2012

High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites

Kiyoshi Ozawa
University of Wollongong, ozawa@uow.edu.au

Karin V. Loscha
Australian National University

Kekini V. Kuppan
Australian National University

Choy Theng Loh
Australian National University

Nicholas E. Dixon
University of Wollongong, nickd@uow.edu.au

See next page for additional authors

Publication Details

Ozawa, K., Loscha, K. V., Kuppan, K. V., Loh, C. Theng., Dixon, N. E., Otting, G. (2012). High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. Biochemical and Biophysical Research Communications, 418 (4), 652-656.

High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites

Abstract
Using aminoacyl-tRNA synthetase/suppressor tRNA pairs derived from Methanocaldococcus jannaschii, an Escherichia coli cell-free protein production system affords proteins with site-specifically incorporated unnatural amino acids (UAAs) in high yields through the use of optimized amber suppressor $tRNA_{CUA}$ opt and optimization of reagent concentrations. The efficiency of the cell-free system allows the incorporation of trifluoromethyl-phenylalanine using a polyspecific synthetase evolved previously for p-cyanophenylalanine, and the incorporation of UAAs at two different sites of the same protein without any re-engineering of the E. coli cells used to make the cell-free extract.

Keywords
high, two, yield, cell, free, protein, synthesis, site, specific, incorporation, unnatural, amino, sites, acids, CMMB

Publication Details
Ozawa, K., Loscha, K. V., Kuppan, K. V., Loh, C. Theng., Dixon, N. E., Otting, G. (2012). High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. Biochemical and Biophysical Research Communications, 418 (4), 652-656.

Authors
Kiyoshi Ozawa, Karin V. Loscha, Kekini V. Kuppan, Choy Theng Loh, Nicholas E. Dixon, and Gottfried Otting

CITATION:

Ozawa, K., Loscha, K. V., Kuppan, K. V., Loh, C. T., Dixon, N. E., and Otting G. (2012) High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. *Biochem. Biopys. Res. Comm.* 418, 652-656.

DOI: 10.1016/j.bbrc.2012.01.069

High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites

Kiyoshi Ozawa[a,b,*], Karin V. Loscha[a], Kekini V. Kuppan[a], Choy Theng Loh[a], Nicholas E. Dixon[b], Gottfried Otting[a,*]

[a]Research School of Chemistry, Australian National University, Canberra ACT 0200 and [b]School of Chemistry, University of Wollongong, NSW 2522, Australia

*To whom correspondence should be addressed. Tel: +61 2 61256507; Fax: +61 2 61250750; Email: ozawa@rsc.anu.edu.au, go@rsc.anu.edu.au

****

Using aminoacyl-tRNA synthetase/suppressor tRNA pairs derived from *Methanocaldococcus jannaschii*, an *Escherichia coli* cell-free protein production system affords proteins with site-specifically incorporated unnatural amino acids (UAA) in high yields through the use of optimized amber suppressor $tRNA_{CUA}^{opt}$ and optimization of reagent concentrations. The efficiency of the cell-free system allows the incorporation of trifluoromethyl-phenylalanine using a polyspecific synthetase evolved previously for *p*-cyano-phenylalanine, and the incorporation of UAAs at two different sites of the same protein without any re-engineering of the *E. coli* cells used to make the cell-free extract.

****

Keywords: aminoacyl-tRNA-synthetase, cell-free protein synthesis, *Escherichia coli*, suppressor tRNA, unnatural amino acids

Abbreviations: AcF, *p*-acetyl-phenylalanine; AzF, *p*-azido-phenylalanine; Bpa, *L*-4-benzoyl-phenylalanine; Bpy, *L*-bipyridylalanine; CECF, continuous exchange cell-free; CF, cell-free; Cm^r, chloramphenicol resistance; CNF, *p*-cyano-phenylalanine; Hco, 7-hydroxy-coumarin-4-yl-ethylglycine; RS, aminoacyl-tRNA-synthetase; tfmF, *L*-4-trifluoromethyl-phenylalanine; $tRNA_{CUA}$, amber suppressor tRNA derived from *M. jannaschii* and optimized for UAA incorporation in *E. coli* in 2001; $tRNA_{CUA}^{opt}$, recently further optimized $tRNA_{CUA}$; UAA, unnatural amino acid.

1. Introduction

Unnatural amino acids (UAA) positioned in proteins at strategically chosen sites open an exciting new range of possibilities for the study and engineering of proteins and their interactions with other molecules. The most successful approach for site-specific incorporation of UAAs is by genetic encoding, where an amber suppressor tRNA is enzymatically charged with the UAA. With the advent of efficient orthogonal suppressor-tRNA ($tRNA_{CUA}$)/aminoacyl-tRNA synthetase (RS) systems, this is possible *in vivo* [1,2]. It has recently been shown that $tRNA_{CUA}$/RS systems evolved for *in vivo* incorporation of UAAs can also be used in cell-free (CF) protein synthesis [3-7]. In general, modern CF systems produce more protein in a given volume of medium than *in vivo* systems [8,9], which would make them attractive for the production of proteins from expensive UAAs. In addition, optimal reagent concentrations vary among different UAAs, mutants and proteins, highlighting the importance of controlling reaction conditions, which is difficult to achieve in *in vivo* systems.

