ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Genetically Expanded Cell-Free Protein Synthesis Using Endogenous Pyrrolysyl Orthogonal Translation System

Yonatan Chemla,[1] Eden Ozer,[1] Orr Schlesinger,[1] Vincent Noireaux,[2] Lital Alfonta[1]

[1]Department of Life Sciences and Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel; telephone: +972 8 6479066; fax: +972 8 6479576; e-mail: alfontal@bgu.ac.il

[2]School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55401

ABSTRACT: Cell-free protein synthesis offers a facile and rapid method for synthesizing, monitoring, analyzing, and purifying proteins from a DNA template. At the same time, genetic code expansion methods are gaining attention due to their ability to site-specifically incorporate unnatural amino acids (UAAs) into proteins via ribosomal translation. These systems are based on the exogenous addition of an orthogonal translation system (OTS), comprising an orthogonal tRNA, and orthogonal aminoacyl tRNA synthetase (aaRS), to the cell-free reaction mixture. However, these components are unstable and their preparation is labor-intensive, hence introducing a major challenge to the system. Here, we report on an approach that significantly reduces the complexity, effort and time needed to express UAA-containing proteins while increasing stability and realizing maximal suppression efficiency. We demonstrate an endogenously introduced orthogonal pair that enables the use of the valuable yet insoluble pyrrolysyl-tRNA synthetase in a cell-free system, thereby expanding the genetic repertoire that can be utilized in vitro and enabling new possibilities for bioengineering. With the high stability and efficiency of our system, we offer an improved and accessible platform for UAA incorporation into proteins.

Biotechnol. Bioeng. 2015;9999: 1–11.

KEYWORDS: amber suppression; cell-free protein synthesis; genetic code expansion; pyrrolysyl-tRNA synthetase

Introduction

The ability to produce a mature and functional protein without the integrity of a living cell is the fundamental principal of cell-free protein synthesis (CFPS) (Rosenblum and Cooperman, 2014). Recent surge in CFPS systems can be attributed to their unique advantages over in vivo methodologies, including fast expression of recombinant proteins from DNA templates (Carlson et al., 2012; Katzen et al., 2005; Swartz, 2006; Zawada et al., 2011), high relative yields of protein produced (up to 6 mg/mL) (Caschera and Noireaux, 2014; Goerke and Swartz, 2009; Kigawa et al., 1999; Spirin et al., 1988; Sun et al., 2013), and high throughput protein screening without a need for gene cloning and other time-consuming steps (Kanter et al., 2007; Yabuki et al., 2007). Moreover, the absence of integral cells allows for manipulation and precise control over the micro-environment, the production and use of toxic (Renesto and Raoult, 2003; Xu et al., 2005), and membrane-non-impermeable molecules, facile monitoring that enables immediate feedback and manipulation of the protein synthesis process. In light of these advantages, CFPS methodologies are being constantly improved and rendered more accessible, thus allowing the production of "hard-to-express" proteins (Jun et al., 2008; Kai et al., 2013; Niwa et al., 2012; Sasaki et al., 2011; Zawada et al., 2011), amino acid replacement, in vitro directed evolution (Fujii et al., 2013; Tawfik and Griffiths, 1998), enzyme bioengineering (Catherine et al., 2013; Venancio-Marques et al., 2012), biotechnology (Ahn et al., 2011; Daube and Bar-Ziv, 2013; Heyman et al., 2012; Tan et al., 2013), and synthetic biology (Shin and Noireaux, 2012).

"Genetic code expansion" corresponds to an experimental attempt to increase the amino acid repertoire that can be incorporated into proteins during ribosome-mediated translation (Budisa, 2004). This approach is based on suppression of the TAG (amber) stop codon, i.e., amber suppression (Wang et al., 2001). To obtain a protein with a site-specifically incorporated UAA, three components are required (other than the gene encoding for the desired protein and the UAA), namely a TAG codon at the location of the new amino acid, an orthogonal amino acyl tRNA synthetase (o-aaRS), and an orthogonal tRNA (o-tRNA). These last two

The authors declare no competing financial interest.
Correspondence to: Lital Alfonta
Contract grant sponsor: European Research Council (ERC)
Contract grant number: 260647
Received 23 December 2014; Revision received 19 February 2015; Accepted 23 February 2015
Accepted manuscript online xx Month 2015;
Article first published online in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/bit.25587

components were termed the orthogonal translation system (OTS) (Hong et al., 2014). One limitation of the TAG suppression approach is that it competes with the endogenous release factor 1 (RF1), an essential factor responsible for termination of translation when the TAG (amber) and TAA (ochre) nonsense codons are encoded. To overcome this limitation, the genomically recoded organism (GRO) *Escherichia coli* strain C321.Δ*prf*A (Lajoie et al., 2013) was developed. In this strain genome, all TAG stop codons are replaced with synonymous TAA stop codons, which permit the deletion of release factor 1 (RF1) and a reassignment of the TAG translation function. As ochre termination is shared with RF2 (Kisselev and Buckingham, 2000), no significant interference with such termination is expected in an organism lacking RF1 (for a simplified explanation, see the highlights video).

When attempts were made to express proteins containing UAAs using CFPS systems, the use of exogenously added o-tRNA (synthetic or native) and purified o-aaRS was the only way reported to achieve cell-free genetic expansion thus far (Albayrak and Swartz, 2013; Goerke and Swartz, 2009; Smolskaya et al., 2013). By controlling the concentration of the orthogonal o-tRNA and o-aaRS translational pair added to the reaction mixture and more recently by using partially recoded and RF1-deficient *E. coli* strains (Hong et al., 2014), suppression efficiency was significantly increased. In addition, initial attempts at in situ production of the orthogonal translational components were reported. In situ expression of o-tRNA using ribozyme-based method (Albayrak and Swartz, 2013) was recently reported and achieved increased yields. Since, it has also been shown that it is possible to achieve incorporation of UAAs using the *Methanocaldococcus jannaschii* (*Mj*) *Mj*-tyrosyl-amino-acyl-tRNA synthetase/tRNA orthogonal pair (Smith et al., 2014).

Yet, major limitations of cell-free genetic expansion methods remain. For instance, insoluble aaRSs purification is highly challenging. As a consequence, cell-free genetically expanded protein synthesis is still limited to the production of soluble and easily over-expressed aaRSs, thus severely limiting the genetic expansion repertoire. To the best of our knowledge, pyrrolysyl synthetase and its mutants (Tuley et al., 2014), which to date facilitate the incorporation of at least 101 UAAs (Wan et al., 2014) comprising more than 52% of the known genetic code expansion repertoire (Dumas et al., 2014), have a highly complex purification process. As a result, pyrrolysyl synthetase has never been used in cell-free systems, thus preventing the use of a large majority of UAAs from cell-free usage. Another outstanding problem is the labor and time-intensive nature of aaRS expression and purification steps, as well as tRNA synthesis, which, if performed correctly, can require several days. Furthermore, even when an orthogonal pair is synthesized and purified, it can only be stored in its active form for a relatively short period of time. Because these two essential components are synthesized and added exogenously, more levels of complication are introduced into the system, thus reducing the reproducibility and consistency of the results and products attained.

We believe that these limitations present a significant barrier to the widespread use of cell-free genetically expanded protein synthesis. As such, reducing this barrier was the initial motivation of the present study. To tackle these limitations, we hypothesized that the introduction and induction of the *Methanosarcina mazei* (Mm) orthogonal pair *Mm*-PylRS/*Mm*-tRNA$_{CUA}^{pyl}$ (i.e., Pyl-OTS) into the GRO strain prior to cell lysis would result in the creation of cell extract that would facilitate cell-free TAG stop codon (amber) suppression (Fig. 1) and enable the utilization of the useful, albeit insoluble Pyl-OTS for cell-free genetic code expansion. Figure 1 offers a schematic representation of the different steps taken using our approach. Initially, a plasmid containing the Pyl-OTS components is transformed into different *E. coli* strains (Fig. 1a). This step is followed by lysis and preparation of active extracts (Fig. 1b). The active extracts are used for the cell-free protein synthesis reaction upon addition of the gene of interest (in this case encoding deGFP), as well as the UAA (Fig. 1c). This results in production of full-length fluorescent deGFP containing a UAA (Fig. 1d).

