Published online 15 April 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 11* ***5949–5963***
doi:10.1093/nar/gkt226

Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation

Cem Albayrak[1] and James R. Swartz[1,2,*]

[1]Department of Chemical Engineering, Stanford University, 381 North-South Mall, Stanford, CA 94305, USA and [2]Department of Bioengineering, Stanford University, 318 Campus Drive, Stanford, CA 94305, USA

Received June 26, 2012; Revised March 8, 2013; Accepted March 12, 2013

ABSTRACT

We describe a new cell-free protein synthesis (CFPS) method for site-specific incorporation of non-natural amino acids (nnAAs) into proteins in which the orthogonal tRNA (o-tRNA) and the modified protein (i.e. the protein containing the nnAA) are produced simultaneously. Using this method, 0.9–1.7 mg/ml of modified soluble superfolder green fluorescent protein (sfGFP) containing either *p*-azido-L-phenylalanine (pAzF) or *p*-propargyloxy-L-phenylalanine (pPaF) accumulated in the CFPS solutions; these yields correspond to 50–88% suppression efficiency. The o-tRNA can be transcribed either from a linearized plasmid or from a crude PCR product. Comparison of two different o-tRNAs suggests that the new platform is not limited by Ef-Tu recognition of the acylated o-tRNA at sufficiently high o-tRNA template concentrations. Analysis of nnAA incorporation across 12 different sites in sfGFP suggests that modified protein yields and suppression efficiencies (i.e. the position effect) do not correlate with any of the reported trends. Sites that were ineffectively suppressed with the original o-tRNA were better suppressed with an optimized o-tRNA (o-tRNA^{opt}) that was evolved to be better recognized by Ef-Tu. This new platform can also be used to screen scissile ribozymes for improved catalysis.

INTRODUCTION

Despite the vast functional and structural variety in proteins, the 20 natural amino acids that comprise these proteins are limited in their chemical diversity. Most of the natural amino acids possess inert site chains, and the ones with reactive side chains typically appear multiple times in a given protein. As a result, precise post-translational modification of proteins is usually impractical with the 20 natural amino acids.

This problem is solved by the site-specific incorporation of non-natural amino acids (nnAAs), which contain a diverse array of reactive side chains. Initially, these nnAAs were chemically coupled to the amber suppressor tRNA (i.e. a tRNA that recognizes the UAG amber stop codon), and >70 nnAAs were site-specifically introduced into proteins using this method (1). Partly due to the fact that stoichiometric amounts of acylated tRNA were needed, the modified protein yields were low (~ 50 μg/ml) (2). Moreover, the nnAAs could only be incorporated *in vitro*, as acylated tRNAs could not be introduced into cells. Biological acylation, in contrast, mimics the incorporation of natural amino acids, and involves dedicated orthogonal tRNAs (o-tRNAs) and synthetases that are specific to each nnAA and act catalytically. The nnAA is covalently coupled to the 3′ terminus of the o-tRNA by the synthetase specific to the nnAA (nnAARS). The acylated o-tRNA then forms a ternary complex with GTP and the *Escherichia coli* elongation factor Ef-Tu, enters the aminoacyl (A) site on the ribosome, outcompetes the release factor 1 (RF1), which would otherwise terminate translation, and adds the nnAA to the nascent polypeptide chain at the amber stop codon.

The first orthogonal components for site-specific incorporation of nnAAs in *E. coli* were derived from the tyrosyl tRNA (*Mj*tRNA^{Tyr}) and the tyrosyl-tRNA synthetase (*Mj*TyrRS) of the archaeon *Methanocaldococcus jannaschii*. There was already some orthogonality between the tyrosyl pairs from *M. jannaschii* and from *E. coli* due to the difference in the modes of tRNA^{Tyr} recognition between the *M. jannaschii* TyrRS and the *E. coli* TyrRS (3,4). First, the anticodon of the *Mj*tRNA^{Tyr} was changed from GUA to CUA to suppress the amber (UAG) stop codon. Then, the amber suppressor *Mj*tRNA^{Tyr} was mutated to be fully orthogonal to the endogenous aminoacyl-tRNA synthetases while retaining

*To whom correspondence should be addressed. Tel: +1 650 723 5398; Fax: +1 650 725 0555; Email: jswartz@stanford.edu
Present address:
Cem Albayrak, Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland.

the ability to be aminoacylated by the *Mj*TyrRS (5). Finally, the amino acid-binding site of the *Mj*TyrRS was mutated to recognize a nnAA so that the resultant enzyme would acylate the evolved o-tRNA exclusively with the particular nnAA. The absence of cross-reactivity between the orthogonal components and the endogenous ones was crucial for maintaining translational fidelity. The Schultz laboratory used this methodology to site-specifically incorporate >30 nnAAs (6) encompassing a vast range of chemical diversity. Advancements in o-tRNA design (7) and the expression of orthogonal components (8,9) led to accumulation of >100 mg/l modified protein *in vivo*. These phenylalanine / tyrosine analogues contain, among others, acetyl (10,11), halide (12,13), benzoyl (14), methyl (15), amino (16), azido (17) and alkyne (18) moieties. The latter two, namely *p*-azido-L-phenylalanine (**1**) and *p*-propargyloxy-L-phenylalanine (**2**, Figure 1), are particularly useful for bioconjugation, as proteins containing these nnAAs can be directly coupled using the copper (I)-catalysed azide–alkyne cycloaddition reaction. More recently, orthogonal pairs derived from the pyrrolysyl tRNA-synthetase pair of the archaea *Methanosarcina barkeri* and *Methanosarcina mazei* were used to incorporate a wide variety of pyrrolysine and phenylalanine analogues during *in vivo* expression in *E. coli*; some of these analogues also contain azide or alkyne moieties (19–24). In one study, a single pyrrolysyl tRNA-synthetase mutant was used to incorporate as many as 12 different phenylalanine analogues (25).

Cell-free protein synthesis (CFPS) with bacterial extracts has been successfully used in this laboratory to produce a large variety of active proteins. These proteins include, but are not limited to, metalloproteins such as ferredoxin (26) and [FeFe] hydrogenase (27), virus-like particles (28), membrane proteins (29,30), transcription factors (31) and proteins containing nnAAs (32–35). Recently, the cell-free accumulation of 700 mg/l of human granulocyte-macrophage colony-stimulating factor (hGM-CSF) was demonstrated at 100-litre scale (36). Amino acid stabilization (37,38), a natural chemical environment (39) and the activation of central metabolism (40) were critical in enabling the inexpensive and efficient cell-free synthesis of these proteins.

The open nature and the direct accessibility of the cell-free platform were particularly important for the synthesis of proteins containing nnAAs. Initial experiments were focused on extract strain development and the optimization of orthogonal component supply. The purified orthogonal synthetase and the nnAA, along with standard reagents, were added to an orthogonal extract in which the o-tRNA was constitutively expressed. Using this platform, proteins containing pAzF or pPaF were synthesized at high concentrations (150–930 μg/ml) and suppression efficiencies (25–96%; suppression efficiency is defined as the ratio of the modified protein yield to the natural protein yield) (32,33). Unfortunately, however, the yields and suppression efficiencies varied significantly from protein to protein, extract to extract and position to position (that is, position in the protein at which the nnAA is incorporated). The dependence of modified protein yields on the site of nnAA incorporation is called the position effect; it is also referred to as the codon or mRNA context effect in the literature (41–44). We have recently measured the o-tRNA concentrations, demonstrated their limitation in this platform and developed a new cell-free method in which the o-tRNA is synthesized *in vitro*, purified and added to the more productive standard extract along with the purified synthetase, nnAA and the standard reagents (45). Despite the improvements in modified protein titers and suppression efficiencies, we needed a more practical and inexpensive method for supplying the o-tRNA, as *in vitro* transcription and purification were both laborious and expensive.

In this article, we describe a modular, scalable and efficient cell-free platform in which the o-tRNA is synthesized simultaneously with the modified protein (Figure 2). In this new cell-free system, T7 RNA polymerase (RNAP) catalyses the transcription of both the mRNA encoding the modified protein and a 'transzyme', which is a fusion of the hammerhead ribozyme (46) and the o-tRNA. The transzyme folds correctly and cleaves itself to liberate the active o-tRNA. The o-tRNA is then acylated with the nnAA by the specific synthetase, and the nnAA is added to the nascent polypeptide chain at the amber stop codon. Using this method, 0.9 to 1.7 mg/ml of super-folder green fluorescent protein (sfGFP) containing either pAzF or pPaF accumulated after 16 h; these yields correspond to 50–88% suppression efficiency. The modified sfGFP yields and suppression efficiencies were measured across 12 different nnAA incorporation sites and using two different o-tRNA sequences. We observed that the position effect could not be explained by any of the reported hypotheses. The variation in modified protein yields and suppression efficiencies across the 12 incorporation sites was reduced when the o-tRNA^{opt} [which was optimized for better recognition by Ef-Tu (9)] was produced *in situ*, instead of the original o-tRNA. Experiments with *Schistosoma* ribozymes showed that this platform can also be used for evolving more active scissile ribozymes, as modified sfGFP fluorescence is directly related to ribozyme catalysis. Finally, the o-tRNA can be produced *in situ* from either a linearized plasmid or a crude PCR product, thereby making the system more practical for large-scale applications. The new cell-free platform can be used to effectively incorporate any of the >30 nnAAs (even poorly soluble ones such as pPaF) site-specifically into any protein.