Here we report an efficient implementation of $tRNA_{CUA}$/RS systems for the incorporation of different UAAs in a continuous exchange CF (CECF) system. CECF systems enhance protein yields by the exchange of low-molecular weight compounds across a dialysis membrane [10]. Our system is based on an *E. coli* CECF system supplied with orthogonal suppressor-tRNA ($tRNA_{CUA}$) and aminoacyl-tRNA synthetases (RS) evolved from *M. jannaschii* [2]. By optimization of the reagent concentrations, the system readily produces proteins with UAAs in yields comparable to those of the wild-type. A recently optimized amber suppressor $tRNA_{CUA}^{opt}$ [11] is found to perform as well or better than the original $tRNA_{CUA}$ [1]. We demonstrate the production of proteins containing UAAs at two different sites, without any need for orthogonal ribosomes [12] or extensively reengineered *E. coli* strains [13-15]. Furthermore, we show that a polyspecific RS enzyme [16] that has not previously been used for the incorporation of trifluoromethyl-phenylalanine (tfmF) delivers excellent protein yields at increased RS concentrations. Our system allows protein expression from conventional pET vectors [17] or linear DNA containing the T7 promoter and terminator. By expressing the genes from linear PCR-amplified DNA [18], the amber mutations and corresponding UAAs are readily introduced at different sites without any cloning step, creating a high-throughput system.

2. Materials and Methods

2.1. Materials

Synthetic oligonucleotides were purchased from GeneWorks (Hindmarsh, SA, Australia) or Integrated DNA Technology, Inc. (Coralville, IA, USA). The sequences of oligonucleotides used are listed in Table S1 (Appendix A). The plasmids pEVOL-*p*AcF [11], pEVOL-*p*AzF [11], pEVOL-*p*CNF [11,16], pEVOL-CouA [11,19], pSup-MjTyrRS-6TRN [20] and pSup-BpaRS-6TRN [20] were obtained from Prof. Peter G. Schultz (Scripps Research Institute, CA, USA). Vent DNA polymerase, restriction enzymes and other enzymes for cloning were from New England Biolabs (Ipswich, MA, USA). The NucleoSpin Extract II kit was obtained from Macherey-Nagel (Düren, Germany). *E. coli* strains BL21Star (DE3) and BL21 (DE3)/pLysS were from Invitrogen (Carlsbad, CA, USA). ProPur IMAC Mini spin columns were purchased

from Nalge Nunc International (USA). *L*-4-Benzoyl-phenylalanine (Bpa) and *L*-4-trifluoro-methylphenylalanine (tfmF) were purchased from Peptech Corporation (Burlington, MA, USA), and *p*-acetylphenylalanine (AcF) and *p*-azidophenylalanine (AzF) were obtained from Mr Ansis Maleckis (University of Latvia). All other standard chemicals necessary for cell-free protein synthesis were as described previously [21].

2.2. $tRNA_{CUA}$ and $tRNA_{CUA}^{opt}$ expression vectors

Total tRNA containing $tRNA_{CUA}$ derived from *M. jannaschii* [1] or a newly optimized version ($tRNA_{CUA}^{opt}$) [11] was purified from *E. coli* cells containing $tRNA_{CUA}$ or $tRNA_{CUA}^{opt}$ expression vectors. To generate an expression vector for $tRNA_{CUA}$, the wild-type *Mj*TyrRS gene was excised from the plasmid pSup-MjTyrRS-6TRN (4379 bp) [20] using *Nde*I and *Pst*I. The resulting vector was treated with T4 DNA polymerase to generate blunt ends and the blunt-ended vector was re-ligated. The resulting plasmid was termed pSUP-6TRN (about 3.2 kb, Cm^r). It contains two sets of three copies of *M. jannaschii* amber suppressor tyrosyl-$tRNA_{CUA}$ ($tRNA_{CUA}$). To generate an expression vector for $tRNA_{CUA}^{opt}$, the vector pEVOL-pAcF (about 6.1 kb, Cm^r) was digested with *Bgl*II and *Pst*I to remove the AcF-RS gene, treated with T4 DNA polymerase to generate blunt ends and re-ligated. The resulting plasmid was termed pKO1474 (about 4 kb, Cm^r). It contains a single copy of the recently optimized *M. jannaschii* amber suppressor tyrosyl-$tRNA_{CUA}$ ($tRNA_{CUA}^{opt}$ [11]).

2.3. Expression and purification of total tRNA containing suppressor $tRNA_{CUA}$

Preparations of total tRNA containing suppressor $tRNA_{CUA}$ were obtained from *E. coli* BL21 Star (DE3) cells harboring the plasmid pSUP-6TRN (for $tRNA_{CUA}$ [20]) or pKO1474/pEVOL-AcF (for the optimized $tRNA_{CUA}^{opt}$ [11]). The cells were grown at 37° C in 20 L of Z medium [21] (pH 7.3) supplemented with 22.4 mM glucose, 10 mg/L thiamine and 33 mg/L chloramphenicol in a fermenter [21,22] until A_{595} = 3.0. While 15–16 L of this culture were used for subsequent preparation of S30 extracts (see below), the remaining 4–5 L were used for purification of $tRNA_{CUA}$ (or $tRNA_{CUA}^{opt}$). 20 g of cells were resuspended in 100 mL of 0.3 M potassium acetate solution (pH 4.8) and the cells were disrupted by being passed twice through a French press at 12,000 psi. Next, an equal volume of water-saturated phenol was added. After vigorous shaking for 1 h at room temperature, the greenish mixture was centrifuged at 18,000 rpm (using a Sorvall SS-34 rotor) for 30 min. The supernatant was subjected to ethanol precipitation. The resulting pellet was air-dried and resuspended in 100 mL of Milli-Q water. It was loaded onto a 2.5 x 8 cm column of Q-Sepharose (Amersham Pharmacia) equilibrated with equilibration buffer (20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.2 M NaCl). The column was washed with 120 mL of equilibration buffer and the total tRNA was eluted with running buffer (20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 M NaCl) at a flow rate of 1.0 mL/min. The eluted tRNA was subjected to ethanol precipitation and resuspended in sterile Milli-Q water to make a 17.5 mg/mL stock solution. Typical yields were about 85 mg of total tRNA. The tRNA concentration was determined by measuring A_{260} with a NanoDrop spectrophotometer (Thermo Scientific, USA), assuming $A_{260} \approx 40$ μg tRNA/mL.