Materials and Methods

Bacterial Strains and Plasmids Transformation

Strains and plasmids used in this study are listed in Table (supporting information, SI section). All transformations of the C321.Δ*prf*A, C321 (RF1+), and BL21 strains were done by electroporation. Parent strains not containing any plasmids were grown in Luria-Bertani (LB) broth (10 g/L NaCl, 10 g/L trypton, and 5 g/L yeast extract) overnight at 30°C (C321 derivatives) or 37°C (BL21 and DH5α strains) for sequential inoculation. Strains were made competent as follows: the cultures were diluted 1/100 in fresh LB and incubated at the relevant temperature while being shaken at 275 rpm to OD_{600} of 0.5–0.7. Cells were then harvested and washed with 10% glycerol in water three times, aliquoted, and stored at −80°C until thawed for transformation. Cells were transformed with DNA using 50–100 ng/μL of template DNA (obtained by mini-prep) and then transformed by electroporation using a MictroPulser (Bio-Rad, Hercules, CA). Transformed cells were then incubated for 1–1.5 h in SOC broth (2% bacto-tryptone, 0.5% bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) at the appropriate temperature and sequentially plated on selective LB-agar plates containing the proper antibiotic. Competent DH5α *E. coli* cells (New England Biolabs, Ipswich, MA) were transformed using the prescribed heat shock protocol and sequentially plated on selective LB-agar plates. All transformed strains are described in Table S2. All relevant gene sequences can be found in the supporting information section (SI).

Cell Extract Preparation

The S30 cell extract (crude extract) was prepared based on research by Kigawa et al. (2004) and Liu et al. (2005) but with significant modifications described in detail and in a visualized protocol by Sun et al. (2013). We have used the exact same methods, reagents, and apparatus as the latter protocol for all bacterial strains described in Table S1. The changes made by us are as follows: we adjusted the selection antibiotics (according to the plasmid used) and the temperature of growth, which influenced the growth incubation times of inoculated cultures. The rationale behind this was to have the different strains undergo the same number of generations. For example, a strain with a doubling time of ∼30 min was incubated for ∼8 h, whereas a strain with a doubling time of 50 min was

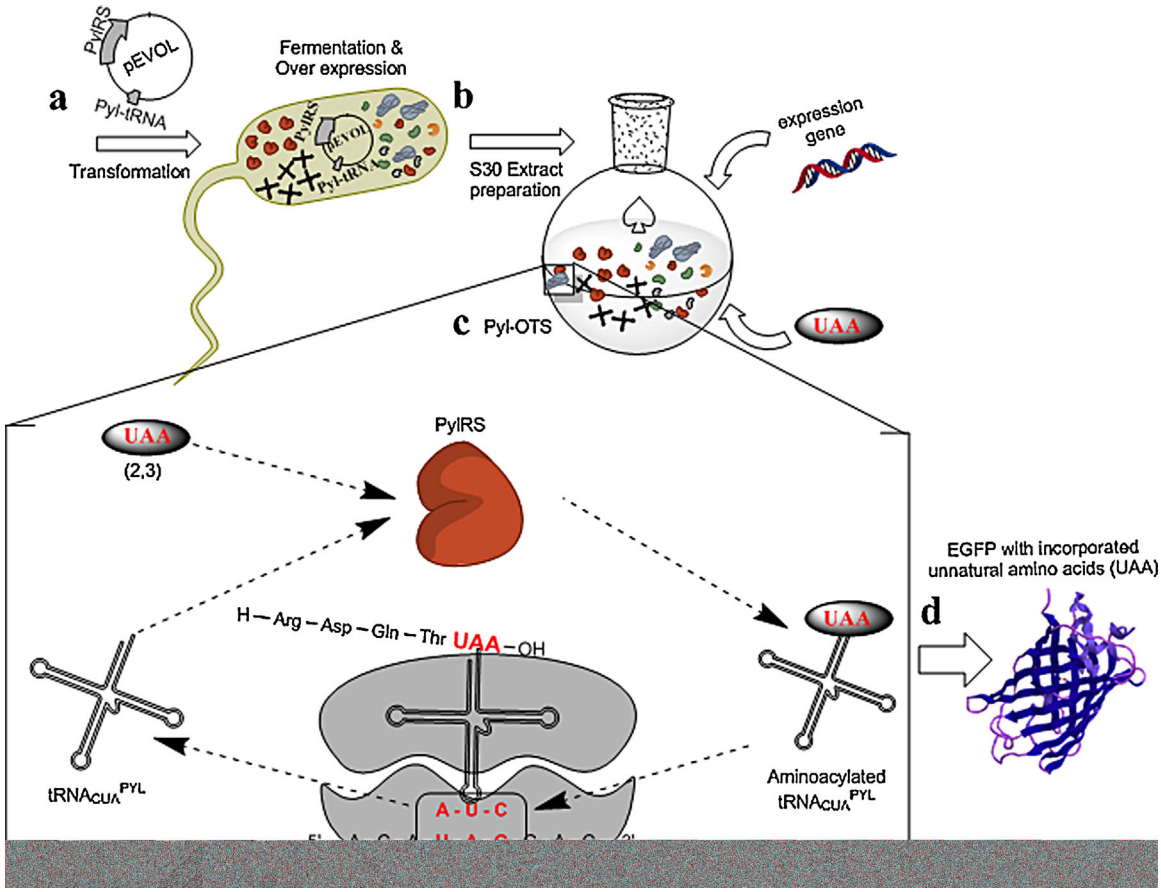

Figure 1. Schematic presentation of an endogenous Pyl-OTS *E. coli* extract. **a)** Transformation of the orthogonal pair (OTS) into *E. coli* strains. **b)** Lysis under special conditions and genetically expanded cell extract preparation. **c)** Cell-free protein synthesis using the produced extract with the addition of the desired gene (deGFP in this manuscript) and the unnatural amino acid. **d)** The result is the facile and rapid synthesis of a genetically expanded protein.

incubated for ~13 h. Additionally, the promoter regulating the expression of the aaRS was induced in early log phase of growth (OD_{600} 0.5–0.7) for pEVOL, 0.5–1% L-arabinose was added, while for plasmid pKD, 1 mM IPTG was added; no induction was performed for pSUP that lacks an inducible promoter, resulting in over-expression of the aaRS, thus enabling cell-free o-tRNA amino-acetylation once exogenous UAA is introduced.