MATERIALS AND METHODS

Plasmid construction

The 'transzyme' template (46) consists of the T7 promoter, the hammerhead ribozyme sequence and the o-tRNA sequence. Two different sequences, which contain either the original J17 o-tRNA template (5) or the optimized o-tRNA (o-tRNA^{opt}) template (9), were constructed by PCR using overlapping oligonucleotides. The transzyme sequence encoding the original o-tRNA is: 5′-GCTTTTA GATCTTAATACGACTCACTATAGGGAGACCGGC TGATGAGTCCGTGAGGACGAAACGGTACCCGG TACCGTC**CCGGCGGTAGTTCAGCAGGGCAGAACG**

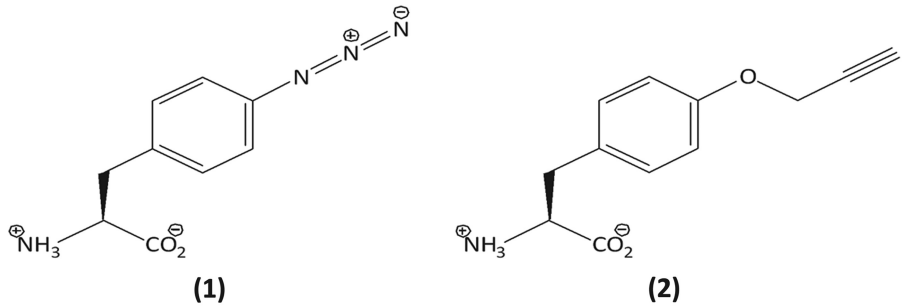

Figure 1. The nnAAs used in this study. **(1)** *p*-azido-L-phenylalanine (pAzF), and **(2)** *p*-propargyloxy-L-phenylalanine.

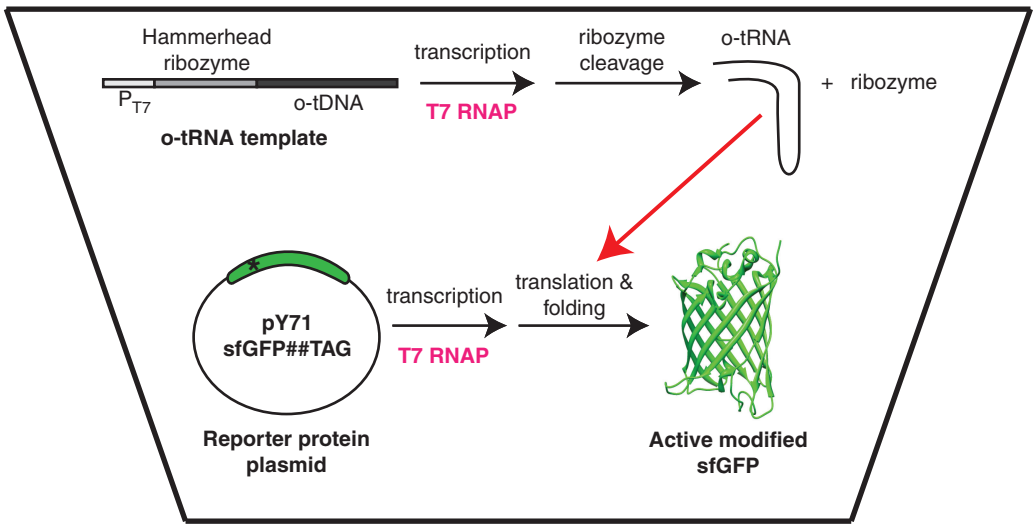

Figure 2. Schematic of the new cell-free platform for site-specific incorporation of nnAAs. The T7 RNA polymerase catalyses the transcription of both the transzyme RNA, which comprises the fusion of the hammerhead ribozyme and the o-tRNA, and the messenger RNA encoding the protein to be modified with the nnAA. Following transcription, the transzyme cleaves itself to release the biologically active o-tRNA needed for nnAA incorporation at the amber stop codon (TAG; depicted by an asterisk). The three-dimensional structure of sfGFP was used to render the modified protein (PDB accession number: 2B3P).

GCGGACTCTAAATCCGCATGGCGCTGGTTCAAAT CCGGCCCGCCGGACCAGGAAGCTTACATCCGTC GACAAAAGC-3′.

The transzyme sequence encoding the o-tRNA^{opt} ('T-Nap1' from reference 7) is: 5′-GCTTTTAGATCTTA ATACGACTCACTATAGGGAGACCGGCTGATGAG TCCGTGAGGACGAAACGGTACCCGGTACCGTC**C CGGCGGTAGTTCAGCAGGGCAGAACGGCGGACT CTAAATCCGCATGGCAGGGGTTCAAATCCCCTCC GCCGGACCA**GGAAGCTTACATCCGTCGACAAAA GC-3′. The o-tRNA sequences are given in bold. The transzyme sequence encoding the original o-tRNA was ligated into the pY71 vector (32) at the BglII and SalI restriction sites to create the pY71 HHotDNA. The pY71HHotDNA^{opt} was created by converting the o-tRNA template to the optimized o-tRNA template (o-tDNA^{opt}) using the QuikChange® protocol (Agilent Technologies, Santa Clara, CA).

The *Schistosoma* transzyme sequence encoding the *Schistosoma* ribozyme (47) 5′ to the o-tRNA is: 5′-GCT TTTAGATCTTAATACGACTCACTATAGGGAGATG AA**GGCA**CGCCGGCTGATGAGTCCCAAATAGGAC GAAATGCCAGAGAGGCATCCCGGCGGTAGTTC AGCAGGGCAGAACGGCGGACTCTAAATCCGCA TGGCGCTGGTTCAAATCCGGCCCGCCGGACCAG GGTCGACAAAAGC-3′. The AGAGA loop (underlined) was added to covalently link the Stem III to the gene encoding the cleaved RNA. Three different variants of the *Schistosoma* transzyme were created by mutating the Stem I 'bulge' nucleotides (written in bold). The first variant is given above; the second and third variants contain the AAAA and GTAC sequences, respectively. The secondary structure of the *Schistosoma* ribozymes was analysed using the Mfold web server (48). All three *Schistosoma* transzyme variants were assembled by PCR using overlapping oligonucleotides, and ligated into the pY71 at the BglII and SalI restriction sites to create the pY71fullHHotDNA plasmids.

Construction of the plasmid pY71sfGFP216TAG was described previously (32). This plasmid encodes a sfGFP with an amber stop codon at position 216 and a C-terminal *Strep*-tag. The additional glycine codon at position 215 and the amber stop codon were deleted to obtain pY71sfGFP, which contains the original sfGFP

sequence (49,50). Twelve amber stop codons were then individually introduced into the sfGFP sequence using the QuikChange® protocol (Agilent Technologies) to create 12 new coding sequences. A stop codon was introduced at the following positions to replace the codon naturally present at that position (the amino acid normally encoded is indicated): Asn23, Ile39, Lys52, Tyr66, Tyr74, Tyr92, Tyr106, Tyr143, Tyr151, Tyr182, Tyr200 and Tyr223. The DNA polymerase from *Pyrococcus furiosus* (51) used in the PCRs was produced and purified in our laboratory as described (52).

After the verification of the plasmid sequences by DNA sequencing, *E. coli* cultures harbouring the plasmids were grown for 16 to 18 h in Terrific Broth (Invitrogen, Carlsbad, CA) and purified using the Qiagen tip-500 columns (Qiagen, Valencia, CA) as per manufacturer instructions. DNA concentrations were calculated from absorbance measurements at 260 nm (1 A_{260} = 50 μg/ml double-stranded DNA); the absorbance measurements were taken using the HP 8452A diode array spectrophotometer (Agilent Technologies).

Preparation of orthogonal tRNA templates and the synthetases specific to nnAAs

The purified plasmids encoding the o-tRNA and o-tRNA^{opt} were restricted with the BstNI enzyme at 60°C for 2 to 3 h. The plasmid concentration and reaction volume did not exceed 320 μg/ml and 1 ml, respectively; digestion reactions at higher plasmid concentrations or larger volumes were incomplete. After every reaction, 5 μl samples were analysed by agarose gel electrophoresis to ensure complete linearization. The linearized plasmids were subsequently ethanol-precipitated, washed once with 70% v/v ethanol, resuspended in 10 mM Bis-Tris, pH 7.5 buffer overnight at 4°C, and stored at −20°C. Concentrations of the linearized plasmids were calculated from absorbance measurements at 260 nm as described previously.

A linear transzyme template encoding the T7 promoter, the hammerhead ribozyme and the o-tRNA^{opt} was also synthesized by PCR using overlapping oligonucleotides. The sequence for this template is: 5′-GCTTTTAGATCT TAATACGACTCACTATAGGGAGACCGGCTGATG AGTCCGTGAGGACGAAACGGTACCCGGTACCG TCCCGGCGGTAGTTCAGCAGGGCAGAACGGCG GACTCTAAATCCGCATGGCAGGGGTTCAAATCC CCTCCGCCGGACCA-3′. Differently from the previous transzyme template sequences, this sequence terminates with the CCA nucleotides that comprise the 3′ terminus of the o-tRNA^{opt}. This linear DNA was subsequently ethanol-precipitated without purification, washed with 70% ethanol, resuspended in 10 mM Bis-Tris buffer, pH 7.5 overnight, and stored at −20°C. The concentrated PCR product was quantified as explained above.

