LETTERS

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome

Heinz Neumann[1]*, Kaihang Wang[1]*, Lloyd Davis[1], Maria Garcia-Alai[1] & Jason W. Chin[1]

The *in vivo*, genetically programmed incorporation of designer amino acids allows the properties of proteins to be tailored with molecular precision[1]. The *Methanococcus jannaschii* tyrosyl-transfer-RNA synthetase–$tRNA_{CUA}$ ($MjTyrRS–tRNA_{CUA}$)[2,3] and the *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase–$tRNA_{CUA}$ ($MbPylRS–tRNA_{CUA}$)[4–6] orthogonal pairs have been evolved to incorporate a range of unnatural amino acids in response to the amber codon in *Escherichia coli*[1,6,7]. However, the potential of synthetic genetic code expansion is generally limited to the low efficiency incorporation of a single type of unnatural amino acid at a time, because every triplet codon in the universal genetic code is used in encoding the synthesis of the proteome. To encode efficiently many distinct unnatural amino acids into proteins we require blank codons and mutually orthogonal aminoacyl-tRNA synthetase–tRNA pairs that recognize unnatural amino acids and decode the new codons. Here we synthetically evolve an orthogonal ribosome[8,9] (ribo-Q1) that efficiently decodes a series of quadruplet codons and the amber codon, providing several blank codons on an orthogonal messenger RNA, which it specifically translates[8]. By creating mutually orthogonal aminoacyl-tRNA synthetase–tRNA pairs and combining them with ribo-Q1 we direct the incorporation of distinct unnatural amino acids in response to two of the new blank codons on the orthogonal mRNA. Using this code, we genetically direct the formation of a specific, redox-insensitive, nanoscale protein cross-link by the bio-orthogonal cycloaddition of encoded azide- and alkyne-containing amino acids[10]. Because the synthetase–tRNA pairs used have been evolved to incorporate numerous unnatural amino acids[1,6,7], it will be possible to encode more than 200 unnatural amino acid combinations using this approach. As ribo-Q1 independently decodes a series of quadruplet codons, this work provides foundational technologies for the encoded synthesis and synthetic evolution of unnatural polymers in cells.

Because each of the 64 triplet codons are used to encode natural amino acids or polypeptide termination, new blank codons are required for cellular genetic code expansion. In principle, quadruplet codons might provide 256 blank codons. Stoichiometrically aminoacylated extended anticodon tRNAs have been used to incorporate unnatural amino acids in response to four-base codons with very low efficiency in *in vitro* systems[11–13] and in limited *in vivo* systems, by the import of previously aminoacylated tRNA[14,15]. In one case a four-base suppressor and amber codon have been used, in a non-generalizable approach, to encode two unremarkable amino acids with low efficiency[16]. Indeed, the inefficiency with which natural ribosomes decode quadruplet codons severely limits their utility for genetic code expansion.

A ribosome must accommodate an extended anticodon tRNA in its decoding centre to be able to decode it[17,18]. Natural ribosomes are very inefficient at, and unevolvable for quadruplet decoding, which would enhance misreading of the proteome. In contrast, orthogonal ribosomes[8], which are specifically addressed to the orthogonal message and are not responsible for synthesizing the proteome, may, in principle, be evolved to efficiently decode quadruplet codons on the orthogonal message (Supplementary Figs 1 and 2). To discover evolved orthogonal ribosomes that enhance quadruplet decoding we first created 11 saturation mutagenesis libraries in the 16S ribosomal RNA of ribo-X (an orthogonal ribosome previously evolved for efficient amber codon decoding on an orthogonal message[9]); taken together these libraries cover 127 nucleotides that are within 12 Å of a tRNA bound in the decoding centre[19] (Supplementary Fig. 3). We used ribo-X as a starting point for library generation because we hoped to discover evolved orthogonal ribosomes that gain the ability to efficiently decode quadruplet codons while maintaining the ability to efficiently decode amber codons on the orthogonal mRNA; thereby maximizing the number of additional codons that can be decoded by the orthogonal ribosome.