2.4. RS enzyme preparations

The *M. jannaschii*-derived RS enzymes BpyRS, BpaRS, AzF-RS, AcF-RS, HcoRS and CNF-RS for the incorporation of different UAAs [2] were individually produced with His_6-tags and purified. The wild-type *Mj*TyrRS was overexpressed and purified without a His_6-tag. Detailed protocols are provided in Appendix A.

2.5. Preparation of E. coli S30 extracts

S30 extracts were prepared from the cell cultures grown for suppressor-tRNA preparation (see above) using *E. coli* BL21 Star (DE3) containing pSup-6TRN (or pKO1474 for the optimized $tRNA_{CUA}^{opt}$), using our published standard procedure [21]. Over-expression of T7 RNA polymerase was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at $A_{600} = 1.0$ and the cells were harvested after a further 2 h incubation at 37 ˚C, $A_{600} = 3.0$ [21,22]. S30 extracts prepared from *E. coli* BL21 Star (DE3) without any additional plasmids (but including induction of T7 RNA polymerase) performed equally well for the incorporation of UAAs in CECF experiments.

2.6. Cell-free reactions

CECF reactions were carried out as described previously [21], using either plasmids or PCR-amplified DNA as DNA templates, in 0.2 or 0.6 mL reaction mixtures in a dialysis bag suspended in a ten times larger volume of outer buffer. The inner buffer was supplied with purified RS and tRNA, and inner and outer buffers contained the 20 natural amino acids and the UAA (each at 1 mM). Except for the additional presence of $tRNA_{CUA}$ and RS and UAA, the reactions were performed at 30 ˚C in 7 hours or overnight as described previously [21].

2.7. WNVpro mutants Trp53TAG and Tyr132TAG

The WNVpro construct and the mutants with amber stop codons replacing W53 and Y132 were described previously [5,6].

2.8. WNVpro double-mutant Glu101TAG/Lys104TAG

The double-amber mutant Glu101TAG/Lys104TAG of the WNVpro-K96A gene was prepared in two rounds of PCR, always using Vent DNA polymerase. First, the site-specific mutation was introduced in two separate PCR reactions (50 μL each) with 20–30 ng of pRSET-5b-WNVpro-K96A template [11]. The first reaction used primer 1375 and the reverse primer 2231 (containing the desired mutations) and the second reaction used the forward primer 2230 and primer 1376 (Table S1). The PCR products were separately purified by 1.5% agarose gel electrophoresis and extraction from the gel using the NucleoSpin Extract II kit (Macherey-Nagel, Düren, Germany) to remove the residual primers and template plasmid. The gel-purified 5' and 3' DNA fragments from the first round of PCR were mixed together in equimolar ratio to produce the template (20–30 ng) for the second round of PCR (50 μL) using the primer pairs 1375 and 1376 (Table S1). The resulting PCR product was digested with *Nde*I and *Eco*RI, followed by agarose gel

purification and isolation using the NucleoSpin Extract II kit. The isolated WNVpro-K96A gene with the double-amber mutation Glu101TAG/Lys104TAG was cloned into the *Nde*I–*Eco*RI sites of the high-copy number pRSET-5b T7-expression vector [23].

2.9. DnaB mutant Phe166TAG

The template DNA of the DnaB Phe166TAG mutant was prepared as linear PCR-amplified DNA without cloning into a vector. Similarly to the double-amber mutant of WNVpro, the template DNA was prepared in two rounds of PCR, using Vent DNA polymerase for all PCR reactions. Starting from plasmid DNA with the *E. coli dnaB* gene in the vector pETMCSIII [24]. The first round used primer 1375 and the reverse primer 2572 (containing the desired mutation) and the forward primer 2571 and primer 1376 in separate reactions. Following gel-purification and mixing in equimolar ratio as described above, the second round of PCR appended the T7 promoter and terminator sequences in two separate PCR reactions (50 µL each) using the primer pairs 1131 and 1134, and 1132 and 1133, respectively (Table S1), with a mixture of 20–30 ng of purified PCR products from the first round. The two PCR products were mixed in an approximately equimolar ratio and the residual primers removed using the NucleoSpin Extract II kit. Following denaturation at 95 ˚C (5 min) and reannealing at room temperature (5 min), this generated DNA with complementary single-stranded 8-nt overhangs suitable for cyclization by the intrinsic ligase activity of the cell extract [18,22]. The reannealed DNA solution was used as the template for subsequent cell-free protein synthesis at a concentration of about 10 µg DNA per mL of reaction mixture.