The synthetases specific to the nnAAs pPaF and pAzF, pPaFRS and pAzFRS, were produced and purified as described (32,33).

Cell extract preparation

The orthogonal S30 cell extract was prepared from KC6 cells (37) harbouring the pDule-tRNA plasmid, as described (33). The lysate was not diluted after homogenization. The specific growth rate of this KC6/o-tRNA culture was 0.57 h^{-1} during the batch fermentation. The standard KC6 extract was prepared as described (53). After homogenization, 0.2 ml S30 buffer (14 mM magnesium acetate, 60 mM potassium acetate, 10 mM Tris acetate, pH 8.2) was added per ml of lysate, and the diluted lysate was dialysed against 100 volumes of S30 buffer at 4°C four times for 30 min each. Finally, the S30 cell extract was centrifuged at 10 000*g* and 4°C for 20 min, and the supernatant was aliquoted, flash-frozen and stored at −80°C. The specific growth rate of the KC6 culture was 0.96 h^{-1} during the fed-batch fermentation.

Cell-free protein synthesis

The PANOx SP (PEP, Amino Acids, NAD, Oxalic Acid, Spermidine and Putrescine) (39) system was used in the cell-free reactions, with several changes in the reagents and reagent concentrations. All of the chemicals were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise stated.

Twenty-five-microlitre PANOx SP reaction solutions contained, unless otherwise noted: 10 mM ammonium glutamate, 20 mM magnesium glutamate, 175 mM potassium glutamate, 1.2 mM ATP, 0.86 mM each of CMP, GMP and UMP, 10 mM dibasic potassium phosphate, 34 μg/ml folinic acid, 171 μg/ml *E. coli* tRNAs (Roche Applied Science, Penzberg, Germany), 33 mM phosphoenolpyruvate (PEP) (Roche Applied Science), 1.5 mM spermidine, 1 mM putrescine, 0.33 mM nicotinamide adenine dinucleotide (NAD), 0.27 mM coenzyme A, 2.7 mM sodium oxalate, 2 mM each of the 20 amino acids, 5 μM L-[^{14}C(U)]-leucine (PerkinElmer, Waltham, MA), 2 mM *p*-azido-L-phenylalanine (pAzF, Chem-Impex International, Wood Dale, IL) or 4 mM *p*-propargyloxy-L-phenylalanine (pPaF, synthesized according to reference 54), 0.3 mg/ml pAzFRS or 0.5 mg/ml pPaFRS (prepared as described above), 100 μg/ml T7 RNA polymerase (prepared according to reference 55), 6 nM sfGFP plasmid, 0.2 mg/ml linearized o-tRNA or o-tRNA^{opt} plasmid template and either 0.24 volume of standard KC6 extract or 0.28 volume of orthogonal extract. The glutamate salts were diluted from a 10-fold concentrated (10×) solution. Similarly, a 10× NTMP Master Mix solution contained ATP, the three NMPs, folinic acid and the *E. coli* tRNAs, and was adjusted to pH 7.3 with dibasic potassium phosphate. The natural amino acids were diluted from a stock solution, which contained a 50 mM concentration of each amino acid. This stock solution was prepared by adding the amino acids in the following order (given in their one-letter code): R, V, W, F, I, L, C, M, A, N, D, E, G, Q, H, K, P, S, T, Y. During the preparation, it was ensured that each amino acid was dissolved before addition of the next, except tyrosine, which is added last and remains suspended in the solution. The reactions were incubated at 30°C for 16 h, unless stated otherwise. CFPS reactions containing pAzF were prepared in a dark room, and the tubes containing the reaction solutions were wrapped in aluminum foil to prevent photodissociation of the aromatic azide (56).

Quantification of natural and modified sfGFPs

Radiolabelled proteins were quantified as follows: The cell-free reactions were stopped by cooling the reaction tubes at −20°C for 4 min. A 4 µl sample from each reaction solution was spotted onto two of three Whatman MM filter papers (Whatman, Springfield Mill, United Kingdom); one of these papers was used to measure the total protein concentration and the other to measure total system radioactivity. The remaining cell-free solution was centrifuged at 20 800*g* and 4°C for 15 min. Four microlitres of the supernatant was spotted on the third Whatman MM filter paper to measure the soluble protein concentration, and the papers were dried overnight or for 1 h under an incandescent light bulb. Trichloroacetic acid precipitation was used to measure the total and soluble protein concentrations as described (57). The leucine content and the molecular weight of the protein were used to convert radioactivity measurements to protein concentrations. A set of 'blank' reactions (which do not contain the protein template) was included in each experiment; the same protocol was applied to these reactions to determine background radioactivity. These measurements were subtracted from the protein radioactivity measurements to obtain accurate protein concentrations. Full-length protein amounts in the soluble fractions were determined by densitometry analysis of the bands on SDS-PAGE gel autoradiograms.

Cell-free samples containing radiolabelled sfGFP were diluted in a sample loading buffer that contains lithium dodecyl sulfate (LDS) (Invitrogen) at pH 8.4 and 50 mM DTT. The solution was subsequently incubated at 95°C for 15 min to ensure complete denaturation of the sfGFP. The samples were loaded onto 10% Bis-Tris SDS-PAGE gels, and run using the MES SDS running buffer (Invitrogen). The gels were subsequently dried and exposed overnight to a storage phosphor screen (Molecular Dynamics, Sunnyvale, CA). Autoradiograms were obtained by imaging the screens on the Typhoon system (GE Healthcare, Uppsala, Sweden), and the band intensities were quantified using the ImageJ software (National Institutes of Health, Bethesda, MD).

The suppression efficiency is defined as the ratio of the yield of full-length modified protein to that of its natural counterpart. In each CFPS experiment, natural sfGFP was synthesized under the same cell-free reaction conditions; the full-length and soluble protein yields were measured via scintillation counting and densitometry.

Characterization of modified sfGFPs

To determine the specific activity of the modified sfGFPs in comparison with that of natural sfGFP, sfGFP23pPaF, sfGFP39pPaF and wild-type sfGFP were produced in CFPS with a C-terminal *Strep*-tag and purified via *Strep*-Tactin affinity chromatography (IBA GmbH, Goettingen, Germany) as per manufacturer instructions. The proteins were produced in 0.5 ml CFPS reactions containing all of the mentioned reagents except L-[^{14}C(U)]-leucine. Fractions obtained from *Strep*-Tactin affinity chromatography were analysed by SDS-PAGE, and those containing sfGFP were pooled and dialysed at 4°C first against 30 volumes of 10 mM potassium phosphate buffer, pH 7.5, then twice against 30 volumes of 20% w/v sucrose dissolved in the same buffer. Proteins were concentrated during the latter dialyses against sucrose-containing phosphate buffer.

Total protein concentrations were determined using the DC protein assay (Bio-Rad, Hercules, CA) with bovine serum albumin (BSA) as the standard for the assay. Full-length and soluble sfGFP concentrations were measured via densitometry analysis on the SDS-PAGE gels of purified proteins. Two-microlitre samples from each protein solution were diluted in 10 mM potassium phosphate buffer at pH 7.4; the fluorescence from these diluted protein solutions was measured between 523 and 548 nm using a Mithras LB940 multimode microplate reader (Berthold Technologies, Bad Wildbad, Germany) with F485 excitation and F535 emission filters. Specific activities were calculated by dividing fluorescence measurements by protein concentrations, measured at different dilutions of the purified proteins.

For mass spectrometry analysis, three additional modified proteins (sfGFP39pPaF, sfGFP143pPaF and sfGFP223pPaF) were produced and purified as explained above. Liquid chromatography/mass spectroscopy (LC/MS) analysis of these proteins was performed at the Stanford University Mass Spectrometry Laboratory. The proteins were precipitated with acetone, reconstituted in a buffer containing 8 M urea and 50 mM ammonium bicarbonate, reduced, alkylated and digested with trypsin overnight at a 1:100 protease-to-protein ratio. The resulting protein fragments were analysed by nano reversed-phase HPLC using an Eksigent 2D NanoLC (Eksigent, Dublin, CA). A fused silica column containing Duragel C18 (Peeke, Redwood City, CA) matrix was used. The initial buffer composition was 98% A (0.1% formic acid in water) / 2% B (0.1% formic acid in acetonitrile), and was increased to 40% B at a flow rate of 600 nl/min. The nanoHPLC was interfaced with a Bruker–Michrom Advance CaptiveSpray (Michrom Bioresources, Auburn, CA), spray source for nanoESI into the mass spectrometer. The LTQ Orbitrap Velos (Thermo Fisher, Waltham, MA) mass spectrometer was set in data-dependent acquisition mode to perform MS/MS on the top 12 most intense multiply charged cations. The .RAW data were searched using Sequest on a Sorcerer platform against the Uniprot database. The data were visualized and validated using Scaffold software (Proteome Software, Portland, OR). Quantification and MS/MS analysis of the identified peptides were conducted using XCalibur (Thermo Fisher) and GPMAW (Lighthouse Data, Odense, Denmark) software.