CFPS Expression Plasmids and Recombinant Gene Design and Construction

The CFPS expression plasmid pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (Addgene #40019) used in this study was described previously (Shin and Noireaux, 2012). The plasmid was used for cell-free protein synthesis, based on the activities of endogenous core RNA polymerase and sigma factor 70. All plasmids were transformed and grown in *E. coli* DH5α cells and harvested using a Qiaprep spin miniprep kit (Qiagen, Hilden, Germany). To mutate various codons in the deGFP gene to amber nonsense codons, a KAPA HiFi PCR Kit was employed (Kapa Biosynthesis, Wilmington, MA), together with a thermo-cycler (Bioer Technologies, Hangzhou, China). The resulting mutated PCR product was then transformed into competent *E. coli* DH5α cells and plated onto selective plates to isolate transformed colonies. Transformed colonies were sequentially incubated overnight and their plasmids were harvested and sequenced. All relevant gene sequences are provided in the SI section

Cell-Free Protein Synthesis

Cell-free reactions were carried out in volumes of 10 μL at 29°C. The 3-PGA reaction buffer is composed of 50 mM HEPES, pH 8, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM coenzyme A, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-phosphoglyceric acid, 1.5 mM each of 20 amino acids, 1 mM DTT, and 2% PEG-8000. A typical cell-free reaction with this system contained 33% (by volume) *E. coli* extract, corresponding to a protein concentration of 10–15 mg/mL, before target protein synthesis. The other 66% of the reaction volume are composed of the plasmids, reaction buffer containing nutrients and the UAA. The concentrations of all reagents in the reaction buffer were fixed, except for magnesium glutamate and potassium glutamate, containing two essential ions for CFPS and molecular interactions involved in transcription and

translation. The cell-free expression system was prepared so as to adjust the concentrations of these two ions for a given strain extract. Optimization of these ion concentrations was achieved by expressing deGFP in the CFPS reactions with a gradient (1–6 mM) of Mg-glutamate while fixing the K-glutamate concentration at 80 mM. Mg-glutamate concentration which exhibited the highest fluorescence was chosen. Sequentially the optimization of the K-glutamate concentration between a gradient of 20–140 mM was performed in the same manner while fixing the Mg-glutamate to the chosen concentration. Detailed protocol of the reaction buffer preparation and the CFPS reaction preparation (including easy to use and formatted excel files) are available at Sun et al. (2013).

For genetic code expansion via amber suppression, the exogenous addition of the relevant UAA (i.e., that UAA recognized by the o-aaRS and charged on the o-tRNA in the extract) is needed. In our system and from past experience (Smolskaya et al., 2013), the optimal UAA concentration is 1 mM. Here, we added various UAAs to different stain extracts containing different endogenous OTSs (L-phospho-serine [Sigma-Aldrich, Rehovot, Israel], L-para-acetyl-L-phenylalanine [synthesized as described by Takasu et al. (2011)], propargyl-L-lysine [synthesized as described by Nguyen et al. (2009)] and N^{ϵ}-Boc-L-lysine [Sigma–Aldrich, Rehovot, Israel]). It is likely more convenient for large-scale use of our system with a single strain extract to add the UAA to the reaction enhancement buffer at a final concentration of 1 mM.

When expressing deGFP, the reaction components were added to A Nunc 384 (120 μL) well plates (Thermo Fisher Scientific, Waltham, MA) in order to sample deGFP expression as reflected by fluorescence intensity periodically (every 30 min). Reactions were incubated for 10 h for optimal deGFP fluorescence.

deGFP Quantification

deGFP fluorescence was measured during the CFPS reaction, with quantification achieved with a deGFP calibration curve (Fig. S6) assembled from readings obtained from pure EGFP (MBL International, Woburn, MA). In addition, a second quantification approach was used; Y35(**2**) and WT deGFP were purified using nickel affinity chromatography. The resulting samples were measured using a commercial Bradford assay (BCA Protein Assay, Thermo Scientific, Waltham, MA).

deGFP Purification and Mass Spectrometry

For LC-MS validation of incorporation of **2**, nickel affinity chromatography purification of 6xhis-tagged deGFP was performed. Five-hundred microliters of CFPS reaction mixture was incubated overnight at 29°C to produce either deGFP Y35X (N-terminal 6xhis tag) or WT deGFP (N-terminal 6xhis tag). The reaction mixture was then diluted with 3 volumes of PB buffer (50 mM PB pH 8, 0.3 mM NaCl, and 10 mM imidazole) and added to a nickel-bead column (Novagen, Madison, WI). Wash (50 mM imidazole) and elution (250 mM imidazole) steps were conducted according to the manufacturer's instructions. The protein-containing eluted fraction was concentrated using a Vivaspin 10 kDa cutoff concentrator (Sartorius, Göttingen, Germany). The resulting concentrated fraction was analyzed by LC-MS (Finnigan Surveyor Autosample Plus/LCQ Fleet, Thermo Scientific, Waltham, MA).

For the click chemistry downstream reaction, size exclusion-based purification of untagged proteins was performed. One-hundred and twenty microliters of CFPS reaction mixture was incubated overnight at 29° to produce deGFP Y35X incorporating (2). The reaction mixture was then diluted (×10) by DDW and subjected to size exclusion chromatography using an ÄKTA apparatus (GE Healthcare, Tel-Aviv, Israel) and the relevant 8 mL fraction was collected. The relevant fractions were determined prior to the purification of the reaction mixture by using commercially purified EGFP (MBL International, Woburn, MA). The fraction was concentrated using a Vivaspin 10 kDa cutoff concentrator (Sartorius, Göttingen, Germany). The resulting concentrated fraction was used for a "click" reaction.

"Click" Reaction

deGFP containing (**2**) was labeled using the Cu(I)-catalyzed azide-alkyne cycloaddition reaction (CuAAC). The protein sample was resuspended in 0.1 M PB, pH 7.5. Tetramethylrhodamine-azide (Tamra-Az) (Sigma-Aldrich, Rehovot, Israel) was added to a concentration of 100 μM. 3,3′,3″-(4,4′,4″-(Nitrilotris(methylene))tris(1H-1,2,3-triazole-4,1-diyl))tris(propan-1-ol) (THPTA), sodium ascorbate, and $CuCl_2$ were added to final concentrations of 400 μM, 2.5 mM, and 200 μM, respectively. The reaction mixture was incubated at room temperature for 1 h. A 20 μL sample was diluted with 4X SDS sample buffer and kept for 10 min at 70°C, after which it was loaded and run on a 12% SDS–PAGE gel. Labeled proteins were visualized in-gel using an ImageQuant LAS 4000 imager (Fujifilm, Tokyo, Japan) set in the Cy3 fluorescence mode (Ex/Em = 520 nm/575 nm).

Results and Discussion

Our main purpose in this study was to establish a methodology that enables the in vitro utilization of Pyl-OTS while maintaining simplicity by creating an endogenous and genetically expanded CFPS system. We sought to confirm the methodology by incorporating two different pyrrolysine Fig. 2, amino acid 1 analogs, Ne-Propargyl-L-lysine (Figure 2, 2) and Ne-Boc-L-lysine (Figure 2, 3) (Fig. 2, amino acids **2**, and **3**) into a model protein deGFP [an EGFP variant optimized for CFPS (Shin and Noireaux, 2010)]. We assessed the reproducibility of the method and further characterized it in a comparative study between different *E. coli* strains with or without RF1.

