boer, H.A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **84**, 4762–4766 (1987).
27. Triman, K.L., Peister, A. & Goel, R.A. Expanded versions of the 16S and 23S ribosomal RNA mutation databases (16SMDBexp and 23SMDBexp). *Nucleic Acids Res.* **26**, 280–284 (1998).
28. Schuwirth, B.S. *et al.* Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**, 827–834 (2005).
29. Korostelev, A., Trakhanov, S., Laurberg, M. & Noller, H.F. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell* **126**, 1065–1077 (2006).
30. Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).
31. Carter, A.P. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348 (2000).
32. Petry, S. *et al.* Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell* **123**, 1255–1266 (2005).
33. Magliery, T.J., Anderson, J.C. & Schultz, P.G. 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 (2001).
34. Cannone, J.J. *et al.* The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2 (2002).
35. Datta, K. & Majumdar, M.K. A method for assay of chloramphenicol acetyltransferase from crude cell extract. *Microbiologica* **8**, 73–77 (1985).
36. Rice, J.B., Libby, R.T. & Reeve, J.N. Mistranslation of the mRNA encoding bacteriophage T7 0.3 protein. *J. Biol. Chem.* **259**, 6505–6510 (1984).
37. Kramer, E.B. & Farabaugh, P.J. The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. *RNA* **13**, 87–96 (2007).
38. Kauer, J.C., Erickson-Viitanen, S., Wolfe, H.R., Jr. & DeGrado, W.F. p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. *J. Biol. Chem.* **261**, 10695–10700 (1986).
39. Chin, J.W., Martin, A.B., King, D.S., Wang, L. & Schultz, P.G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **99**, 11020–11024 (2002).
40. 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).
41. Schlieker, C. *et al.* Substrate recognition by the AAA+ chaperone ClpB. *Nat. Struct. Mol. Biol.* **11**, 607–615 (2004).
42. Weibezahn, J. *et al.* Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* **119**, 653–665 (2004).
43. Steer, B.A. & Schimmel, P. Major anticodon-binding region missing from an archaeobacterial tRNA synthetase. *J. Biol. Chem.* **274**, 35601–35606 (1999).
44. Bossi, L. Context effects: translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message. *J. Mol. Biol.* **164**, 73–87 (1983).
45. Youngman, E.M., Cochella, L., Brunelle, J.L., He, S. & Green, R. Two distinct conformations of the conserved RNA-rich decoding center of the small ribosomal subunit are recognized by tRNAs and release factors. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 545–549 (2006).
46. Stadtman, T.C. Selenocysteine. *Annu. Rev. Biochem.* **65**, 83–100 (1996).
47. Crick, F.H. Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* **19**, 548–555 (1966).
48. Rackham, O. & Chin, J.W. Cellular logic with orthogonal ribosomes. *J. Am. Chem. Soc.* **127**, 17584–17585 (2005).
49. Murray, I.A. & Shaw, W.V. O-Acetyltransferases for chloramphenicol and other natural products. *Antimicrob. Agents Chemother.* **41**, 1–6 (1997).
50. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/11–12 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).