Measurement of o-tRNA concentrations produced in CFPS

These 25 µl CFPS reactions included 8 µM [^{3}H]-UTP (PerkinElmer) instead of L-[^{14}C(U)]-leucine. A new 10× Master Mix solution was made, in which the UMP was replaced with UTP so that the initial ratio of radioactive to non-radioactive UTP concentrations was known. The sfGFP yields from cell-free reactions containing each of

the two Master Mix solutions were comparable. The stock solution of [^{3}H]-UTP was concentrated by evaporation under vacuum at room temperature. The concentrations of tritium-labelled ribonucleic acids were measured using the same protocol as that for radiolabelled proteins. Concentrations of intact o-tRNA and o-tRNA^{opt} were determined by densitometry analysis of the soluble bands on the denaturing polyacrylamide gel autoradiogram.

Six-microlitre samples of cell-free solutions containing tritium-labelled RNA were first diluted in 2× TBE-Urea sample buffer (Invitrogen), then denatured at 95°C for 5 min prior to loading onto a 15% denaturing polyacrylamide gel (TBE-Urea gel, Invitrogen). The gel was run at 180 V for 80 min in 1× TBE buffer, pH 8.3 (Invitrogen), washed with diethyl pyrocarbonate (DEPC)-treated water, dried, and exposed for 4 days to a tritium storage phosphor screen (GE Healthcare). The tritium autoradiogram was generated using the Typhoon system, and the band intensities were quantified using the ImageJ software. For size comparison, the o-tRNA was also transcribed *in vitro* and purified by denaturing polyacrylamide gel electrophoresis as explained in reference 45. Two-microlitre samples of the purified uncleaved transzyme, o-tRNA and hammerhead ribozyme solutions were run on a 15% denaturing polyacrylamide gel as explained above.

RESULTS

In situ orthogonal tRNA production increases modified protein yields

In the standard cell-free platform for nnAA incorporation, the o-tRNA is produced in the extract source strain, and the nnAA and the purified nnAARS are added as separate reagents. We have previously shown that the o-tRNA is a limiting reagent in this cell-free platform; i.e. that the supplementation with additional o-tRNA increases modified protein yields (45). Since *in vitro* production and purification of the o-tRNA are both laborious and expensive, we sought a more practical method for supplying the cell-free reaction with adequate o-tRNA.

The hammerhead ribozyme was used to produce the o-tRNA *in vitro*, because the T7 RNA polymerase does not efficiently transcribe RNAs that start with pyrimidines such as the o-tRNA that requires a cytosine at its 5′ terminus (58). Effective synthesis of transcripts starting with pyrimidines was demonstrated using the 'transzyme' approach where the hammerhead ribozyme sequence is inserted between the T7 promoter and the gene encoding the desired transfer RNA (46). The plasmid containing these sequences is restricted using the BstNI enzyme that yields the correct CCA 3′ terminus required for tRNA activity. From this linearized template, the T7 RNA polymerase transcribes the ribozyme–tRNA fusion RNA which subsequently cleaves itself to release the tRNA. Placing the 5′ terminus of the tRNA sequence immediately after the ribozyme cleavage site ensures that the released tRNA has the correct 5′ and 3′ termini required for biological activity (46).

To develop a more practical and inexpensive method of o-tRNA supply, we investigated the *in situ* production of o-tRNA during the cell-free reaction, where both the o-tRNA and the modified protein are produced simultaneously (Figure 2). The linear template encoding the ribozyme–tRNA ('transzyme') fusion and the protein template are added to the cell-free reaction solution along with the nnAA, the nnAARS and the standard reagents for CFPS. During the incubation, T7 RNA polymerase catalyses the transcription of both the mRNA for protein translation and the transzyme (i.e. ribozyme–tRNA fusion). Once the transzyme is transcribed, the o-tRNA is liberated by autocleavage of the hammerhead ribozyme and acylated with the nnAA by the orthogonal synthetase. The acylated o-tRNA then enters the aminoacyl (A) site in the ribosome, suppresses the amber stop codon, and adds the nnAA to the nascent polypeptide chain.

To explore the feasibility of this approach and to identify the non-limiting concentrations of the o-tRNA template, we added different concentrations of the linearized o-tRNA template to the cell-free solution and measured the accumulation of full-length and soluble sfGFP with the nnAA pPaF incorporated at position 39 (sfGFP39pPaF). This new method greatly increased the modified protein yields from 260 μg/ml to ~1.2 mg/ml of full-length and soluble sfGFP39pPaF. The full-length protein yield was directly proportional to the o-tRNA template concentration up to 150 μg/ml; addition of o-tRNA template in excess of 150 μg/ml did not increase the modified protein yields. The suppression efficiency, which is defined as the ratio of the full-length and soluble modified protein yield to that of the wild-type protein, was increased from 15–70% (Figure 3A, diamonds).

Next, we wanted to see whether the increase in protein accumulation in the new cell-free platform was due to faster or more prolonged production of the modified protein. Therefore, we measured the accumulation of sfGFP23pPaF (i.e. full-length and modified sfGFP containing pPaF at position 23) over the course of 24 h in two different sets of reactions. In one set, the cell-free reaction solution contained 28% v/v of the orthogonal extract, in which the o-tRNA had been expressed constitutively during the growth of the source strain (Figure 3B, 'o-extract'). In the other set, the cell-free reactions contained, in addition to the standard reagents, 200 μg/ml linearized o-tRNA template to ensure that the o-tRNA template was not limiting (Figure 3B, 'transzyme'). A comparison of these two platforms showed that the modified protein was produced faster when it was synthesized simultaneously with the o-tRNA. In the standard CFPS reactions with the orthogonal extract, about half of the protein was produced within the first 2 h, and another 100 μg/ml modified protein accumulated after 14 h. At the end of 16 h, about 250 μg/ml full-length and soluble sfGFP23pPaF accumulated in the CFPS solution, which amounted to a suppression efficiency of ~15% (Figure 3B, 'o-extract'). In the 'transzyme' platform, there was a burst of protein

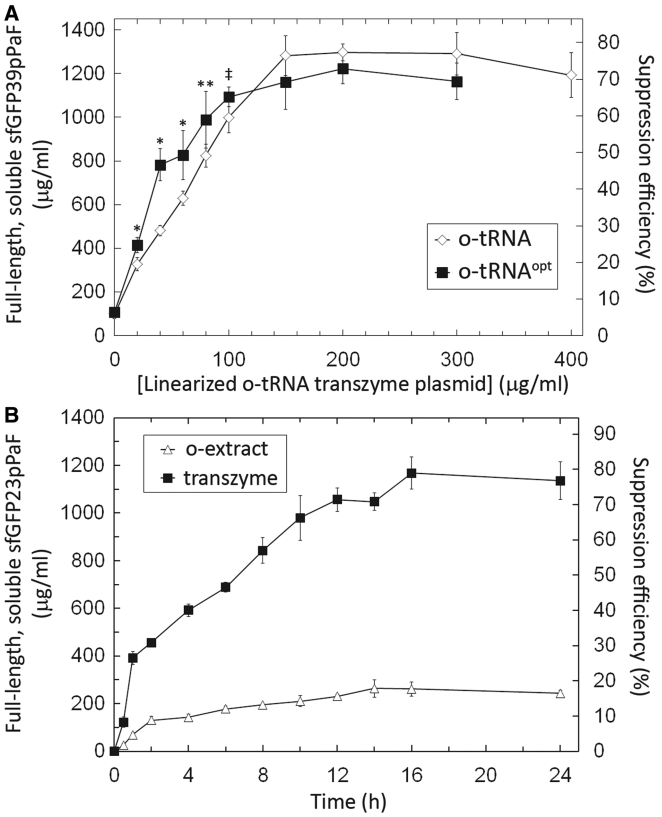

Figure 3. (**A**) Titration of linearized o-tRNA and o-tRNA^{opt} plasmid templates in the CFPS reactions containing pY71sfGFP39TAG. Error bars indicate ±1 standard deviation from six reactions. (*$P < 0.01$, **$P = 0.025$, ***$P = 0.023$) (**B**) Production of sfGFP containing pPaF at position 23 (sfGFP23pPaF) over 24 h. The 'o-extract' curve was generated from CFPS reactions in which the cell extract contained the o-tRNA. The 'transzyme' values were obtained from reactions, in which the o-tRNA and sfGFP23pPaF were produced simultaneously. A different set of reactions was incubated for each of the time points. Error bars indicate ±1 standard deviation for three reactions.

production within the first hour, fueled by the conversion of phosphoenolpyruvate (PEP) to ATP. This initial period was followed by slower protein production, which was probably fueled by oxidative phosphorylation (40). At the end of 16 h, about 1.2 mg/ml of modified sfGFP had accumulated, corresponding to a suppression efficiency of ~75% (Figure 3B, 'transzyme').