To select orthogonal ribosomes that efficiently decode quadruplet codons using extended anticodon tRNAs we combined each orthogonal (O)-ribosome library with a reporter construct (O-*cat* (AAGA 146)–$tRNA^{Ser\,2}_{UCUU}$). The reporter contains a chloramphenicol acetyl transferase gene that is specifically translated by O-ribosomes[9], an in-frame AAGA quadruplet codon and $tRNA^{Ser\,2}_{UCUU}$ (a designed variant of $tRNA^{Ser\,2}$ that is aminoacylated by *E. coli* seryl-tRNA synthetase and decodes the AAGA codon[9,20]). The orthogonal *cat* gene is read in frame, and confers chloramphenicol resistance, only if $tRNA^{Ser\,2}_{UCUU}$ efficiently decodes the AAGA codon and restores the reading frame. Clones surviving on chloramphenicol concentrations that kill cells containing ribo-X and the *cat* reporter have four distinct sequences. Clone ribo-Q4 has double mutations: C1195A and A1196G, ribo-Q3 has the three mutations: C1195T, A1196G and A1197G; ribo-Q2 and ribo-Q1 have two mutations: A1196G and A1197G, ribo-Q2 also has eight other non-programmed mutations. Although the entire decoding centre was mutated, the selected mutations are spatially localized and might accommodate an extended anticodon–codon interaction in the decoding centre (Fig. 1a). The chloramphenicol resistance of cells containing $tRNA^{Ser\,2}_{UCCU}$ and *cat* with two AGGA codons is greatly enhanced when the *cat* gene is translated by ribo-Q ribosomes in place of unevolved ribosomes (Fig. 1b, c). Indeed the chloramphenicol resistance of cells containing two AGGA codons read by the ribo-Q ribosomes approaches that of a wild-type *cat* gene. This suggests that ribo-Q1 may decode quadruplet codons with an efficiency approaching that for triplet decoding and with a much greater efficiency than the unevolved ribosome. The enhancement in quadruplet decoding efficiency is maintained for a variety of quadruplet codon–anticodon interactions (Supplementary Fig. 4).

Natural ribosomes decode triplet codons with high fidelity (error frequencies ranging from 10^{-2} to 10^{-4} errors per codon have been

[1]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK.
*These authors contributed equally to this work.