3. Results

3.1. CECF reactions

Our CECF system is based on the system described by Kigawa et al. [10], which we modified in several aspects [21], most importantly by using an S30 extract from the OmpT protease deficient *E. coli* strain BL21 Star (DE3) and by inducing the cells grown to produce the S30 extract with IPTG at A_{600} = 1 to produce T7 RNA polymerase, obviating the need for additional polymerase during the CF reactions. In addition, $tRNA_{CUA}$ derived from *M. jannaschii* [1] or a newly optimized version ($tRNA_{CUA}^{opt}$) [11] was provided together with the natural tRNAs by supplying a total tRNA mixture purified from *E. coli* cells containing $tRNA_{CUA}$ or $tRNA_{CUA}^{opt}$ expression vectors, and the *M. jannaschii*-derived RS enzymes for the incorporation of different UAAs [2] were provided as purified proteins following expression with a His_6-tag.

3.2. Comparison of $tRNA_{CUA}$ and $tRNA_{CUA}^{opt}$

Fig. 1 shows that the protein yields in CECF synthesis with UAAs depended on the sequence context of the amber stop codon. Total tRNA containing the $tRNA_{CUA}^{opt}$ optimized previously for *in vivo* expression [11] invariably performed as well or better than the original $tRNA_{CUA}$ [1] for a range of different proteins (Fig. 1 and unpublished results).

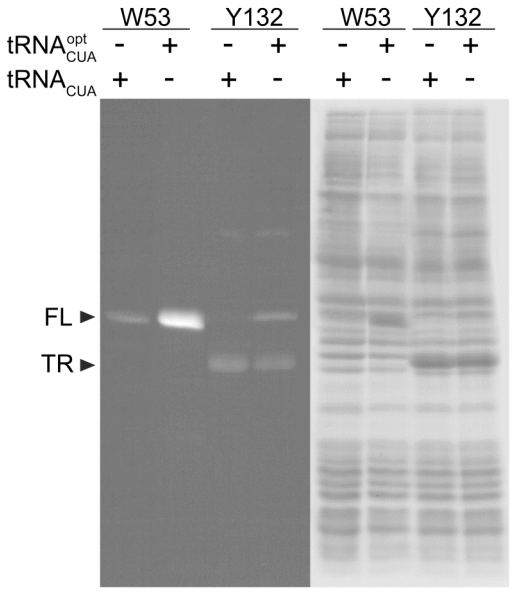

Fig. 1. Incorporation of 7-hydroxycoumarin-4-yl-ethylglycine (Hco) into the West Nile virus protease (WNVpro) amber mutants Trp53TAG and Tyr132TAG, using $tRNA_{CUA}$ [1] or $tRNA_{CUA}^{opt}$ [11] in the CECF reaction. The concentrations of tRNA (263 μg/mL) and HcoRS (180 μg/mL) were the same in all reactions. FL: full-length product. TR: truncated product. The truncated product fluoresces weakly because of the high tryptophan content of WNVpro (K. V. Loscha, unpublished results). (A) SDS-PAGE photographed under UV (312 nm). (B) Same gel as in (A) following staining with Coomassie blue.

3.3. Incorporation of UAAs at two different sites in the same protein

Using $tRNA_{CUA}^{opt}$, we tested the possibility of incorporating two UAAs at two different sites, using the double mutant WNVpro with amber codons at positions 101 and 104. Fig. 2 shows that CECF reactions with $tRNA_{CUA}^{opt}$ delivered good yields with a range of different UAAs. As might be expected for different RS activities, different UAAs were incorporated with different efficiencies under otherwise similar conditions. Most notably, for *p*-acetyl-phenylalanine (AcF), bipyridyl-phenylalanine (Bpy) and Hco, the expression yields approached that of the wild-type protein without UAAs (Fig. 2).

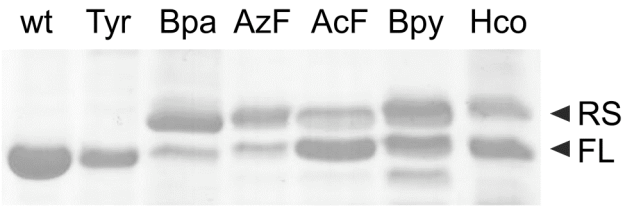

Fig. 2. Incorporation of UAAs into WNVpro at two different sites. The codons for Glu101 and Lys104 were replaced by amber codons. In all experiments, the concentration of total tRNA containing $tRNA_{CUA}^{opt}$ was 175 μg/mL and the UAA was present at 1 mM. Wild-type and mutant *Mj*TyrRS were supplied as purified proteins. Except for the wild-type *Mj*TyrRS, all RS had an N-terminal His_6-tag. The lanes are labeled with the UAA used. wt: control without amber codons. Tyr: double amber mutant made using 1.1 mg/mL *Mj*TyrRS; Bpa: 1.1 mg/mL BpaRS; AzF: 1.1 mg/mL AzF-RS; 1.1 mg/mL AcF-RS; Bpy: 1.4 mg/mL BpyRS; 0.87 mg/mL HcoRS. The proteins were purified over a Ni-NTA column. The bands of the His-tagged RS enzymes were used as internal references for estimating the yield of the WNVpro mutants. The SDS-PAGE (15%) gel was stained with Coomassie blue for visualization.*3.4. Incorporating trifluoromethyl-phenylalanine using a polyspecific RS.*

p-benzoyl-phenylalanine (Bpa) and *p*-azido-phenylalanine (AzF) produced full-length protein in relatively poor yields in the double-mutant of Fig. 2. In the case of Bpa, expression yields similar to wild-type were obtained by increasing the concentration of $tRNA_{CUA}^{opt}$, whereas the incorporation of AzF increased only little (Fig. 3A). This may be attributed to different activities of the respective RS mutants.