System Validation; Cell-Free Incorporation of Pyrrolysine Derivatives Into deGFP Using an Endogenously Encoded Pyrrolysyl-tRNA Synthetase/$tRNA_{CUA}^{Pyl}$ (Pyl-OTS)

The recoded *E. coli* strain (GRO):C321.Δ*prf*A generated by Lajoie et al. (2013) was chosen as the best candidate for our system. This strain has been developed to replace the message encoded by the amber nonsense codon, making it the obvious choice upon which we based our efforts. A cell-free system in which genomic TAG stop codons are replaced will direct the translation of a UAG triplet that is

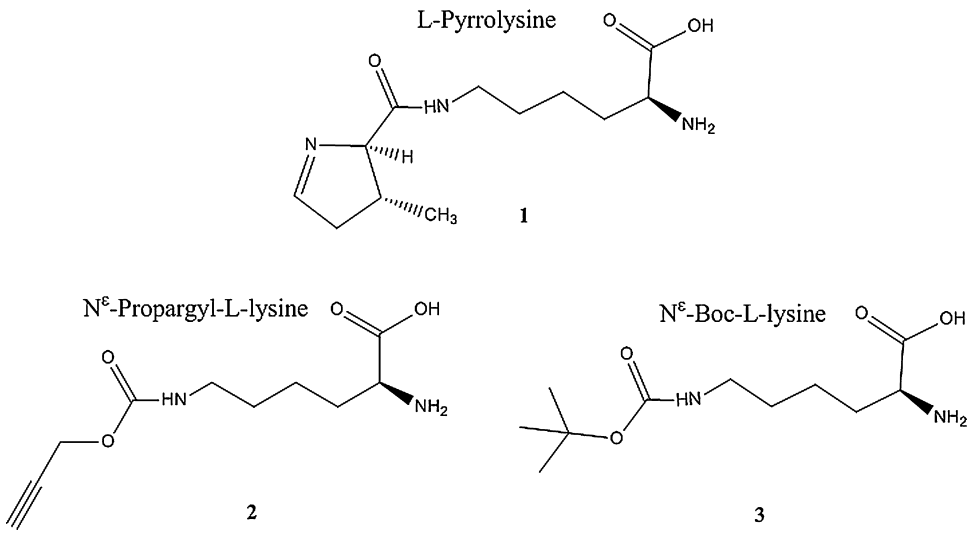

Figure 2. Chemical Structures of **1**: L-pyrrolysine, the natural substrate of the Pyl-OTS; **2**: N^{ϵ}-Propargyl-L-lysine and **3**: N^{ϵ}-Boc-L-lysine.

completely replaced, now being translated as a sense codon instead of a nonsense codon. Upon the endogenous addition of the OTS, no addition of exogenous components (other than the UAA and gene of interest) is necessary. Hence, this novel platform truly corresponds to a "genetically expanded cell-free protein synthesis."

To confirm the functionality of the proposed system, the medium-low copy number pEVOL (Young et al., 2010) plasmid containing the OTS genes mM-PylRS/mM-$tRNA_{CUA}^{pyl}$ was transformed into the GRO strain. Since amber suppression can be toxic for the host cell (Young et al., 2010), we tested the effects and toxicity of the Pyl-OTS harbored in the GRO strain (Fig. S1). Changes in growth rates with/without inducer and in the face of different inducer concentrations were measured. In agreement with previous findings (Hong et al., 2014; Lajoie et al., 2013), no toxicity was observed. Next, we prepared Pyl-OTS containing *E. coli* S30 cell extracts and reaction enhancement buffer (containing amino acids, energy sources, nucleotides, co-factors, and crowding agents) needed for the subsequent CFPS reaction (see materials and methods section). Upon establishment, the system was tested for the incorporation of two pyrrolysine analogs; (**2**) & (**3**) (Fig. 2) into deGFP [both analogs are known to be recognized by Pyl-OTS (Li et al., 2009; Wan et al., 2014)]. Accordingly, the deGFP-encoding gene was sub-cloned into the pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid (Table S1) under the control of a mutated bacteriophage λ promoter (OR2-OR1-Pr), additionally, we have mutated the Y35X codon position of deGFP (where X denotes TAG). At t=0, the pBEST deGFP plasmid was added to the reaction mixture. To each mixture, a different concentration of UAA, (either **2** or **3**) was added. As a negative control, no UAA was added, while as a positive control, WT deGFP (without amber mutations) was used. Reaction mixtures were incubated and the subsequent fluorescence of the deGFP produced was monitored in real time using a fluorescence plate reader. Figure 3, panel A shows an increase in fluorescence resulting from the expression and maturation of a full-length deGFP containing **2**.

The results show the UAA concentration-dependence of deGFP fluorescence. The UAA and orthogonal tRNA are specifically recognized by PylRS, reflected in ribosomal translation of the TAG codon as a sense codon. In the absence of the UAA (as is the case in the negative control), the ribosome will either stall, causing the nascent polypeptide to be degraded or truncated and triggering mRNA cleavage cascade (Hayes and Sauer, 2003; Liu et al., 2011; Sunohara et al., 2004), or rarely, "read-through" the nonsense codon (i.e., a mis-incorporation of a natural amino acid at the amber mutation position). This proposed mechanism is further strengthened by the observation that the expression levels of UAA-incorporating deGFP are similar to the expression levels of the WT deGFP positive control, yet significantly higher than that of the negative control. The relatively low fluorescence (∼130 [AU]; Fig. 3 panel A) of the negative control in the absence of the UAA is partially due to background noise of the reaction mixture itself but mainly due to ribosomal read-through (background suppression) (O'Donoghue et al., 2012). It is important to note that in the presence of the UAA no background suppression is observed as was verified in by ESI, vide infra (Fig. 4). Examining the fluorescence results revealed that optimal protein yields were attained with 1 or 2 mM UAA, in agreement with previous findings (Young et al., 2010). Henceforth, we used a UAA concentration of 1 mM.

Due to the possibility that incorporation of UAA and changing the native amino acid could result in decreased fluorescence of the deGFP model protein rather than eliciting a decrease in protein yields, Western blot analysis with an anti-GFP antibodies that allows for assessment of protein expression levels along with fluorescence was performed. Anti-GFP immunoblotting was, moreover, performed as an additional confirmation of the cell-

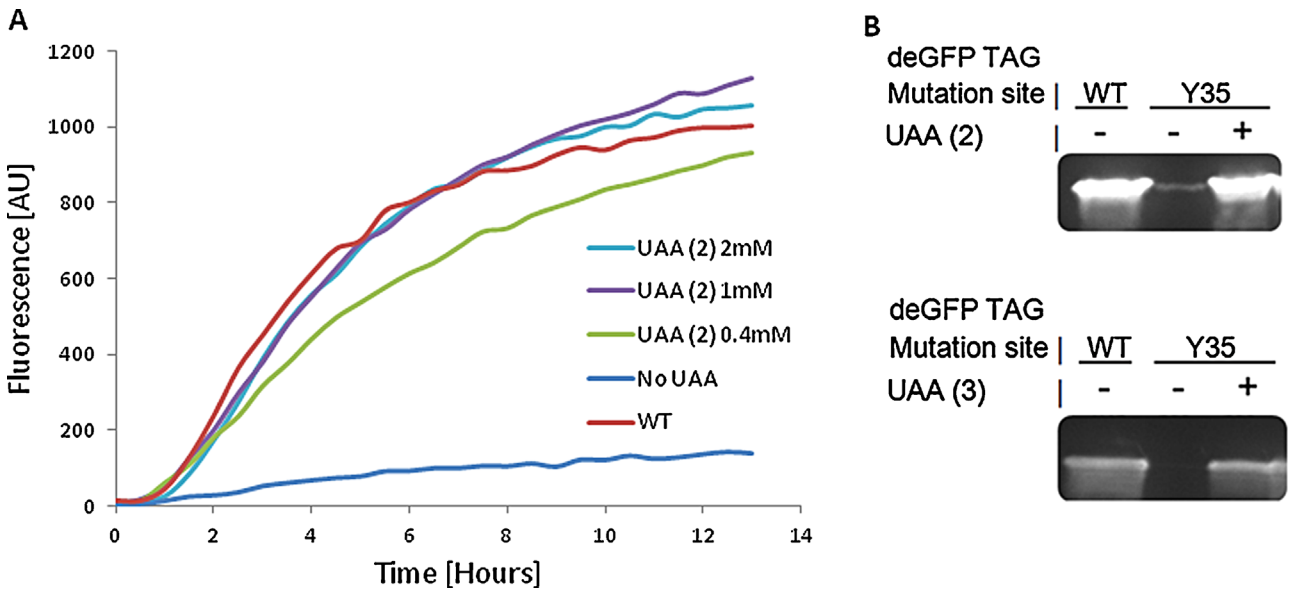

Figure 3. Genetically expanded (Pyl-OTS) cell-free protein synthesis of deGFP. **(A)** Expression kinetics of Y35X deGFP, measured as fluorescence intensity, in the presence of varying concentrations of (**2**) (**B**) Anti-GFP Western blot comparing the reaction results between WT deGFP and Y35X deGFP with or without 1 mM of (**2**) or (**3**).