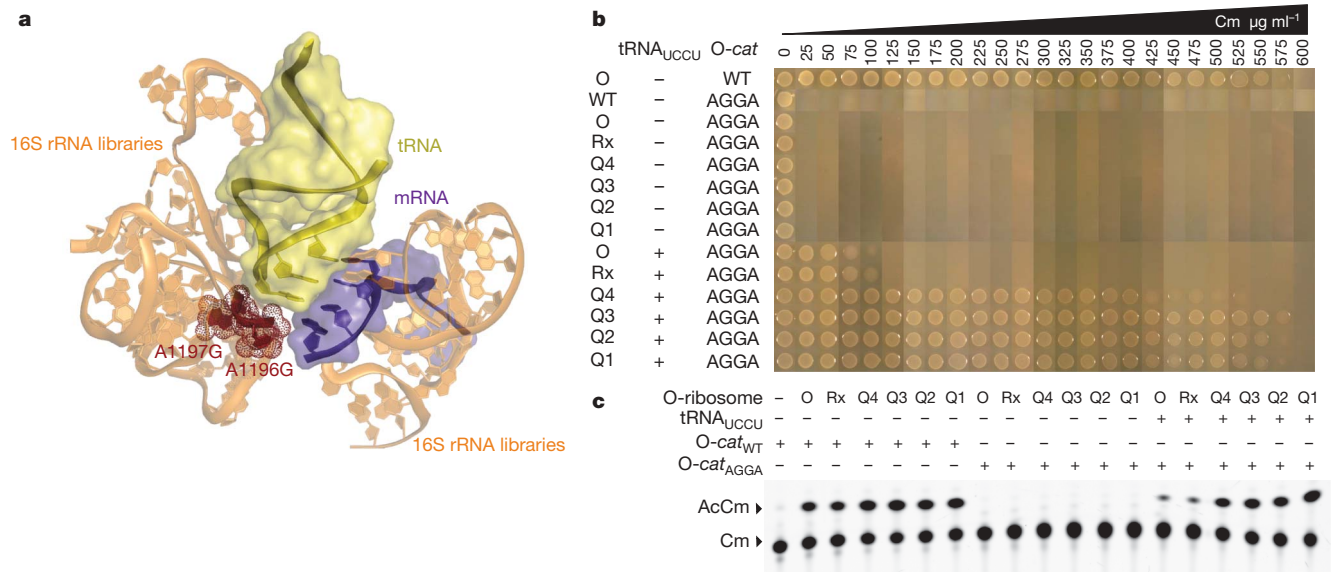

Figure 1 | Selection and characterization of orthogonal quadruplet decoding ribosomes. **a**, Mutations in quadruplet decoding ribosomes form a structural cluster close to the space potentially occupied by an extended anticodon tRNA. Selected nucleotides are shown in red. Image created using Pymol (http://www.pymol.org) and Protein Data Bank (PDB) accession 2J00. **b**, Ribo-Qs substantially enhance the decoding of quadruplet codons. The tRNA$^{Ser\,2}_{UCCU}$-dependent enhancement in decoding AGGA codons in the O-*cat* (AGGA 103, AGGA 146) gene was measured by survival on increasing concentrations of chloramphenicol (Cm). WT, wild type. **c**, As in **b**, but measuring CAT enzymatic activity directly by thin-layer chromatography. AcCm, acetylated chloramphenicol; O, O-ribosome; Q1–Q4, ribo-Q1–Q4; Rx, ribo-X.

reported[21–23]). To explicitly compare the fidelity of triplet decoding and quadruplet decoding for the evolved orthogonal ribosomes and the progenitor ribosome we used two independent methods: the incorporation of ^{35}S-cysteine into a protein, which contains no cysteine codons in its gene[9], and variants of a dual luciferase system[9,23] (Supplementary Fig. 5). We find that the triplet and quadruplet decoding translational fidelity is the same for the evolved ribosome (ribo-Q1) and unevolved and wild-type ribosomes, and the fourth base of the codon–anticodon interaction is discriminated equally well by all ribosomes (Supplementary Fig. 5).

To demonstrate that the enhanced amber-decoding properties of ribo-X are maintained in ribo-Q1 we compared the efficiency of incorporating *p*-benzoyl-L-phenylalanine (Bpa) into a recombinant glutathione *S*-transferase–maltose binding protein (GST–MBP) fusion in response to an amber codon on an orthogonal mRNA using orthogonal ribosomes and a previously evolved Bpa-tRNA synthetase–tRNA$_{CUA}$ pair[3] (BpaRS–tRNA$_{CUA}$) (Fig. 2). Ribo-Q1 and ribo-X incorporate Bpa with a comparable and high efficiency in response to the amber codons in the orthogonal mRNA (compare lanes 4 and 6 and lanes 10 and 12 in Fig. 2a). Ribo-X and ribo-Q1 are substantially more efficient than the wild-type ribosome at incorporating Bpa by means of amber suppression (compare lanes 4 and 6 to lane 2, and lanes 10 and 12 to lane 8 in Fig. 2a).

To demonstrate the utility of ribo-Q1 for incorporating unnatural amino acids in response to quadruplet codons we compared the efficiency of incorporating *p*-azido-L-phenylalanine (AzPhe) into a recombinant GST–MBP fusion in response to a quadruplet codon using ribo-Q1 or the wild-type ribosome. To direct the incorporation of AzPhe we used the AzPheRS*–tRNA$_{UCCU}$ pair (a variant of the *p*AzPheRS-7–tRNA$_{CUA}$ pair[24] derived from the MjTyrRS–tRNA$_{CUA}$ pair for the incorporation of AzPhe as described later). We find that ribo-Q1 substantially increases the efficiency of incorporation of AzPhe in response to a quadruplet codon, and even allows the incorporation of AzPhe in response to two quadruplet codons for the first time (compare lanes 2 and 6 and lanes 4 and 8, Fig. 2b). The site and fidelity of incorporation of AzPhe were further confirmed by analysis of tandem mass spectrometry (MS/MS) fragmentation series of the relevant tryptic peptides (Supplementary Fig. 7).

To take advantage of ribo-Q1 for the incorporation of multiple distinct unnatural amino acids in recombinant proteins, we required mutually orthogonal aminoacyl-tRNA synthetase–tRNA pairs. We demonstrated that the MbPylRS–tRNA$_{CUA}$ pair[4,5] and MjTyrRS–tRNA$_{CUA}$ pair[2], each of which have previously been evolved to incorporate a range of unnatural amino acids[1,6,7,25], are mutually orthogonal in their aminoacylation specificity (Supplementary Fig. 8). We created the AzPheRS*–tRNA$_{UCCU}$ pair, which is derived from the MjTyrRS–tRNA$_{CUA}$ pair, by a series of generally applicable directed evolution steps (Supplementary Figs 9–11). The MbPylRS–tRNA$_{CUA}$ pair and AzPheRS*–tRNA$_{UCCU}$ pair are mutually orthogonal: they decode distinct codons, use distinct amino acids and are orthogonal in their aminoacylation specificity (Supplementary Fig. 12).