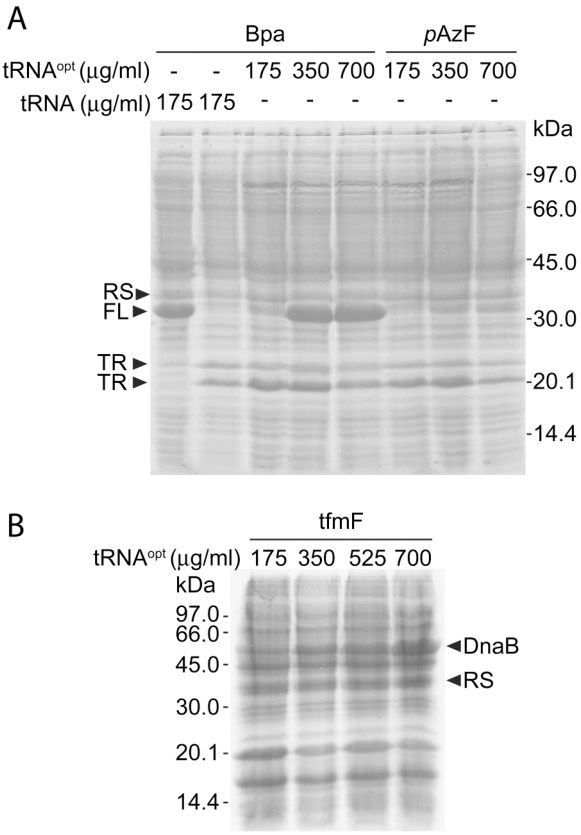

Fig. 3. Optimization of tRNA concentration. The SDS-PAGE (15%, Coomassie blue stained) gels showing CECF-produced WNVpro and DnaB. A) Wild-type and mutant WNVpro produced in the presence of synthetase (either 0.54 mg/mL BpaRS or AzF-RS) and the matching UAA (1 mM Bpa or AzF). Left-most lane: wild-type WNVpro. All other lanes show the double-amber mutant WNVpro Glu101TAG/Lys104TAG (where TAG denotes the UAA). The concentrations of total tRNA containing $tRNA_{CUA}$ or $tRNA_{CUA}^{opt}$ are indicated at the top. FL: full-length protein. TR: truncation products resulting from the two amber stop codons. B) DnaB Phe166tfmF produced in the presence of CNF-RS (1.4 mg/mL), tfmF (1 mM) and varying amounts of total tRNA containing $tRNA_{CUA}^{opt}$ as indicated at the top.

The yields of WNVpro Glu101Bpa/Lys104Bpa were maximal already at 350 µg/mL (Fig. 3A). In contrast, using twice as much $tRNA_{CUA}^{opt}$ more than doubled the protein yield when *L*-4-trifluoromethyl-phenylalanine (tfmF) was incorporated into the Phe166TAG mutant of the *E. coli* DnaB helicase, using the polyspecific RS enzyme evolved for incorporation of *p*-cyano-phenylalanine (CNF) [16] (Fig. 3B). To compensate for non-optimal activity with tfmF, the DnaB Phe166tfmF mutant was made in the presence of 2.6-fold more RS enzyme than the WNVpro double-amber mutant. Fig. 4 shows that over 2 mg/mL of DnaB F166tfmF could be produced at high RS and $tRNA_{CUA}^{opt}$ concentrations, with virtually no premature truncation of the protein.

4. Discussion

Early experiments employed cell-free protein synthesis to demonstrate the possibility of site-specific incorporation of UAAs into proteins following chemical loading of suppressor tRNA with the UAA [25]. Subsequent development of orthogonal tRNA/RS systems that achieve

specific and efficient loading of the suppressor tRNA with the UAA *in vivo* [1,2] shifted attention to *in vivo* systems.

CECF protein synthesis adds distinct advantages to the orthogonal tRNA/RS systems for producing proteins with UAAs. (i) CF synthesis tends to produce significantly higher protein yields per amino acid than *in vivo* protein expression [8,9], which is important for expensive UAAs. (ii) The possibility to adjust the concentrations of suppressor tRNA and RS is critical for high protein yields, as optimal conditions depend on the sequence context of the amber codons and on the UAA. Adjusting the concentrations of tRNA and RS is much harder to achieve *in vivo*. For example, the present study demonstrates that the polyspecific CNF-RS [16] can efficiently incorporate tfmF into proteins, if CNF-RS is supplied at increased concentration. Similarly high concentrations of the RS enzyme would be difficult to achieve *in vivo* without impeding the expression of the target protein. (iii) We showed that two UAAs can readily be incorporated into the same protein without need to impair the activity of RF1, a feat that has previously been achieved only with heavily reengineered systems [13,26.28-30].

A $tRNA_{CUA}$/RS system [2] has been implemented earlier in a CF system operating in batch mode [3,4]. We prefer to use a CECF system because it produces more protein from a smaller reaction volume, reducing the amount of macromolecular reagents (S30 extract, DNA, tRNA and purified RS), which tend to be more costly than the UAAs. In the reported batch mode system, inactivation of the OmpT protease proved deleterious to protein yields [3]. In contrast, we obtain high yields from conventional T7 vectors using the S30 extract from the *E. coli ompT* strains BL21 Star (DE3) and Rosetta (DE3).