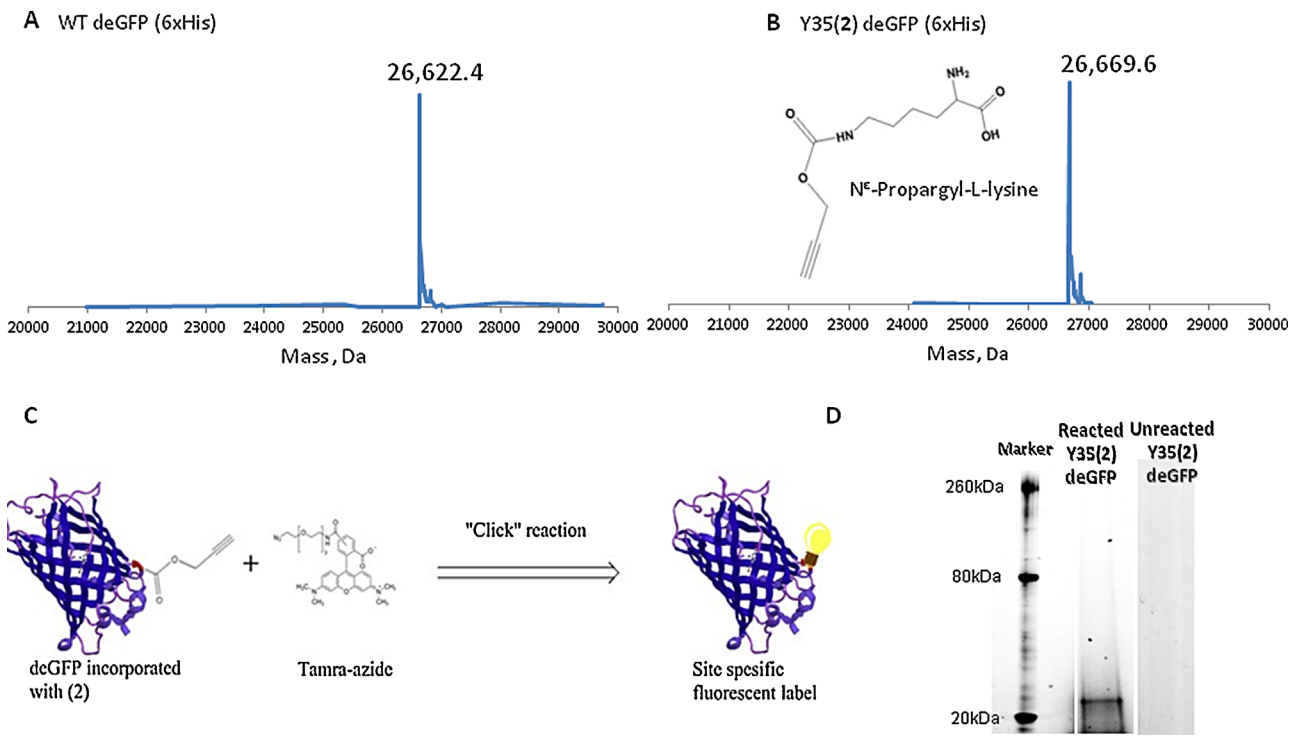

Figure 4. Verification of incorporation of (**2**) into deGFP. **A)** Deconvoluted ESI mass spectrum of purified WT deGFP; **B)** Deconvoluted ESI mass spectrum of purified deGFP Y35X (with incorporated **2**) **C)** Illustration of site-specific incorporation of (**2**) into deGFP and a sequential "Click" reaction to Tamra-azide fluorescent dye. **D)** Image of SDS-PAGE containing cell-free purified deGFP including (**2**) at position 35. Left lane contains Y35(**2**) deGFP after a "click" reaction with Tamra-azide, right lane contains un-reacted Y35(**2**) deGFP under the same experimental conditions.

free incorporation of (**2**) and (**3**) into deGFP (Fig. 3, panel B). The results show similar bands (that can be correlated to expression) of the positive control and UAA-incorporating deGFP, while significantly weaker negative control bands were obtained. As mentioned before, the weak negative control bands are possibly due to ribosomal read-through, which may occur in the absence of a UAA. Finally, the system was successfully tested for incorporation of (**2**) into two additional sites (K136X and D193X) in deGFP (Fig. S3) and successfully tested for incorporation of (**2**) into two different sites in the same protein (Fig. S4) thus corroborating the system's generality for UAA incorporation into various sites.

Verification of the Incorporation of (2) Via Mass Spectrometry and a "Click" Reaction

Using electrospray ionization-mass spectrometry (ESI-MS), we compared the masses of WT deGFP that was expressed using our methodology (Fig. 4, panel A) with that of deGFP containing (**2**) (Y35X) (Fig. 4, panel B). The aim of this comparison was to verify the correct incorporation of UAA and to exclude the possibility of a ribosomal read-through (i.e., background suppression) of the system when the protein was expressed in the presence of (**2**). Furthermore, we wanted to confirm the incorporation of the bio-orthogonal propargyl functional group into deGFP by subjecting cell-free-produced and purified deGFP to a catalyzed Huisgen 1,3-cyclo-addition ("click") reaction. In this reaction, the fluorescent Tamra-azide dye (shown in Fig. 4) was ligated to deGFP site specifically where (**2**) had been incorporated (Fig. 4, panel C). The results of this reaction were observed using SDS–PAGE in gel fluorescence analysis using specific filters for the Tamra fluorophore (excitation: 520 nm; emission 575 nm); (Fig. 4, panel D).

The MS results show a mass difference of 47.1 Da, a value that is in a good agreement with the calculated mass difference between deGFP containing (**2**) and WT deGFP with a tyrosine at position 35, which is a difference of 47 Da. Furthermore, only a single peak which corresponds to a total mass of $26{,}669.6 \pm 2.2$ Da was observed upon ESI-MS analysis of purified deGFP Y35X (a value that coincides well with the calculated mass of 26,670 Da for our mutant protein), confirming that no background suppression had occurred in the presence of the UAA. These results imply that the presence of (**2**) in the reaction led to no detectable background suppression by natural amino acids instead of (**2**) (detailed ESI-MS results, Fig. S5). The "click" reaction results (Fig. 4, panel D) clearly show a fluorescent band corresponding to ca. 27 kDa, a value in agreement with the calculated molecular mass of deGFP of 26.6 kDa, including an added Tamra-azide group which elevates its mass to ca. 27 kDa. The possibility of detecting GFP fluorescence (excitation: 485 nm; emission: 525 nm) was excluded by a control experiment where Y35 (**2**) deGFP, that did not undergo a "click" reaction, was checked for Tamra fluorescence. No fluorescence was observed under the same conditions used for the imaging of the gel shown in Figure 4, panel D. Lastly, when protein expression was attempted in the absence of (**2**) followed by an attempted "click" reaction using Tamra-azide, Tamra fluorescence was not observed (data not shown). The combination of these results brought us to the conclusion that our system is suitable for downstream applications and that (**2**) was the only incorporated amino acid in response to the UAG stop codon.

Suppression Efficiency of the System and Long-Term Stability

To confirm the usefulness of the special advantages conferred by the GRO strain, we also tested several other strains and variations of the GRO strain as listed in Table S1 (a complete description of all plasmids and strains used in this study is provided in Table S2 in the SI section). These other strains vary from the GRO strain in terms of genotype and induction conditions, as well as RF1 effects. To examine the compatibility of other OTS vectors with our methodology, we also tested several other OTSs (Table S1).