To demonstrate the simultaneous incorporation of two useful unnatural amino acids into a single protein we combined the MbPylRS–MbtRNA$_{CUA}$ pair, the AzPheRS* tRNA$_{UCCU}$ pair and

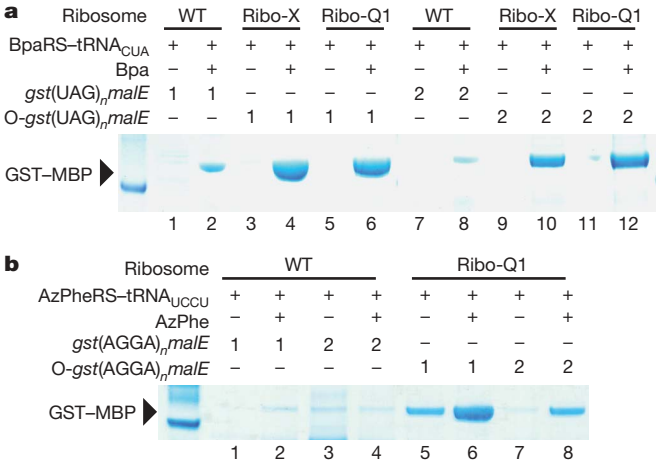

Figure 2 | Enhanced incorporation of unnatural amino acids in response to amber and quadruplet codons with ribo-Q1. **a**, Ribo-Q1 incorporates Bpa as efficiently as ribo-X. The entire gel is shown in Supplementary Fig. 6. **b**, Ribo-Q1 enhances the efficiency of AzPhe incorporation in response to the AGGA quadruplet codon using AzPheRS*–tRNA$_{UCCU}$. The gel showing the ratio of GST–MBP to GST as well as MS/MS spectra of the single and double AzPhe incorporations are shown in Supplementary Fig. 7. (UAG)$_n$ or (AGGA)$_n$ describes the number of amber or AGGA codons (n) between *gst* and *malE*.

ribo-Q1 in *E. coli*. We used these components to produce full-length GST–calmodulin containing AzPhe and *N6*-[(2-propynyloxy)carbonyl]-L-lysine (CAK; which we recently discovered is an efficient substrate for MbPylRS[7]) (Fig. 3) in response to an AGGA and UAG codon in an orthogonal gene. Production of the full-length protein required the addition of both unnatural amino acids. We further confirmed the incorporation of AzPhe and CAK at the genetically programmed sites by MS/MS sequencing of a single tryptic fragment containing both unnatural amino acids (Fig. 3).

To begin to demonstrate that emergent properties may be programmed into proteins by combinations of unnatural amino acids we genetically directed the formation of a triazole cross-link, by a copper catalysed Huisgen [2+3] cycloaddition reaction ('click reaction')[10]. We first encoded AzPhe and CAK at positions 1 and 149 in calmodulin (Fig. 4). After incubation of calmodulin incorporating the azide (AzPhe) and alkyne (CAK) at these positions with Cu(I) for 5 min we observe a more rapidly migrating protein band in SDS–polyacrylamide gel electrophoresis (SDS–PAGE). MS/MS sequencing unambiguously confirms that the faster mobility band results from the product of a bio-orthogonal cycloaddition reaction between AzPhe and CAK. Our results demonstrate the genetically programmed proximity acceleration of a new class of asymmetric, redox-insensitive cross-link that can be used to specifically constrain protein structure