In principle, access to the reaction mixture in CF systems opens the possibility to inactivate the release factor RF1, ensuring that the amber stop codon is only recognized by $tRNA_{CUA}$. In our hands, however, preliminary experiments using anti-RF1 antibodies [31] or RF1 aptamers [32] gave no improvements over the results shown in Fig. 2. Notably, however, the performance of the CECF system described here already works reliably with different proteins and UAAs. As expected and demonstrated here, the system allows protein synthesis from linear PCR-amplified DNA [18] also with UAAs, making it easy to generate many different amber mutants as required for systematic mutation studies in a high-throughput approach [33].

Acknowledgements

We thank Prof. Peter G. Schultz for the vectors pEVOL-*p*AcF, pEVOL-*p*AzF, pEVOL-*p*CNF, pEVOL-CouA, pSUP-BpaRS-6TRN and pSUP-MjTyrRS-6TRN, Ms Thi Hoang Duong Nguyen for a sample of Bpy and an expression vector of WNVpro mutants, Dr Zhi-Qiang Xu for the plasmid pETMCSIII-*dnaB*, and Mr Ansis Maleckis for samples of AcF and AzF. This work was supported by the Australian Research Council, including a fellowship to K. O.

Disclosure statement

The authors have nothing to disclose.

Declarations of interest

The authors declare no conflicts of interest.

References

[1] L. Wang, P.G. Schultz, A general approach for the generation of orthogonal tRNAs, Chem. Biol. 8 (2001) 883-890.

[2] C.C. Liu, P.G. Schultz, Adding new chemistries to the genetic code, Annu. Rev. Biochem. 79 (2010) 413-444.

[3] A.R. Goerke, J.R. Swartz, High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids, Biotech. Bioeng. 102 (2009) 400-416.

[4] B.C. Bundy, J.R. Swartz, Site-specific incorporation of *p*-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation, Bioconj. Chem. 21 (2010) 255-263.

[5] T.H.D. Nguyen, K. Ozawa, M. Stanton-Cook et al., Pseudocontact shifts in protein NMR spectra generated by a genetically encoded Co^{2+}-binding amino acid, Angew. Chem. Int. Ed. 50 (2011) 692-694.

[6] I.N. Ugwumba, K. Ozawa, L. de la Cruz et al., Using a genetically encoded fluorescent amino acid as a site-specific probe to detect binding of low-molecular weight compounds, Assay Drug Dev. Technol. 9 (2011) 50-57.

[7] I.N. Ugwumba, K. Ozawa, Z.-Q. Xu et al., Improving a natural enzyme activity through incorporation of unnatural amino acids, J. Am. Chem. Soc. 133 (2011) 326-333.

[8] T. Torizawa, M. Shimizu, M. Taoka et al., Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol, J. Biomol. NMR 30 (2004) 311-325.

[9] J. Yokoyama, T. Matsuda, S. Koshiba et al., An economical method for producing stable-isotope labeled proteins by the *E. coli* cell-free system, J. Biomol. NMR 48 (2010) 193-201.

[10] T. Kigawa, T. Yabuki, Y. Yoshida et al., Cell-free production and stable-isotope labeling of milligram quantities of proteins, FEBS Lett. 442 (1999) 15-19.

[11] T.S. Young, I. Ahmad, J.A. Yin et al., An enhanced system for unnatural amino acid mutagenesis in *E. coli*, J. Mol. Biol. 395 (2010) 361-374.

[12] K. Wang, H. Neumann, S.Y. Peak-Chew et al., Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion, Nat. Biotechnol. 25 (2007) 770-777.

[13] T. Mukai, A. Hayashi, F. Iraha et al., Codon reassignment in the *Escherichia coli* genetic code, Nucleic Acids Res. 38 (2010) 8188-8195.

[14] F.J. Isaacs, P.A. Carr, H.H. Wang et al., Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement, Science 333 (2011) 348-353.

[15] D.B.F. Johnson, J. Xu, Z. Shen et al., RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites, Nat. Chem. Biol. 7 (2011) 779-786.

[16] D.D. Young, T.S. Young, M. Jahnz et al. An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity, Biochemistry 50 (2011) 1894-1900.

[17] F.W. Studier, B.A. Moffat, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, J. Mol. Biol. 189 (1986) 113-130.

[18] P.S.C. Wu, K. Ozawa, S.P. Lim et al., Cell-free transcription/translation from PCR amplified DNA for high-throughput NMR studies, Angew. Chem. Int. Ed. 46 (2007) 3356–3358.

[19] J. Wang, J. Xie, P.G. Schultz, A genetically encoded fluorescent amino acid, J. Am. Chem. Soc. 128 (2006) 8738-8739.

[20] Y. Ryu, P.G. Schultz, Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*, Nat. Methods 3 (2006) 263-265.