One of our preliminary assumptions was that by introducing the OTS into $RF1^+$ *E. coli* strains, we would not successfully expand the genetic code by amber suppression using endogenously introduced components. This assumption was based on the premise that the concentration of o-tRNA that accumulated in the *E. coli* cell prior to lysis would be insufficient for competition with RF1, thus resulting in low expression due to limited suppression efficiency. In addition, we have shown in a previous report (Smolskaya et al., 2013) that suppression efficiency in the presence of RF1 is mostly dependent on suppressor tRNA concentrations; we experimented to test this assumption. Additionally, in the absence of RF1, we inquired whether the concentrations of endogenous o-tRNA introduced would suffice to prevent ribosome stalling in the GRO strain. These two points could be addressed by calculating the "suppression efficiency" (Wang et al., 2001) of each OTS in each strain. This term is commonly calculated using equation 1:

$$\%\text{Suppression efficiency} = [(\text{UAGprotein})/(\text{WTprotein})] \times 100 \quad \text{(Eq.1)}$$

As listed in Table S 1, we produced the various cell-free extracts and compared suppression efficiencies. We compared the pEVOL Pyl-OTS in three different *E. coli* strains, DH5α, BL21 (DE3), and C321.Δ*prf*A. Such comparison between $RF1^+$ strains (DH5α and BL21 (DE3)) and the $RF1^-$ strain (C321.Δ*prf*A) enabled us to assess the advantages of our system and the effects of RF1 on the behavior of the system (Fig. 5, panel A). Furthermore, we wished to identify the crucial components that enabled the efficient functioning of the system. To this end, we produced crude extracts made using different growth and expression conditions (Table S1). First, to identify the effect of the PylRS concentration on the system, we prepared an extract from non-induced GRO cells transformed with the Pyl-OTS. Second, given the failed attempt to create two other functional C321.Δ*prf*A OTS strains, one with pSUP-pACF OTS (encoding for p-acetyl-L-phenylalanine) and the other with pKD-Sep OTS (encodes for phospho-serine), we tested the compatibility of the methodology with other OTS plasmids (Fig. 5, panel A).

From the results shown in Figure 5, panel A, it can be seen that as hypothesized, the GRO strain expressing Pyl-OTS exhibits near-perfect suppression efficiency (100%), while the same OTS expressed in other strains shows significantly lower suppression efficiencies. These results were expected because of the special

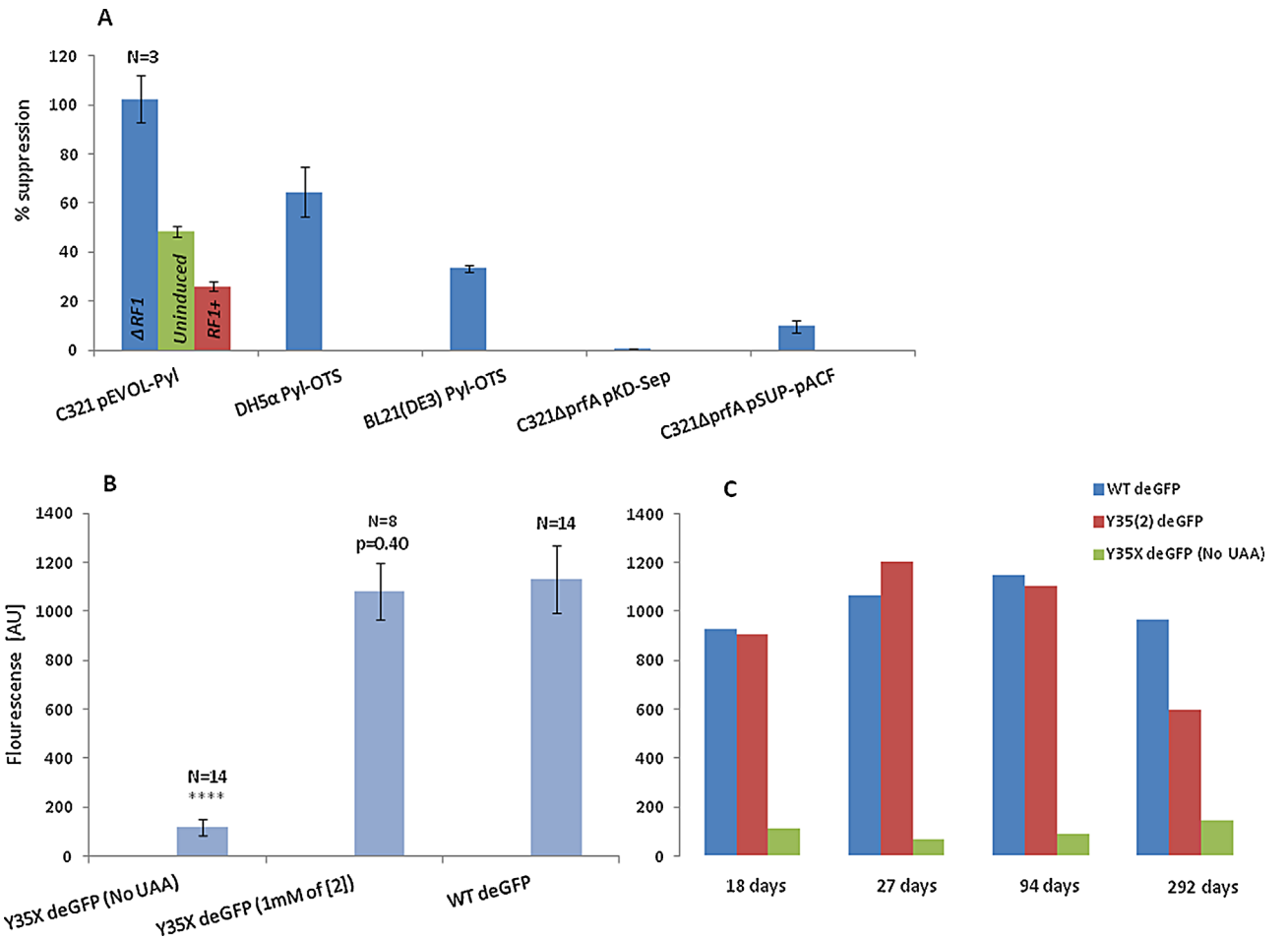

Figure 5. Suppression efficiency, reproducibility and long term stability of the system. **A)** Suppression efficiencies of different *E. coli* strains calculated from deGFP fluorescence. In the ΔRF1 strain, the number (N) of different reactions carried out to obtain suppression efficiency was $N = 3$, while in all other strains $N = 2$. A type 3 nested ANOVA test comparing suppression efficiencies was performed. **B)** The stability and reproducibility of the system was calculated as the standard deviation of N independent reactions, collected on different dates and using different aliquots of extract, reaction buffer and deGFP expression plasmid. A type 3 nested ANOVA test comparing WT deGFP expression and that of Y35X deGFP and Y35X deGFP with no UAA added was calculated. $P < 0.0001$ is indicated as ****. (**C**) Yields of genetically expanded CFPS reactions using a single batch of both lysate and buffer monitored over the course of time.

attributes of the GRO strain (i.e., the complete reassignment of the amber codon). The hypothesis is further reinforced when comparing the low $RF1^{+}$ and significantly higher $RF1^{-}$ suppression efficiencies of the recoded C321 strains and when comparing between the C321 strain to the DH5α and BL21 strains which contain RF1. Thus, we are confident that the absence of RF1 is crucial for the unprecedented suppression efficiency of the system. The reduced efficiency of the non-induced GRO Pyl-OTS strain could be explained by the lower cellular concentrations of the pyrrolysyl-tRNA synthetase that sequentially reduced the rate of o-tRNA aminoacylation with the UAA and ultimately, the overall yield of the genetically expanded CFPS.