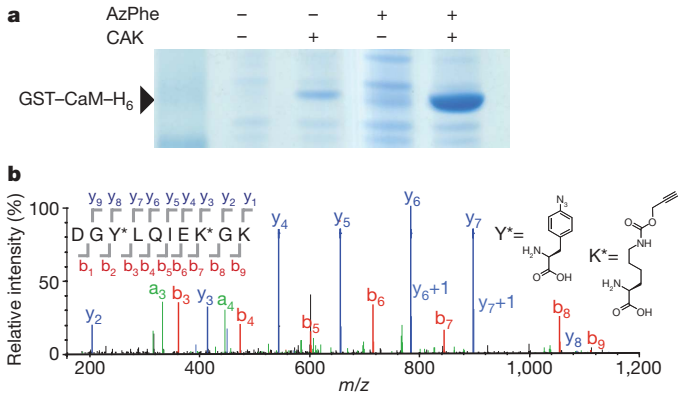

Figure 3 | Encoding an azide and an alkyne in a single protein by orthogonal translation. **a**, Expression of GST–CaM–His_6 (a GST–calmodulin–His6 fusion) containing two unnatural amino acids. An orthogonal gene producing a GST–CaM–His_6 fusion that contains an AGGA codon at position 1 and an amber codon at position 40 of calmodulin was translated by ribo-Q1 in the presence of AzPheRS*–$tRNA_{UCCU}$ and MbPylRS–$tRNA_{CUA}$. The entire gel is shown in Supplementary Fig. 13. **b**, LC–MS/MS analysis of the incorporation of two distinct unnatural amino acids into the linker region of GST–MBP. Y*, AzPhe; K*, CAK.