Modified protein production and Ef-Tu recognition of the acylated orthogonal tRNA

The Schultz laboratory recently reported a modified o-tRNA optimized for better recognition by the endogenous elongation factor Ef-Tu (9). This new o-tRNA sequence, dubbed the o-tRNA^{opt} sequence, was among the nine identified from their first tRNA library and enabled more effective site-specific incorporation of several nnAAs including pAzF (7). This step in the nnAA insertion pathway, namely the recognition of the acylated o-tRNA by the *E. coli* Ef-Tu, was a potential bottleneck that we had not addressed. We thus cloned this o-tRNA^{opt} template 3′ to the hammerhead ribozyme sequence, and

Table 1. Accumulation of o-tRNA and o-tRNA^{opt} at different tRNA template concentrations

Original or optimized o-tRNA	[Template]	[Accumulated o-tRNA]
Original	150 μg/ml (127 nM)	280 ± 40 μg/ml (12.6 ± 2.0 μM)
Original	200 μg/ml (169 nM)	380 ± 20 μg/ml (17.0 ± 1.0 μM)
Optimized	150 μg/ml (127 nM)	170 ± 60 μg/ml (7.4 ± 2.7 μM)
Optimized	200 μg/ml (169 nM)	200 ± 60 μg/ml (8.9 ± 2.7 μM)

Error bars indicate ±1 standard deviation for three reactions. The o-tRNA concentrations from the reactions containing less o-tRNA (or o-tRNA^{opt}) template could not be accurately quantified via densitometry because of weak tritium signals from the autoradiograms.

compared the performance of the o-tRNA^{opt} with that of the original o-tRNA in our new cell-free platform by titrating each of the two linearized transzyme templates (Figure 3A). At limiting template concentrations (i.e. below 150 μg/ml), Ef-Tu recognition was indeed a limiting factor: cell-free reactions containing the o-tRNA^{opt} template were more productive than the ones containing equal concentrations of the o-tRNA template. However, at template concentrations ≥150 μg/ml, there was no difference in productivity between the two different o-tRNAs; 1.1 to 1.2 mg/ml soluble and full-length sfGFP39pPaF accumulated in both sets of CFPS reactions.

Quantification of *in situ* orthogonal tRNA synthesis

We then quantified the o-tRNA (or o-tRNA^{opt}) synthesized during cell-free reactions incubated with 20, 40, 60, 80, 100, 150 or 200 μg/ml of the linearized o-tRNA template. Tritiated UTP (^{3}H-UTP) was added to the CFPS reaction solution to label the ribonucleic acids produced during the reaction, and the amount of synthesized o-tRNA was determined by scintillation counting and densitometry, similar to the quantification of modified protein labelled with ^{14}C-leucine (Supplementary Figure S1). UMP was replaced with UTP in these cell-free reactions to increase the accuracy of the synthesized RNA quantification.

As expected, the amount of accumulated o-tRNA and o-tRNA^{opt} was directly proportional to the template concentrations (Table 1). It was surprising to see that the only bands on the tritium autoradiogram were due to the mRNA and the o-tRNA; the uncleaved transzyme and the hammerhead ribozyme are apparently degraded by the ribonucleases in the CFPS solution. We suspect that the hammerhead ribozyme may not cleave itself efficiently in the CFPS reaction due to the difference in optimal temperature and Mg^{2+} concentrations between those for CFPS (20 mM Mg^{2+}, 30°C) and for ribozyme cleavage (30 mM Mg^{2+}, 60°C) (46). The need for high levels of accumulated o-tRNA and o-tRNA^{opt} is consistent with results from our previous study indicating low turnover for the nnAARS (45).

Non-natural amino acid incorporation and the position effect

The position effect (also called the codon context effect in the literature) refers to the dependence of nonsense suppression efficiency (nnAA incorporation, in this case) on the position of the nonsense codon (i.e. amber stop codon, UAG). The nucleotide following the amber stop codon on the mRNA has been suggested to be most influential on the suppression efficiency, with the stop codons followed by a purine being more efficiently suppressed than those followed by a pyrimidine (43,44,59,60). Previous research from our lab also suggested a trend between modified dihydrofolate reductase (DHFR) yields (which is directly related to nnAA incorporation at the UAG codon) and the proximity of the insertion site to the N-terminus of the DHFR protein (33). To further explore these influences, we examined pPaF incorporation in sfGFP at 12 different sites; the locations of these sites in the folded protein are shown in Figure 4A. We replaced, one at a time, all of the tyrosines, which is chemically the most similar natural amino acid to the nnAAs, and several other amino acids on the surface of the protein to minimize chemical or structural perturbations of the natural sfGFP fold. We then quantified the full-length and soluble modified sfGFP that accumulated in five different sets of cell-free reactions (Figure 4B). The first two sets contained all of the reagents except the o-tRNA ('no o-tRNA'; Figure 4B) or the nnAA ('no pPaF'; Figure 4B), respectively, and were conducted to assess non-specific suppression of the amber stop codon at different sites; in this case, the full-length protein contained a natural amino acid at the indicated position. The third set ('o-extract'; Figure 4B) was incubated using the standard platform and contained the orthogonal extract. The fourth ('transzyme'; Figure 4B) and fifth sets ('optimized'; Figure 4B) contained 200 μg/ml of the linearized plasmid templates coding for the o-tRNA and the o-tRNA^{opt}, respectively. Natural sfGFP was produced in a different set of reactions, and the modified proteins yields were divided by the natural protein yield to calculate the suppression efficiencies. Finally, the nucleotide following the amber stop codon at each position is indicated below the graph in Figure 4B.

In the standard platform, modified sfGFP yields were very similar across all of the 12 positions; the highest modified protein yield was obtained when the UAG codon was at position 143 ('o-extract'; Figure 4B). The non-specific read-through in the absence of the o-tRNA was also highest at this position. The greatest differences in protein yields across the various UAG codon positions were observed when the cell-free reactions contained the linearized o-tRNA template ('transzyme'; Figure 4B). Modified, full-length and soluble yields varied between 0.5 mg/ml for sfGFP74pPaF and 1.5 mg/ml for sfGFP223pPaF. Variation in yields and suppression efficiencies was reduced when the o-tRNA^{opt}, instead of the o-tRNA, was simultaneously expressed with the modified protein ('optimized', Figure 4B). Full-length and soluble sfGFP74pPaF yields were increased to 0.9 mg/ml, while the yields of modified sfGFP containing pPaF at more readily suppressed sites (such as 39, 151 or 223) remained around 1.4 mg/ml.

Regardless of the manner of o-tRNA supply or the o-tRNA sequence used, we did not observe any correlation between the modified protein yields (i.e. the suppression efficiency) and the residue position along the polypeptide backbone or between the protein yield and the nucleotide following the stop codon. The least efficiently suppressed stop codons (at positions 74 and 182) were followed by a cytosine, which agrees with the published correlation of $A = G > C > U$ for suppression efficiency (44). However, some of the other codons followed by a cytosine (at positions 52 and 200) were just as efficiently suppressed as others followed by purines. Neither did the presence of a uracil at the second position after the UAG codon or the proximity of the codon to the 5′ end lead to better suppression as suggested by earlier studies (33,44). There were no structural factors that contributed to the position effect either. Both residues 74 and 143 are in flexible coils, but position 74 was one of the least efficiently suppressed positions, while position 143 was one of the most efficiently suppressed. Of the most efficiently suppressed positions, residue 39 is in an alpha helix, residue 143 is in a flexible coil, while residues 106 and 151 are in beta strands.

Of the 12 sites, the least efficiently suppressed was at residue 74. We then tested if increasing the o-tRNA^{opt} supply could improve the soluble and full-length sfGFP74pPaF yields. We titrated the linearized o-tRNA^{opt} template and compared sfGFP74pPaF accumulation with the accumulation of sfGFP39pPaF, one of the most efficiently suppressed positions (Figure 4C). For both products, the modified protein yields followed similar trends, peaking with 200 μg/ml of linearized o-tRNA^{opt} template. Thus, using the o-tRNA^{opt} does not avoid the effects of different nnAA incorporation positions.

Substituting the natural amino acids with pPaF at these 12 sites also affected the fluorescence of the modified sfGFP. The incorporation of this nnAA at residue 66 abolished the fluorescence, which was not surprising since replacement of this tyrosine with other nnAAs was shown to alter the spectral properties of GFP (13). The substitution of pPaF for tyrosine at residue 92 reduced the fluorescence of the protein (emission between 523 and 548 nm) by 60–70%. Other substitutions did not alter the specific activity substantially (i.e. reduced by ≤20%). Two modified sfGFP proteins were purified via *Strep*-Tactin affinity chromatography, and their specific activities were measured and compared with that of natural sfGFP. The specific activities were 95 ± 1 AU/ng for sfGFP, 85 ± 2 AU/ng for sfGFP23pPaF and 86 ± 3 AU/ng for sfGFP39pPaF. Finally, specific activities and soluble yields of pAzF-containing sfGFP proteins were similar to those containing pPaF at the mentioned sites (data not shown).