One key challenge in working with genetically expanded CFPS systems is the low reproducibility of the system due to the use of exogenous components. From our own experience, this is a major challenge. It is simply hard to reproduce the quality and exact quantities of exogenous OTS components (o-tRNA and o-aaRS) added to the CFPS reaction. Moreover, according to our experience, the storage time of active OTS components is very short (i.e., several weeks), translating into a frequent need for refreshing OTS component stocks (Table S2). To assess the long-term stability and reproducibility of our system, we conducted a series of CFPS experiments on different occasions over a period of several months. These reactions involved three different batches of plasmids, several ("N"; denotes the exact number in Fig. 5 panel B) different aliquots of the endogenous and genetically expanded extracts containing the Pyl-OTS and different aliquots of reaction enhancement buffer. The results of these studies are presented in Figure 5, panel B. Finally, to test the long storage effects on the genetically expanded CFPS, reactions were carried out at different time points and up to 10 months after the preparation of the lysate and buffer (Fig. 5, panel C). After ~10 months the system still had stable yields (ca. 85% of average yields) but suppression efficiency was reduced to ca. 60%. Leading us to conclude that the genetically expanded lysate is stable for at least 3 month, after which the decay in suppression efficiency slowly becomes significant.

From the results, we conclude that the introduced endogenous and genetically expanded CFPS system significantly reduces the previously described challenges that arise from using exogenous OTS components. The CFPS of wild type (WT) deGFP was relatively stable, with a standard deviation of 12% over 14 experiments. The CFPS of Y35X deGFP with 1 mM of (**2**) (i.e., genetically expanded CFPS) was equally stable, with a standard deviation of 11% over eight different reactions. The *P*-value calculated shows no significant difference between the values of the CFPS of WT deGFP and of the genetically expanded CFPS of deGFP. The negative control reactions (i.e., Y35X deGFP with no UAA added) showed consistently negligible expression levels. These results both prove the stability and reproducibility of the system and also strengthen the claim that the system achieved 100% suppression efficiency. Based on the latter results, we demonstrate that the endogenous and genetically expanded CFPS is nearly as stable as the CFPS system that did not include genetic code expansion components (Shin and Noireaux, 2012). From our results and to the best of our knowledge, the stability achieved by our system is significantly improved over that relying on exogenously added OTS components. Moreover, we are certain that this stability could be further improved by future optimization of the growth conditions and the cell extract preparation. From integrating the results of this research we were able to demonstrate that some of the limitations of existing genetically expanded CFPS methodologies were overcome. For example, when background suppression appears due to the incorporation of natural amino acids in the in vivo system, it can be overcome by increasing the concentration of the UAA in a controllable manner, hence eliminate the competition with endogenous levels of natural amino acids. In addition, close to 100% suppression efficiency allows for multiple UAAs incorporation into the same protein. Among other uses and advantages (See table S3 in the SI section), we are confident that the presented system will enable new Pyl-OTS UAAs derivatives that are not compatible with in vivo methodologies to be utilized (such as toxic UAAs). That, in our opinion, is the greatest utility and advantage of our system over in vivo methodologies.

Expression Yields of the System

In this study, the question of protein synthesis yields was secondary to the proof of concept, stability, genetic expansion capability, high fidelity and generality of the system. Still, we found it important to understand whether we achieved yields that would enable future optimization and enhancement. Accordingly, a calibration curve was plotted using pure EGFP fluorescence (Fig. S6) to calculate an average yield of WT deGFP of 0.218 ± 0.03 mg/mL. The average yield of (**2**) incorporated into Y35X deGFP was 0.209 ± 0.02 mg/mL. These results are in a good agreement with a Bradford assay conducted for direct quantification of purified deGFP, calculated to be 0.22 mg/mL (Fig. S7). These yields, which are low compared to those reported for the parent CFPS system [i.e., the system used by Sun et al. (2013)], with yields of ca. 0.75 mg/mL (Shin and Noireaux, 2010; Sun et al., 2013), could be explained by the fact that our system is a prototype and as such, it has yet to undergo optimization, mostly by the improvement of the cell extract preparation e.g., removal of the residual OTS plasmid and identification of stronger promoters. Furthermore, it is important to note that the absolute yield of the presented system is low compared to in vivo methodologies, as a result, we recommend the utilization of the presented system when one or more of the presented CFPS advantages are required (for a comprehensive comparison between CFPS and in vivo methodologies see Table S3 in the SI section).

Conclusions

Although our system has yet to undergo optimization, we feel that it is important to compare the advantages and disadvantages of this new system to existing systems and methodologies so to emphasize our motivation in developing this new approach and, more importantly, its current limitations and benefits. We thus compared our endogenous and genetically expanded CFPS system, an exogenous genetically expanded CFPS system and the earlier in vivo genetic expansion methodologies in Table S3: The major advantages of our system are; (i) rapid conversion from genes into genetically expanded proteins and sequential downstream application (overnight compared to 3–5 days), (ii) the potential support of high throughput screening approaches, and (iii) alleviation of the need for synthesis and purification of the OTS, thereby significantly increasing the reproducibility, storage time and accessibility of the system, as compared to exogenous CFPS systems. This is particularly meaningful to large-scale CFPS in industry.

Looking forward, optimization, increased yield and scale-up are the main challenges facing the system. Even at present, various possibilities for future system utilization exist, such as easy and rapid screening of substrates for PylRS and its derivatives, without the constraints presented by the membrane and requirement for living cells required by in vivo methodologies.

In summary, we report the development of a simple CFPS methodology that enables genetic code expansion through amber suppression utilizing the useful, albeit insoluble *Mm*PylRS/*Mm*tRNA$_{CUA}^{Pyl}$ (Pyl-OTS), for the first time in vitro. Using the *E. coli* GRO strain, the system achieved an unprecedented stable 100% amber suppression efficiency and produced stable yields of >0.2 mg/mL of genetically expanded deGFP with a long term stability of at least 3 months, while utilizing only endogenous OTS components.

This research was supported by the European Research Council (ERC) grant agreement no. 260647 (L.A). We gratefully acknowledge the help of Ms. Rachel Gregor in mass spectrometry analysis and in narrating the video highlights of this study. We thank Prof. Edward Lemke for providing the pEVOL Pyl-OTS plasmid and Prof. Peter G. Schultz for providing the pSUP pACF plasmid. We acknowledge the help with synthesis services by Dr. Yohai Dayagi from BGU.

References

Ahn J-H, Kim J-H, Reuel NF, Barone PW, Boghossian A a, Zhang J, Yoon H, Chang AC, Hilmer AJ, Strano MS. 2011. Label-free, single protein detection on a near-infrared fluorescent single-walled carbon nanotube/protein microarray fabricated by cell-free synthesis. Nano Lett 11:2743–2752.

Albayrak C, Swartz JR. 2013. Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation. Nucleic Acids Res 41:5949–5963.

Budisa N. 2004. Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. Angew Chem Int Ed Engl 43:6426–6463.