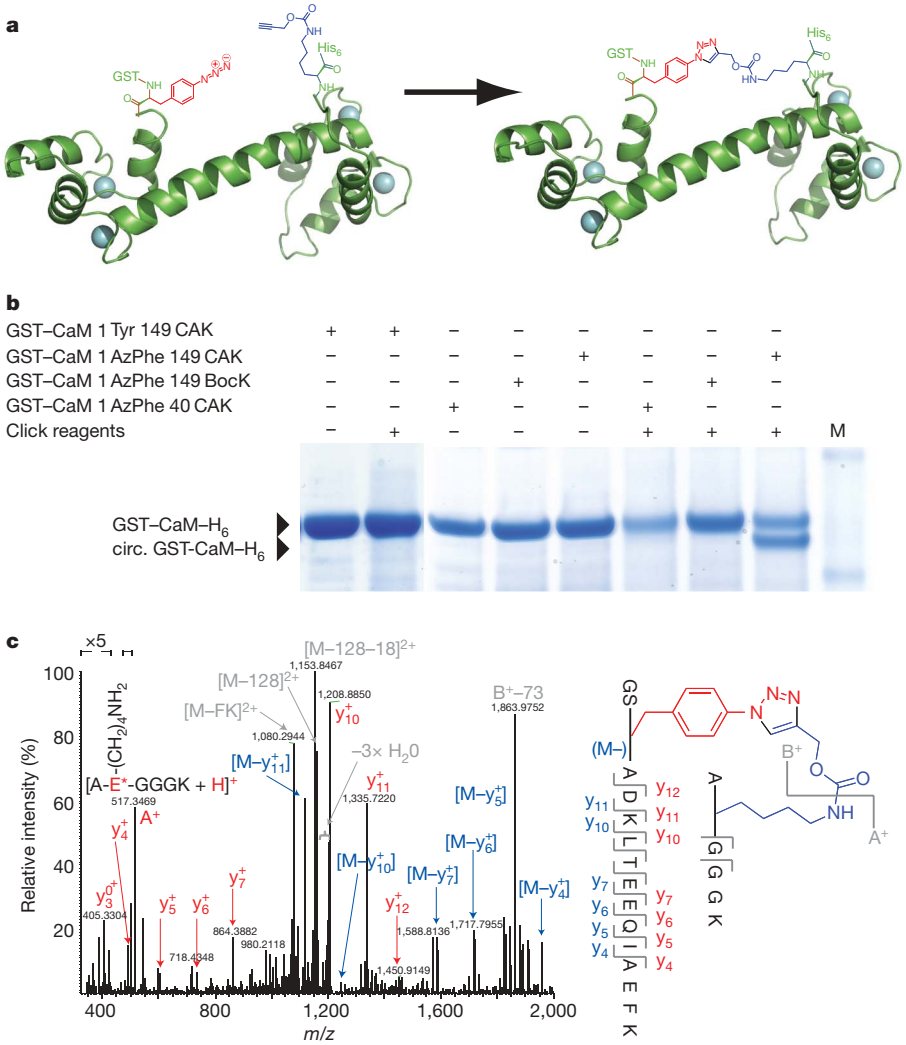

Figure 4 | Genetically directed cyclization of calmodulin by a Cu(I)-catalysed Huisgen's [2+3]-cycloaddition. **a**, Structure of calmodulin indicating the sites of incorporation of AzPhe and CAK and their triazole product. Image created using Pymol (http://www.pymol.org) and Protein Data Bank (PDB) accession 4CLN. **b**, GST–CaM–His_6 1 AzPhe 149 CAK specifically cyclizes with Cu(I)-catalyst. BocK, N^{ε}-*tert*-butyl-oxycarbonyl-L-lysine; circ., circularized protein. The first two lanes are from a separate gel. M denotes molecular mass marker. **c**, LC–MS/MS confirms the triazole formation. The MS/MS spectra of a doubly charged peptide containing the crosslink (m/z = 1,226.6092, which is within 1.8 p.p.m. of the mass expected for cross-linked peptide).

on the nanometre scale. Unlike existing protein cyclization methods for recombinant proteins[26,27], these cross-links can be encoded at any spatially compatible sites in a protein, not just placed at the termini. In contrast to the chemically diverse cyclization methods that can be accessed with peptides by solid-phase peptide synthesis[28], these cross-links can be encoded into proteins of essentially any size. Given the importance of disulphide bonds in natural therapeutic proteins and hormones, the utility of peptide stapling strategies[29], the importance of peptide cyclization[30], and the improved stability of proteins cyclized by native chemical ligation[26], it will be interesting to investigate the enhancement of protein function that may be accessed by combining the encoding of these cross-links with directed evolution methods. By combining the numerous variant MjTyrRS–$tRNA_{CUA}$ and MbPylRS–$tRNA_{CUA}$ pairs reported for the incorporation of unnatural amino acids[1,6,7] (after appropriate anticodon conversion using the steps reported here) with ribo-Q1, it will be possible to encode more than 200 amino acid combinations in recombinant proteins.

METHODS SUMMARY

Methods for cloning, site-directed mutagenesis and library construction are described in the Supplementary Materials. Ribosome libraries were screened for quadruplet suppressors using a modification of the strategy to discover ribo-X[9].

E. coli genehogs or DH10B were used in all protein expression experiments using LB medium supplemented with appropriate antibiotics and unnatural amino acids. Proteins were purified by affinity chromatography using published standard protocols.

Translational fidelity of evolved O-ribosomes was measured by misincorporation of ^{35}S-labelled cysteine[9]. In brief, GST–MBP was produced by the O-ribosome in the presence of ^{35}S-cysteine. The protein was purified, cleaved with thrombin, which cleaves the linker between GST and MBP, and analysed by SDS–PAGE and phospho-imaging. A modified dual-luciferase assay was used to measure the fidelity of translation of O-ribosomes[9]. Luminescence from a luciferase mutant containing an inactivating missense mutation in this assay is a measure of translational inaccuracy of the ribosome. The dual luciferase reporter system was translated by the O-ribosome, extracted in the cold and luciferase activity measured using the Dual-Luciferase Reporter Assay System (Promega).

Liquid chromatography (LC)–MS/MS of proteins was performed by NextGen Science. Proteins were excised from Coomassie-stained SDS–PAGE gels, digested with trypsin and analysed by LC–MS/MS. Total protein mass was obtained by electrospray ionization (ESI)-MS; purified protein was dialysed into 10 mM ammonium bicarbonate, pH 7.5, mixed 1:1 with 1% formic acid in 50% methanol, and total mass was determined in positive ion mode.

Cyclization reactions were performed for 5 min at room temperature on purified protein in 50 mM sodium phosphate, pH 8.3, in the presence of 1 mM ascorbic acid, 1 mM $CuSO_4$ and 2 mM bathophenathroline. Details of all methods can be found in the Supplementary Materials.

Received 26 June 2009; accepted 7 January 2010.
Published online 14 February 2010.