Mass spectroscopy analysis of the modified protein

In our new platform, the o-tRNA synthesis starts at the same time as the modified protein synthesis; hence there is

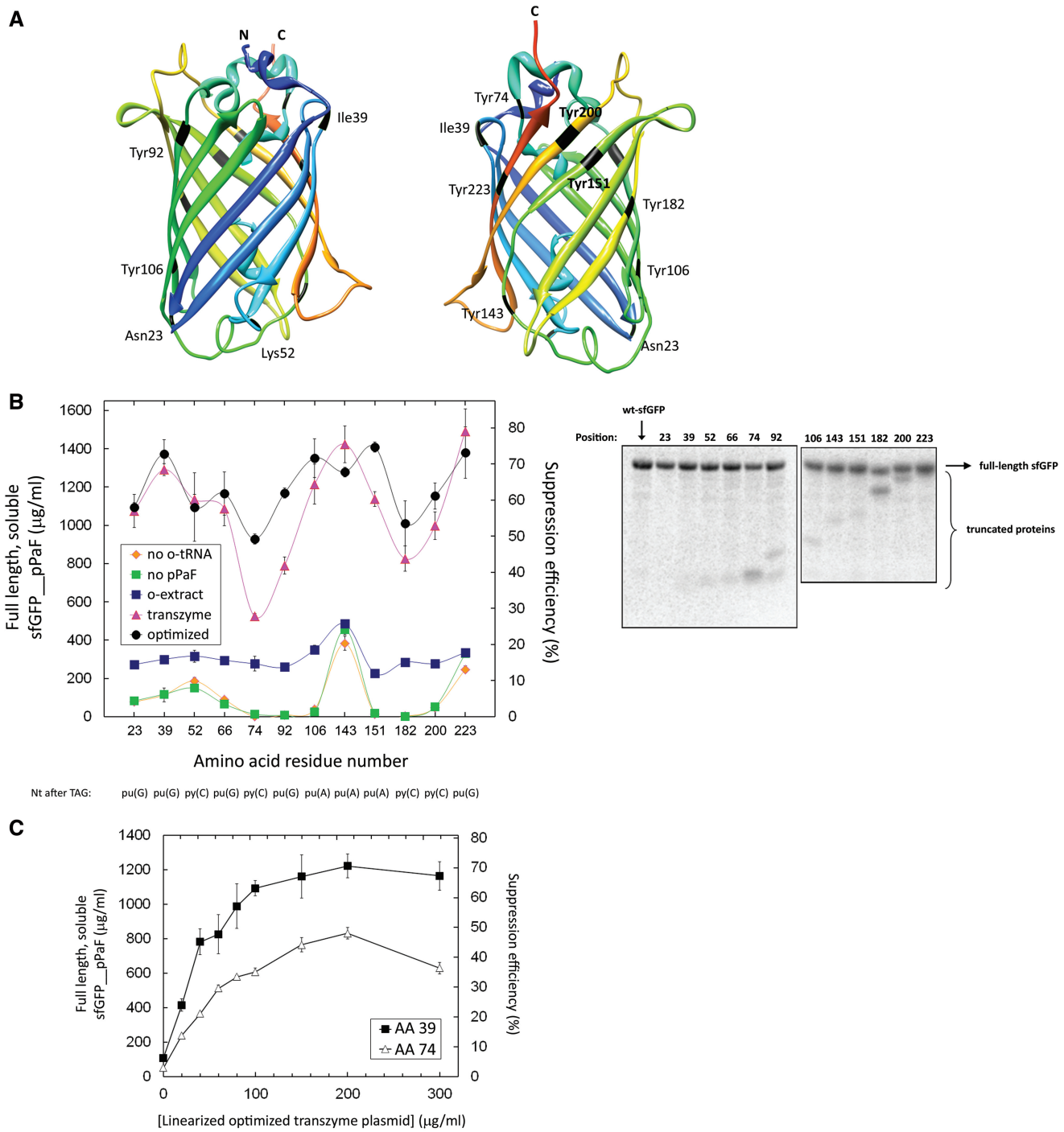

Figure 4. Codon context and nnAA incorporation. (**A**) sfGFP crystal structure; nnAA incorporation sites are labelled in black. The colour scheme shows the position along the primary sequence, with blue and green indicating the N-terminal regions, and orange and red depicting the C-terminal ones. The chromophore was not given in the crystal structure, hence the absence of Tyr66. (**B**) Full-length soluble modified protein yields as a function of the incorporation site under different cell-free conditions. 'No o-tRNA' and 'no pPaF' depict CFPS reactions in which the o-tRNA or the nnAA was omitted from the reaction solutions, respectively. The 'o-extract' series includes the reactions in which the orthogonal extract containing the constitutively expressed o-tRNA was used. The 'transzyme' and 'optimized' reactions contained 200 μg/ml of the linearized transzyme plasmid coding for o-tRNA and o-tRNA^{opt}, respectively. Suppression efficiency is defined as the ratio of accumulated full-length soluble protein containing the nnAA over its natural counterpart (i.e. full-length soluble sfGFP without the nnAA). The bottom line indicates the nucleotides that lie immediately 3′ to the amber stop codon at indicated positions. Error bars depict ±1 standard deviation for three reactions. The right panel shows the SDS-PAGE autoradiogram of radiolabelled sfGFPs. The position of nnAA incorporation is indicated above each lane. (**C**) Titration of linearized o-tRNA^{opt} plasmid templates in the CFPS reactions containing either pY71sfGFP39TAG or pY71sfGFP74TAG. Error bars indicate ±1 standard deviation for nine sfGFP39pPaF reactions, $n = 3$ for the sfGFP74pPaF reactions.

no o-tRNA in the reaction solution at the beginning of the reaction. Since protein synthesis is fastest in the first hour (Figure 3B), we were concerned that some of the protein initially synthesized may contain a natural amino acid instead of the nnAA due to non-specific read-through of the amber stop codon in the absence of adequate o-tRNA supply. Furthermore, for several modified sfGFPs, full-length and soluble protein accumulated when no o-tRNA was added to the reaction solution (Figure 4B). To examine this potential issue, sfGFP39pPaF, sfGFP143pPaF and sfGFP223pPaF were produced, purified via *Strep*-Tactin affinity chromatography, digested using trypsin; and the resulting fragments were analysed by LC/MS to ascertain the homogeneity of the product (Supplementary Figure S2). In the cases of sfGFP39pPaF and sfGFP143pPaF, >99% of the protein contained pPaF at the expected position (Supplementary Figure S2A and B). No wild-type peptide was identified in the sfGFP223pPaF sample (Supplementary Figure S2C); hence all of the detectable protein contained pPaF at position 223. Even though non-specific read-through of the amber stop codon was observed in the absence of one of the orthogonal components, only the modified protein (with the nnAA at the expected position) accumulated in the CFPS solutions in the presence of all three orthogonal components.

Comparison of different ribozymes

We also investigated the effect of using different ribozymes for self-cleaving the transzyme, in place of the hammerhead ribozyme, on the modified protein yields. Our hope was to identify a ribozyme that self-cleaves and liberates the o-tRNA more efficiently. The crystal structure of the full-length hammerhead ribozyme from *Schistosoma mansoni* revealed tertiary interactions that contribute to its 1000-fold higher catalytic efficiency over minimal hammerhead ribozymes such as the one initially used (47). The secondary structures of the three different *Schistosoma* ribozyme variants are given in Supplementary Figure S3. The changes made to the original sequence are explained below.

We first changed the nucleotides that form hydrogen bonds in Stem I of the *Schistosoma* ribozyme to allow these nucleotides to base-pair with the 5′ end of the o-tRNA (47). We also added the five-nucleotide AGAGA loop to covalently link Stem III to the sequence encoding the o-tRNA. Finally, we created three variants of this altered *Schistosoma* ribozyme by mutating the four-nucleotide bulge in Stem I. The first variant contained nucleotides that are able to form additional hydrogen bonds with the o-tRNA. The second variant mimicked the secondary structure of the original *Schistosoma* ribozyme, and the third variant had the original GUAC sequence at the Stem I bulge. We then replaced the hammerhead ribozyme with these variants, titrated different amounts of the linearized o-tRNA templates (with each template coding for a different ribozyme–tRNA fusion) and measured full-length and soluble sfGFP23pPaF concentrations (Figure 5). The hammerhead ribozyme was substantially more efficient than the

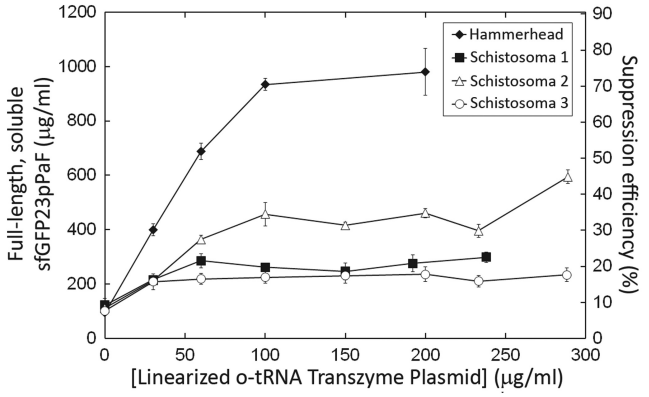

Figure 5. Comparison of modified protein yields obtained with the hammerhead ribozyme and with the three *Schistosoma* ribozyme variants. The four different linearized o-tRNA plasmid templates (each coding for a different ribozyme) were titrated into cell-free reactions at the indicated concentrations. Error bars indicate ±1 standard deviation for three reactions.

Schistosoma ribozymes in liberating the o-tRNA *in situ*, as evidenced by the modified sfGFP23pPaF accumulation. Among the three *Schistosoma* ribozyme variants, the second one was the most active; ~450 μg/ml full-length and soluble sfGFP23pPaF accumulated in the cell-free reactions that contained this variant sequence fused to the o-tRNA template.

Sequential orthogonal tRNA and protein template additions

Finally, we examined whether we could further increase modified protein yields by accumulating o-tRNA in the cell-free reaction before starting protein synthesis. We included the linearized o-tRNA^{opt} plasmid along with the other reagents in the initial cell-free solution, and incubated the reaction for different periods before adding the modified protein template (i.e. sfGFP gene with an amber stop codon at position 39). All of the reactions were incubated at 30°C for 16 h, and the full-length and soluble sfGFP39pPaF yields were determined via scintillation counting and densitometry (Figure 6).