Carlson ED, Gan R, Hodgman CE, Jewett MC. 2012. Cell-free protein synthesis: Applications come of age. Biotechnol Adv 30:1185–1194.
Caschera F, Noireaux V. 2014. Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. Biochimie 99: 162–168.
Catherine C, Lee K-H, Oh S-J, Kim D-M. 2013. Cell-free platforms for flexible expression and screening of enzymes. Biotechnol Adv 31:797–803.
Daube SS, Bar-Ziv RH. 2013. Protein nanomachines assembly modes: Cell-free expression and biochip perspectives. Wiley Interdiscip Rev Nanomed Nanobiotechnol 5:613–628.
Dumas A, Lercher L, Spicer CD, Davis BG. 2014. Designing logical codon reassignment – Expanding the chemistry in biology. Chem Sci 6(1):50–69.
Fujii S, Matsuura T, Sunami T, Kazuta Y, Yomo T. 2013. In vitro evolution of α-hemolysin using a liposome display. Proc Natl Acad Sci USA 110:1–5.
Goerke AR, Swartz JR. 2009. High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. Biotechnol Bioeng 102: 400–416.
Hayes CS, Sauer RT. 2003. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. Mol Cell 12:903–911.
Heyman Y, Buxboim A, Wolf SG, Daube SS, Bar-Ziv RH. 2012. Cell-free protein synthesis and assembly on a biochip. Nat Nanotechnol 7:374–378.
Hong SH, Ntai I, Haimovich AD, Kelleher NL, Isaacs FJ, Jewett MC. 2014. Cell-free protein synthesis from a release factor 1 deficient escherichia coli activates efficient and multiple site-specific nonstandard amino acid incorporation. ACS Synth Biol 3:398–409.
Jun SY, Kang SH, Lee K-H. 2008. Continuous-exchange cell-free protein synthesis using PCR-generated DNA and an RNase E-deficient extract. Biotechniques 44:387–391.
Kai L, Dötsch V, Kaldenhoff R, Bernhard F. 2013. Artificial environments for the co-translational stabilization of cell-free expressed proteins. PLoS ONE 8:e56637.
Kanter G, Yang J, Voloshin A, Levy S, Swartz JR, Levy R. 2007. Cell-free production of scFv fusion proteins: An efficient approach for personalized lymphoma vaccines. Blood 109:3393–3399.
Katzen F, Chang G, Kudlicki W. 2005. The past, present and future of cell-free protein synthesis. Trends Biotechnol 23:150–156.
Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S. 2004. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. J Struct Funct Genomics 5:63–68.
Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S. 1999. Cell-free production and stable-isotope labeling of milligram quantities of proteins. FEBS Lett 442:15–19.
Kisselev LL, Buckingham RH. 2000. Translational termination comes of age. Trends Biochem Sci 25:561–566. http://www.ncbi.nlm.nih.go
Lajoie M, Rovner A, Goodman D, Aerni H, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg JA, Rohland N, Schultz PG, Jacobson JM, Rinehart J, Church GM, Isaacs FJ. 2013. Genomically recoded organisms expand biological functions. Science 342:357–360.
Li X, Fekner T, Ottesen JJ, Chan MK. 2009. A pyrrolysine analogue for site-specific protein ubiquitination. Angew Chem Int Ed Engl 48:9184–9187.
Liu CC, Qi L, Yanofsky C, Arkin AP. 2011. Regulation of transcription by unnatural amino acids. Nat Biotechnol 29:164–168.
Liu D, V, Zawada JF, Swartz JR. 2005. Streamlining Escherichia coli S30 extract preparation for economical cell-free protein synthesis. Biotechnol Prog 21: 460–465.
Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. 2009. 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.
Niwa T, Kanamori T, Ueda T, Taguchi H. 2012. Global analysis of chaperone effects using a reconstituted cell-free translation system. Proc Natl Acad Sci USA 109:8937–8942.
O'Donoghue P, Prat L, Heinemann IU, Ling J, Odoi K, Liu WR, Söll D. 2012. Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNAPyl for genetic code expansion. FEBS Lett 586:3931–3937.
Renesto P, Raoult D. 2003. From genes to proteins: In vitro expression of rickettsial proteins. Ann NY Acad Sci 990:642–652.
Rosenblum G, Cooperman BS. 2014. Engine out of the chassis: Cell-free protein synthesis and its uses. FEBS Lett 588:261–268.
Sasaki Y, Asayama W, Niwa T, Sawada S, Ueda T, Taguchi H, Akiyoshi K. 2011. Amphiphilic polysaccharide nanogels as artificial chaperones in cell-free protein synthesis. Macromol Biosci 11:814–820.
Shin J, Noireaux V. 2010. Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. J Biol Eng 4:1–9.
Shin J, Noireaux V. 2012. An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. ACS Synth Biol 1:29–41.
Smith MT, Hawes AK, Shrestha P, Rainsdon JM, Wu JC, Bundy BC. 2014. Alternative fermentation conditions for improved Escherichia coli-based cell-free protein synthesis for proteins requiring supplemental components for proper synthesis. Process Biochem 49:217–222.
Smolskaya S, Zhang ZJ, Alfonta L. 2013. Enhanced yield of recombinant proteins with site-specifically incorporated unnatural amino acids using a cell-free expression system. PLoS ONE 8:e68363.
Spirin a S, Baranov VI, Ryabova a L, Ovodov SY, Alakhov YB. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242:1162–1164.
Sun ZZ, Hayes a C, Shin J, Caschera F, Murray RM, Noireaux V. 2013. Protocols for implementing an escherichia coli based TX-TL cell-free expression system for synthetic biology. J Vis Exp 79:1–15.
Sunohara T, Jojima K, Tagami H, Inada T, Aiba H. 2004. Ribosome stalling during translation elongation induces cleavage of mRNA being translated in Escherichia coli. J Biol Chem 279:15368–15375.
Swartz J. 2006. Developing cell-free biology for industrial applications. J Ind Microbiol Biotechnol 33:476–485.
Takasu A, Kondo S, Ito A, Furukawa Y, Higuchi M, Kinoshita T, Kwon I. 2011. Artificial extracellular matrix proteins containing phenylalanine analogues biosynthesized in bacteria using T7 expression system and the PEGylation. Biomacromolecules 12:3444–3452.
Tan C, Saurabh S, Bruchez MP, Schwartz R, Leduc P. 2013. Molecular crowding shapes gene expression in synthetic cellular nanosystems. Nat Nanotechnol 8:602–608.
Tawfik DS, Griffiths AD. 1998. Man-made cell-like compartments for molecular evolution. Nat Biotechnol 16:652–656.
Tuley A, Wang Y-S, Fang X, Kurra Y, Rezenom YH, Liu WR. 2014. The genetic incorporation of thirteen novel non-canonical amino acids. Chem Commun (Camb) 50:2673–2675.
Venancio-Marques A, Liu Y-J, Diguet A, di Maio T, Gautier A, Baigl D. 2012. Modification-free photocontrol of β-lactam conversion with spatiotemporal resolution. ACS Synth Biol 1:526–531.
Wan W, Tharp JM, Liu WR. 2014. Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool. Biochim Biophys Acta 1844:1059–1070.
Wang L, Brock a, Herberich B, Schultz PG. 2001. Expanding the genetic code of Escherichia coli. Science 292:498–500.
Xu Z, Chen H, Yin X, Xu N, Cen P. 2005. High-level expression of soluble human beta-defensin-2 fused with green fluorescent protein in Escherichia coli cell-free system. Appl Biochem Biotechno 127:53–62.
Yabuki T, Motoda Y, Hanada K, Nunokawa E, Saito M, Seki E, Inoue M, Kigawa T, Yokoyama S. 2007. A robust two-step PCR method of template DNA production for high-throughput cell-free protein synthesis. J Struct Funct Genomics 8: 173–191.
Young TS, Ahmad I, Yin JA, Schultz PG. 2010. An enhanced system for unnatural amino acid mutagenesis in E. coli. J Mol Biol 395:361–374.
Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ. 2011. Microscale to manufacturing scale-up of cell-free cytokine production?a new approach for shortening protein production development timelines. Biotechnol Bioeng 108:1570–1578.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.