[21] M. Apponyi, K. Ozawa, N.E. Dixon et al., Cell-free protein synthesis for analysis by NMR spectroscopy, in: B. Kobe, M. Guss, T. Huber (Eds.), Methods in Molecular Biology, Vol. 426, Humana Press, Totowa, 2008, pp. 257–268.
[22] K. Ozawa, S. Jergic, A.Y. Park et al., The proofreading exonuclease subunit ε of *Escherichia coli* DNA polymerase III is tethered to the polymerase subunit α via a flexible linker, Nucleic Acids Res. 36 (2008) 5074-5082.
[23] R. Schoepfer, The pRSET family of T7 promoter expression vectors for *Escherichia coli*, Gene 124 (1993) 83-85.
[24] C. Neylon, S.E. Brown, A.V. Kralicek et al., Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: A model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. Biochemistry 39 (2000) 11989–11999.
[25] C.J. Noren, S.J. Anthony-Cahill, M.C. Griffith et al., A general method for site-specific incorporation of unnatural amino acids into proteins, Science 244 (1989) 182-188.
[26] M. Ryden, J. Murphy, R. Martin et al., Mapping and complementation studies of the gene for release factor 1, J. Bacteriol. 168 (1986) 1066-1069.
[27] Y. Shimizu, A. Inoue, Y. Tomari et al., Cell-free translation reconstituted with purified components, Nat. Biotechnol. 19 (2001) 751-755.
[28] K. Wang, H. Neumann, S.Y. Peak-Chew et al., Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion, Nat. Biotechnol. 25 (2007) 770-777.
[29] Y. Doi, T. Ohtsuki, Y. Shimizu et al., Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system, J. Am. Chem. Soc. 129 (2007) 14458-14462.
[30] E. Seki, N. Matsuda, T. Kigawa, Multiple inhibitory factor removal from an *Escherichia coli* cell extract improves cell-free protein synthesis, J. Biosci. Bioeng. 108 (2009) 30-35.
[31] M. Gerrits, J. Strey, I. Claußnitzer et al., Cell-free synthesis of defined protein conjugates by site-directed cotranslational labeling, in: W.A. Kudlicki, F. Katzen, R.P. Bennett (Eds.), Cell-Free Protein Expression, Landes Bioscience, Austin, 2007, pp. 166–180.
[32] S. Sando, A. Ogawa, T. Nishi et al., *In vitro* selection of RNA aptamer against *Escherichia coli* release factor 1, Bioorg. Med. Chem. Lett. 17 (2007) 1216–1220.
[33] K.A. Daggett, M. Layer, T.A. Cropp, A general method for scanning unnatural amino acid mutagenesis, ACS Chem. Biol. 4 (2009) 109-113.

Appendix A

Supplementary Material

High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at multiple sites

Kiyoshi Ozawa, Karin V. Loscha, Kekini V. Kuppan, Choy Theng Loh, Nicholas E. Dixon, Gottfried Otting

Table S1. PCR and mutagenesis primers used in this study[1]

1131	5'-PO_4-TTAGCTGGTCGATCCCGCGAAATTAATACG-3' (30-mer)
1132	5'-PO_4-CCAGCTAACAAAAAACCCCTCAAGACCCG-3' (29-mer)
1133	5'-PO_4-TCGATCCCGCGAAATTAATACG-3' (22-mer)
1134	5'-PO_4-CAAAAAACCCCTCAAGACCCG-3' (21-mer)
1375	5'-TCGATCCCGCGAAATTAATACGACTCAC-3' (28-mer)
1376	5'-CAAAAAACCCCTCAAGACCCGTTTAGAG-3' (28-mer)
1398	5'-TTTTTTTTCATATGGACGAATTTGAAATG-3' (29-mer)
1402	5'-TTTACGCGTTTATAATCTCTTTCTAATTGG-3' (30-mer)
1447	5'-TTTTTTTTCATATGGACGAATTTGAAATGATAAAGAG-3' (37-mer)
1448	5'-TTTGAATTCTTATAATCTCTTTCTAATTGGCTCTAAAATC-3' (40-mer)
2230[2]	5'-GATTGTCGTG**TAG**CCAGGG**TAG**AATGTGAAAAACGTC-3' (37-mer)
2231[2]	5'-GTTTTTCACATT**CTA**CCCTGG**CTA**CACGACAATCATTTG-3' (39-mer)
2571[3]	5'-GAATCCCGCGTC**TAG**AAAATTGCCGAAAG-3' (29-mer)
2572[3]	5'-CTTTCGGCAATTTT**CTA**GACGCGGGATTC-3' (29-mer)

[1]
[2] Codons of mutated amino acids are underlined. Mutated bases are shown in bold.
[3] Forward and reverse primers to generate the Glu101TAG/Lys104TAG double mutant of WNVpro.
Forward and reverse primers to generate the Phe166TAG mutant of *E. coli* DnaB.

Synthetic protocols

Expression constructs of *Mj*TyrRS, BpaRS, AzF-RS, AcF-RS, CNF-RS and HcoRS

The wild-type *Mj*TyrRS gene (about 1.1 kb) was amplified from the plasmid pSup-MjTyrRS-6TRN using the primers 1398 and 1402, and cloned into the *Nde*I and *Mlu*I sites of pETMCSI [1]. The resulting T7-expression vector was termed pKO1377.

The mutant *Mj*TyrRS gene encoding BpaRS was amplified from the plasmid pSup-BpaRS-6TRN using the primers 1447 and 1448, and cloned into the *Nde*I and *Eco*RI sites of

pETMCSIII [1], resulting in N-terminally His_6-tagged BpaRS in the T7-expression vector pKO1349.

The gene encoding AzF-RS was amplified from the plasmid pEVOL-AzF [2] using the primers 1398 and 1402, and cloned into the *Nde*I and *Mlu*I sites of pETMCSIII to yield the T7-vector pKO1350.

The gene encoding AcF-RS was amplified from pEVOL-AcF [2] using the primers 1447 and 1448, and cloned into the *Nde*I and *Eco*RI sites of pETMCSIII to yield the T7-vector pKO1420.

The CNF-RS gene was PCR-amplified from pEVOL-pCNF [2] using the primers 1398 and 1402, and cloned into the *Nde*I and *Mlu*I sites of pETMCSIII, resulting in the T7-expression vector pCT1537.