Delaying the protein template addition by up to 30 min was beneficial at 100 and 200 μg/ml o-tRNA^{opt} template concentrations. At 200 μg/ml o-tRNA^{opt} template, when the protein template was added at t = 15 min, full-length and soluble sfGFP39pPaF yields were increased from 1.37 to 1.67 mg/ml, corresponding to a suppression efficiency of 88%. At 100 μg/ml o-tRNA^{opt} template, modified protein yields were also increased from 1.00 to 1.24 mg/ml, when the protein template was added at t = 30 min. Therefore, we could synthesize almost the same amount of modified protein by adding only half as much o-tRNA^{opt} template if the protein template addition is delayed by 30 min. At 200 μg/ml o-tRNA^{opt} template, the modified protein yields were reduced to 1.11 mg/ml, when the protein template was added 2 h after starting the CFPS incubation. This finding was not surprising, as there is a well-documented burst of protein production early in the PANOx SP cell-free system as was also shown in Figure 3B.

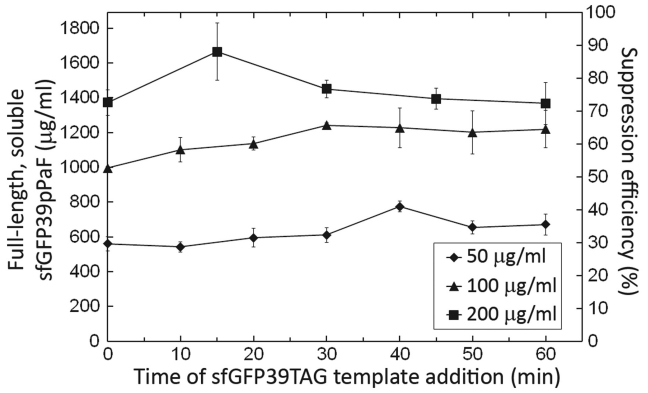

Figure 6. Delaying the protein plasmid template addition. 50, 100 or 200 μg/ml of the linearized o-tRNA^{opt} plasmid template was added along with the other standard reagents at the beginning of the reaction; the sfGFP39TAG template was added at the indicated times. Error bars indicate ±1 standard deviation for three reactions.

Producing the o-tRNA template by PCR

Several steps are required in the preparation of the linearized o-tRNA template. After the ribozyme–tRNA fusion (i.e. the 'transzyme' construct) was cloned into a plasmid, the cells harbouring this plasmid were grown in rich medium. The plasmid was then purified using a commercial kit and linearized using the BstNI restriction enzyme so that the synthesized o-tRNA will have the 3′–CCA terminus required for aminoacylation. Since a very small fraction (<10%) of the plasmid sequence codes for the 'transzyme' construct and the concentration of the linearized o-tRNA plasmid in the CFPS solution is high (200 μg/ml), we sought a more convenient method for producing the 'transzyme' template. Such a method would be particularly useful for high-throughput screens of different tRNA and ribozyme sequences as well as millilitre-scale cell-free reactions.

We therefore synthesized this o-tRNA template from six overlapping oligonucleotides using PCR. The o-tRNA template could just as easily be synthesized from the plasmid using end primers. The PCR product was subsequently ethanol-precipitated, and, without additional purification, was titrated into the cell-free reaction. In Supplementary Figure S4, the primary abscissa shows the actual o-tRNA template concentrations in the CFPS solution, while the secondary abscissa indicates the calculated equivalent mass concentration of the template if the o-tRNA template were part of the linearized plasmid. At 34 μg/ml of transzyme template (which is the molar equivalent of 400 μg/ml linearized o-tRNA^{opt} plasmid), the full-length and soluble sfGFP39pPaF yields were comparable with those obtained with 200 μg/ml of linearized o-tRNA^{opt} plasmid (Figures 3A and 4B). The addition of the Gam protein (at 6.67 μg/ml), which inhibits the intracellular exonuclease RecBCD (61,62), did not increase the accumulation of the modified sfGFP. (The KC6 extract source strain has the *endA* gene deleted from its genome (37,38); this gene codes for Endonuclease I, a non-sequence-specific endonuclease.) The suppression efficiencies may have been higher (~100%) than usual (~70%) because the wild-type sfGFP yields were slightly lower in this set of cell-free reactions (~1.2 instead of 1.6–1.9 mg/ml full-length and soluble protein).

Taken together, these results show that the cell-free reactions can be supplied with adequate o-tRNA^{opt} simply by adding a crude PCR product that encodes the hammerhead ribozyme and the o-tRNA^{opt}. Since a few millilitres of this PCR reaction can produce an amount of o-tRNA^{opt} template equivalent to several milligrams of linearized o-tRNA plasmid, this method can be used to quickly produce enough o-tRNA^{opt} template for millilitre-scale cell-free reactions.

DISCUSSION

We describe here a new method of synthesizing modified proteins that contain nnAAs at specific positions. In the previous protocol, the o-tRNA, one of the three reagents required for site-specific incorporation of nnAAs, was constitutively expressed during the growth of the extract source strain (i.e. the orthogonal extract). The nnAA and the purified synthetase enzyme, evolved separately for each nnAA, were then added to the CFPS solutions along with the standard reagents. With this method, the nnAAs pAzF (33) and pPaF (32,63) were site-specifically incorporated into a variety of proteins to allow the modified proteins to be directly conjugated via Click chemistry (32). Despite the substantial improvement over *in vivo* yields, cell-free yields of modified proteins were almost always lower than their natural counterparts. We previously measured the o-tRNA concentrations and demonstrated that the modified protein yields could be increased by supplementing the cell-free reactions with additional o-tRNA, which was synthesized *in vitro* and subsequently purified (45). However, *in vitro* transcription and purification of the o-tRNA are both labour-intensive and costly. We thus sought a more convenient and inexpensive method for providing adequate o-tRNA to the cell-free reactions.

In this new platform, the o-tRNA is synthesized along with the modified protein in the cell-free reaction. T7 RNA polymerase catalyses the transcription of both the o-tRNA and the mRNA encoding the modified protein. The o-tRNA is synthesized as a fusion with a minimal hammerhead ribozyme. This transzyme cleaves itself during the cell-free reaction, and liberates the o-tRNA. The o-tRNA is acylated with the nnAA by the specific aminoacyl-tRNA synthetase; the acylated o-tRNA then enters the ribosome to suppress the amber stop codon, and adds the nnAA to the nascent polypeptide chain. By using the T7 promoter, the o-tRNA is synthesized inexpensively and at adequate concentrations. Because the standard cell extract that does not harbour the o-tRNA is more productive and the o-tRNA is supplied at higher concentrations, the cell-free yields of modified protein and suppression efficiencies were increased several fold. Here, we report the cell-free accumulation of 0.93 ± 0.03 to 1.67 ± 0.16 mg/ml of pPaF-containing sfGFP, which correspond to 50–88% suppression efficiency (Figures 4B and 6). Mass spectrometry analysis of the purified

cell-free product demonstrated that only the desired modified protein was synthesized (Supplementary Figure S2). To the best of our knowledge, these yields and suppression efficiencies are the highest reported for proteins containing these nnAAs and are among the highest numbers reported for any of the >30 nnAAs site-specifically incorporated via amber stop codon suppression. Previously, the best reported cell-free yields of modified proteins were 660 μg/ml for *E. coli* dihydrofolate reductase (eDHFR10pAzF) (33), and 271 μg/ml for sfGFP216pPaF (32). *In vivo* incorporation of pPaF has been particularly difficult; the best reported modified yields were 2 mg/l for myoglobin (Myo4pPaF) (18). A more recent article describes an optimized *in vivo* expression system, where researchers were able to site-specifically incorporate pPaF and pAzF into GFP with 10% and 25–85% suppression efficiency, respectively (9). (The large variation in suppression efficiency of pAzF-containing GFP was due to the position effect.) Since absolute yields of modified or wild-type GFP were not reported, we cannot make further comparisons.

In addition, the new platform is modular in two different ways: (i) the experimenter does not need to prepare a special orthogonal extract containing o-tRNA, and this is both more convenient and results in higher protein yields and (ii) the o-tRNA concentration can be adjusted to efficiently incorporate >30 nnAAs site-specifically using *E. coli*-based protein synthesis platforms (64). The experimenter can now use the *E. coli*-based commercial cell-free kits and incorporate these nnAAs into virtually any protein by supplementing the cell-free reactions with a crude PCR product that codes for the ribozyme–tRNA fusion, the hexahistidine-tagged synthetase specific to the nnAA (which can be conveniently purified using immobilized metal ion affinity chromatography) and the nnAA itself. In addition, higher concentrations of modified proteins in the crude cell-free solution will increase the yields after downstream purification and covalent coupling processes. This method is now used routinely in our laboratory for site-specific incorporation of the nnAAs pAzF and pPaF into a variety of different proteins.

Adding 200 μg/ml (169 nM) of linearized plasmid template to the CFPS reaction (a plasmid concentration that is much higher than the 6 nM standard concentration for protein templates) produced about 8.9 μM of o-tRNA^{opt}. We believe this high o-tRNA template concentration was required because of limited hammerhead ribozyme activity under the CFPS conditions. The hammerhead ribozyme cleaves itself optimally at 60°C and 30 mM Mg^{2+} (46); in contrast, the CFPS reactions are incubated at 30°C and contain 20 mM Mg^{2+}. However, this is a very complex mixture, and we do not know what the concentration of free Mg^{2+} is in CFPS; the free Mg^{2+} concentration may very well change over the course of the reaction. Under these suboptimal conditions, the uncleaved ribozyme–tRNA transcript would accumulate. We could not verify this hypothesis, as the ribozyme and the uncleaved ribozyme–tRNA fusion were not observed on the tritium radiogram (Supplementary Figure S1). These RNAs were probably digested by the endogenous ribonucleases in the cell-free reaction, and this may also help to explain the high plasmid concentration requirement.