1. Xie, J. & Schultz, P. G. A chemical toolkit for proteins–an expanded genetic code. *Nature Rev. Mol. Cell Biol.* **7,** 775–782 (2006).
2. Steer, B. A. & Schimmel, P. Major anticodon-binding region missing from an archaebacterial tRNA synthetase. *J. Biol. Chem.* **274,** 35601–35606 (1999).
3. 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).
4. Srinivasan, G., James, C. M. & Krzycki, J. A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296,** 1459–1462 (2002).
5. Polycarpo, C. *et al.* An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc. Natl Acad. Sci. USA* **101,** 12450–12454 (2004).
6. Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Genetically encoding N^{ε}-acetyllysine in recombinant proteins. *Nature Chem. Biol.* **4,** 232–234 (2008).
7. Nguyen, D. P. *et al.* Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc.* **131,** 8720–8721 (2009).
8. Rackham, O. & Chin, J. W. A network of orthogonal ribosome·mRNA pairs. *Nature Chem. Biol.* **1,** 159–166 (2005).
9. Wang, K., Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nature Biotechnol.* **25,** 770–777 (2007).
10. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Edn Engl.* **41,** 2596–2599 (2002).
11. Hohsaka, T. & Sisido, M. Incorporation of non-natural amino acids into proteins. *Curr. Opin. Chem. Biol.* **6,** 809–815 (2002).
12. Ohtsuki, T., Manabe, T. & Sisido, M. Multiple incorporation of non-natural amino acids into a single protein using tRNAs with non-standard structures. *FEBS Lett.* **579,** 6769–6774 (2005).
13. Murakami, H., Hohsaka, T., Ashizuka, Y. & Sisido, M. Site-directed incorporation of p-nitrophenylalanine into streptavidin and site-to-site photoinduced electron transfer from a pyrenyl group to a nitrophenyl group on the protein framework. *J. Am. Chem. Soc.* **120,** 7520–7529 (1998).
14. Rodriguez, E. A., Lester, H. A. & Dougherty, D. A. *In vivo* incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. *Proc. Natl Acad. Sci. USA* **103,** 8650–8655 (2006).
15. 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).
16. Anderson, J. C. *et al.* An expanded genetic code with a functional quadruplet codon. *Proc. Natl Acad. Sci. USA* **101,** 7566–7571 (2004).
17. Atkins, J. F. & Bjork, G. R. A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment. *Microbiol. Mol. Biol. Rev.* **73,** 178–210 (2009).
18. Stahl, G., McCarty, G. P. & Farabaugh, P. J. Ribosome structure: revisiting the connection between translational accuracy and unconventional decoding. *Trends Biochem. Sci.* **27,** 178–183 (2002).
19. Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313,** 1935–1942 (2006).
20. 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).
21. Khazaie, K., Buchanan, J. H. & Rosenberger, R. F. The accuracy of Qβ RNA translation. 1. Errors during the synthesis of Qβ proteins by intact *Escherichia coli* cells. *Eur. J. Biochem.* **144,** 485–489 (1984).
22. Laughrea, M., Latulippe, J., Filion, A. M. & Boulet, L. Mistranslation in twelve *Escherichia coli* ribosomal proteins. Cysteine misincorporation at neutral amino acid residues other than tryptophan. *Eur. J. Biochem.* **169,** 59–64 (1987).
23. 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).
24. Chin, J. W. *et al.* Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **124,** 9026–9027 (2002).
25. Mukai, T. *et al.* Adding L-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun.* **371,** 818–822 (2008).
26. Camarero, J. A., Pavel, J. & Muir, T. W. Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization. *Angew. Chem. Int. Ed.* **37,** 347–349 (1998).
27. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C. & Benkovic, S. J. Production of cyclic peptides and proteins *in vivo*. *Proc. Natl Acad. Sci. USA* **96,** 13638–13643 (1999).
28. Li, P. & Roller, P. P. Cyclization strategies in peptide derived drug design. *Curr. Top. Med. Chem.* **2,** 325–341 (2002).
29. Walensky, L. D. *et al.* Activation of apoptosis *in vivo* by a hydrocarbon-stapled BH3 helix. *Science* **305,** 1466–1470 (2004).
30. Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A. & Walsh, C. T. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **407,** 215–218 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to P. B. Kapadnis for synthesizing CAK and to W. An for assistance with some experiments. J.W.C. is grateful to the ERC and the MRC for funding. K.W. is grateful to Trinity College, Cambridge for a fellowship.

Author Contributions K.W., H.N., L.D. and J.W.C. planned the experiments. K.W. selected and characterized ribo-Q, with help from L.D. K.W. and L.D. characterized amber and quadruplet incorporation by protein expression and mass spectrometry. L.D. and M.G.-A. performed protein expression experiments. H.N. demonstrated the mutual orthogonality of synthetase tRNA pairs, evolved synthetases, and characterized the double incorporation and protein cyclization, with help from M.G.-A. H.N., K.W., L.D. and J.W.C. analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.W.C. (chin@mrc-lmb.cam.ac.uk).