The gene encoding HcoRS (CouRS-D8 mutant) was excised with *Nde*I and *Eco*RI from pETMCSI-HcoRS [3], and cloned into the same sites of pETMCSIII to add an N-terminal His_6-tag, resulting in the T7-vector pETMCSIII-HcoRS. The enzyme corresponds to the CouRS-D8 mutant described by Wang et al. [4].

Only the enzymes BpaRS, AzF-RS and HcoRS contained the D286R mutation that enhances the recognition of the tRNA by the RS enzymes [5].

Expression and purification of RS enzymes

The expression and purification of N-terminally His_6-tagged BpyRS was described previously [6].

Wild-type *Mj*TyrRS was produced by growing *E. coli* cells (BL21 (DE3)/pLysS) harbouring pKO1377 aerobically for two days at room temperature in an auto-induction medium [7] supplemented with 100 mg/L ampicillin and 34 mg/L chloramphenicol; 2 liters of cell culture yielded about 15.4 g of cells. The supernatant of the French press lysate was subjected to an ammonium sulfate (AS) precipitation (0.35 g AS per mL supernatant). The supernatant of the AS precipitation was dialyzed against DEAE buffer (50 mM Tris-HCl, pH 7.9, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 10% glycerol) and subsequently loaded onto a DEAE-Toyopearl 650 M column (2.5 x 13 cm). Bound *Mj*TyrRS was eluted with a gradient of 0–600 mM NaCl in the same buffer. The *Mj*TyrRS containing fractions (analyzed by SDS-PAGE) were combined and dialyzed against P-buffer (10 mM sodium phosphate, pH 6.8, 1 mM EDTA and 1 mM DTT). The dialyzed sample was loaded onto a phosphocellulose column (2.5 x 4.3 cm) and bound *Mj*TyrRS was eluted with a gradient of 0–1M NaCl in the same buffer, yielding about 3.1 mg *Mj*TyrRS (>95% purity indicated by SDS-PAGE).

His_6-tagged BpaRS was produced by growing BL21 (DE3)/pLysS/pKO1349 aerobically at 30 ˚C in 1 L of LB medium supplemented with 25 mg/L thymine, 100 mg/L ampicillin and 34 mg/L chloramphenicol. Overexpression of BpaRS was induced by addition of 1 mM isopropyl-β-D-1-thiogalactoside (IPTG) at $A_{595} = 0.56$ and the induced culture was grown to $A_{595} = 1.23$ (3 h), yielding 3.55 g of cells. About 116 mg of purified BpaRS was obtained by loading the supernatant of the French press lysate onto a 5-mL column of Ni-NTA (Pharmacia, Uppsala, Sweden) in buffer N (50 mM HEPES–KOH buffer, pH 7.5, 300 mM NaCl, 20 mM imidazole,

5% (v/v) glycerol). The elution used a linear imidazole gradient (20 to 500 mM) in buffer N. The concentrated BpaRS (664 μM) was stored at –70 ˚C in storage buffer (50 mM Tris-HCl, 2 mM DTT, 100 mM NaCl) at pH 7.6.

His_6-tagged AzF-RS and AcF-RS were produced by aerobically growing *E. coli* cells (BL21 (DE3)/pLysS) harboring pKO1350 or pKO1420, respectively, for two days at room temperature in an auto-induction medium [7] supplemented with 100 mg/L ampicillin and 34 mg/L chloramphenicol; 2 liters of cell culture yielded about 18 g and 26 g of cells, respectively. The supernatant of the French press lysate was loaded onto a 5-mL column of Ni-NTA resin in buffer N. The synthetases were eluted with a gradient of 20–500 mM imidazole in buffer N, yielding about 206 mg of AzF-RS and 197 mg of AcF-RS. Concentrated solutions of AzF-RS (1.18 mM) and AcF-RS (2.21 mM) were stored at –70 ˚C in storage buffer at pH 6.5 and pH 7.0, respectively.

His_6-tagged CNF-RS and HcoRS were produced by growing BL21 (DE3)/pLysS/pCT1537 and BL21(DE3)/pLysS/pETMCSIII-HcoRS, respectively, using the same protocol as for AzF-RS.

The concentrations of purified proteins were determined using calculated UV absorption coefficients [8].

References

[1] C. Neylon, S.E. Brown, A.V. Kralicek et al., Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: A model derived from DNA-binding studies of mutant proteins by surface plasmon resonance, Biochemistry 39 (2000) 11989–11999.

[2] T.S. Young, I. Ahmad, J.A. Yin et al., An enhanced system for unnatural amino acid mutagenesis in *E. coli*, J. Mol. Biol. 395 (2010) 361–374.

[3] I.N. Ugwumba, K. Ozawa, L. de la Cruz et al., Using a genetically encoded fluorescent amino acid as a site-specific probe to detect binding of low-molecular weight compounds, Assay Drug Dev. Technol. 9 (2011) 50–57.

[4] J. Wang, J. Xie, P.G. Schultz, A genetically encoded fluorescent amino acid, J. Am. Chem. Soc. 128 (2006) 8738-8739.

[5] T. Kobayashi, O. Nureki, R. Ishitani et al., Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion, Nat. Struct. Biol. 10 (2003) 425–432.

[6] T.H.D. Nguyen, K. Ozawa, M. Stanton-Cook et al., Pseudocontact shifts in protein NMR spectra generated by a genetically encoded Co^{2+}-binding amino acid, Angew. Chem. Int. Ed. 50 (2011) 692–694.

[7] F.W. Studier, Protein production by auto-induction in high density shaking cultures, Protein Express. Purif. 41 (2005) 207–234.

[8] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 319–326.