Several other alterations were investigated to further improve this methodology and increase cell-free yields of modified proteins: (i) the protein template addition was delayed to allow for o-tRNA accumulation before starting protein synthesis (Figure 6), (ii) the hammerhead ribozyme was replaced with the catalytically more efficient *Schistosoma* ribozyme (35, Figure 5) and (iii) an o-tRNA template titration was performed for two different amber stop codons to see whether difficult sites (such as residue 74 on sfGFP) can be better suppressed by providing more o-tRNA^{opt} (Figure 4C). Delaying the protein template addition was beneficial. Full-length and soluble sfGFP39pPaF yields were increased by ~20% when the protein template was added 15 min after the reaction was initiated with 200 μg/ml o-tRNA^{opt} template. Our results indicated that this strategy can also be used to reduce the concentrations of o-tRNA^{opt} template without reducing modified protein yields.

The hammerhead ribozyme performed better than the three *Schistosoma* ribozyme variants. We had altered the wild-type *Schistosoma* sequence substantially so that it would hybridize to the o-tRNA and also added a loop to form a contiguous sequence until the cleavage site (Supplementary Figure S3); these changes most likely reduced the catalytic activity of the resultant ribozymes and led to lower sfGFP39pPaF accumulation (Figure 5). In hindsight, the *Schistosoma* ribozyme was probably a poor choice, as it hybridizes extensively with its substrate: 10 nucleotides in the *Schistosoma* ribozyme form hydrogen bonds with the substrate, whereas only 4 nucleotides do so in the minimal hammerhead ribozyme (46). Therefore, when the substrate is changed, these 10 nucleotides in the *Schistosoma* ribozyme must be altered to preserve the hydrogen bonding, which, in turn, probably affects its catalytic activity.

On the other hand, our platform provides the ability to screen ribozymes for improved performance. Full-length and fluorescent sfGFP is obtained only when the ribozyme–tRNA transcript cleaves itself to release the biologically active o-tRNA; as a result, fluorescence is directly coupled to the catalytic activity of the ribozyme. The dynamic range of the signal can be adjusted by changing the tRNA template concentrations, as demonstrated in Figure 5. Finally, the ethanol-precipitated crude PCR product is nearly as effective as the purified linearized plasmid (Supplementary Figure S4). These attributes are also useful for a protocol in which a library of ribozymes would be fused to the o-tRNA^{opt}, synthesized simultaneously with a modified sfGFP protein in CFPS, and screened for fluorescence. The CFPS reactions producing the highest fluorescence would indicate the most effective ribozyme.

The open nature of the CFPS platform was crucial in both the development of this methodology and the analysis of the position effect as it pertains to nnAA incorporation. We evaluated the site-specific incorporation of pPaF into sfGFP at 12 different sites and investigated the effect of Ef-Tu recognition of the o-tRNA on codon

context by comparing two different o-tRNAs (i.e. the o-tRNA and the o-tRNA^{opt}). Our hope was to identify a factor that correlates with the suppression efficiency, to predict the sites most amenable to suppression. Factors reported to be important for the position effect were as follows: the identity of the nucleotide following the stop codon, with purines being better than pyrimidines ($G \geq A > C \geq U$) (43,44,59,60); a beneficial effect of thymidine as the second nucleotide after the stop codon (44); a beneficial effect of lysine as the third amino acid before the stop codon site (65); and the identity of the amino acid two residues before the stop codon site, with acidic amino acids allowing for better suppression than basic ones (66). In addition, previous research in our laboratory showed better nnAA incorporation at amber stop codons closer to the N-terminus of the eDHFR protein (33). Our results for site-specific nnAA incorporation do not directly correlate with any of these factors. They agree with those of Ayer and Yarus, who demonstrated that *in vivo* suppression of the amber stop codon is independent of base-pairing interactions between the nucleotide after the stop codon and the one preceding the anticodon on the suppressor tRNA (67).

We were puzzled to see in Figure 4B that the position effect was less pronounced when the o-tRNA was limiting in the cell-free reactions containing the orthogonal extract (o-extract). These observations suggest that the position effect is not correlated to the availability of the o-tRNA or rather the ternary complex of aminoacylated o-tRNA, Ef-Tu and GTP. If it were correlated, then we would have expected difficult-to-suppress sites such as position 74 to be much less effectively suppressed at limiting o-tRNA concentrations than observed with the o-extract (Figure 4B). If anything, the o-tRNA limitation would have exacerbated the position effect, as more time would have been available for RF1 to enter the A site of the ribosome and terminate translation at the amber stop codon. Further investigation is thus needed to better understand the position effect and its underlying causes.

When the original o-tRNA was synthesized *in situ*, we noticed a large difference across the 12 sites, with the suppression efficiencies varying between 28 and 79%. The poor suppression efficiencies were significantly improved to narrow the range (to 49–75%), when the o-tRNA^{opt} [optimized for better recognition by the *E. coli* elongation factor Ef-Tu (7,9)] and modified sfGFP were simultaneously produced (Figure 4B). These results suggest that the suppression of difficult sites is more effective when the ternary complex between Ef-Tu, GTP and the aminoacylated o-tRNA is formed more effectively.

In addition to the position effect, modified protein yields are also dependent on how well the wild-type protein is synthesized. Therefore, the suppression efficiency, which is defined as the ratio of the modified protein yield to the native protein yield, provides a useful and normalized metric for assessing nnAA incorporation. The suppression efficiency can be expressed as:

$$\textit{Suppression Efficiency} = 1 - \textit{Fraction of } \text{RF1 } \textit{insertion} \quad (1)$$

where *Fraction of* RF1 *insertion* indicates the fraction of the time RF1 suppresses the amber stop codon in place of the ternary complex and prematurely terminates translation. This fraction would be expected to vary directly with the RF1 concentration and inversely with the concentration of the ternary complex.

The ultimate goal in site-specific nnAA incorporation is reaching 100% suppression efficiency, i.e. when the nnAA incorporation pathway would no longer be limiting. Across the 12 incorporation sites, we were able to increase the suppression efficiency from ~25% to ~50–75% by removing limitations in orthogonal component supply and in Ef-Tu recognition of the acylated o-tRNA. Even though the RF1 concentration is greatly reduced in the CFPS platform (from about 7.6 μM to 380 nM) due to the dilution of endogenous macromolecules (68), RF1 is still present in the CFPS platform and it can outcompete the aminoacylated o-tRNA as evidenced by the accumulation of prematurely truncated protein products (Figure 4B). We believe this competition with RF1 limits the modified protein yields and suppression efficiencies in the new platform. Elimination of RF1 through various methods has been shown to increase amber stop codon suppression both *in vivo* and *in vitro* (69–73). Recently, researchers were able to convert the 314 amber stop codons in the chromosome to TAA codons across different strains using multiplex automated genome engineering (MAGE), which would permit the removal of the *prfA* gene (which codes for RF1) (74). It would be interesting to repeat these experiments using an extract that is generated from a strain devoid of RF1, and observe its effects on suppression efficiency and the codon context. Apart from competition by RF1, the presence of one or several nnAAs may perturb the three-dimensional fold and induce a limitation in protein folding or may render the modified protein more prone to aggregation, thereby effectively reducing modified protein yields and suppression efficiencies.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

The authors thank Dr Aaron Goerke for the pK7pAzFRS-His6 plasmid, Dr Brad Bundy for the pY71sfGFP216TAG plasmid; Drs William Yang and Kedar Patel for the pPaF, pPaFRS and T7 RNAP reagents; Dr Phil Smith for the Pfu DNA polymerase and the Gam protein; and Dr Chris Adams at the Stanford University Mass Spectrometry Laboratory for LC/MS analysis. The mass spectrometry described was supported by Award Number S10RR027425 from the National Center for Research Resources.

FUNDING

Wallace H. Coulter Foundation; Stanford University School of Engineering. Funding for open access

charge: Department of Chemical Engineering or School of Engineering at Stanford University.

Conflict of interest statement. None declared.

REFERENCES

1. Cornish,V.W., Mendel,D. and Schultz,P.G. (1995) Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed. Engl.*, **34**, 621–633.
2. Ellman,J., Mendel,D., Anthony-Cahill,S., Noren,C.J. and Schultz,P.G. (1991) Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol.*, **202**, 301–336.
3. Bonnefond,L., Giegé,R. and Rudinger-Thirion,J. (2005) Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems. *Biochimie*, **87**, 873–883.
4. Kobayashi,T., Nureki,O., Ishitani,R., Yaremchuk,A., Tukalo,M., Cusack,S., Sakamoto,K. and Yokoyama,S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.*, **10**, 425–432.
5. Wang,L. and Schultz,P.G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol.*, **8**, 883–890.
6. Liu,C.C. and Schultz,P.G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.*, **79**, 413–444.
7. Guo,J., Melançon,C.E., Lee,H.S., Groff,D. and Schultz,P.G. (2009) Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew. Chem. Int. Ed. Engl.*, **48**, 9148–9151.
8. Ryu,Y. and Schultz,P.G. (2006) Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods*, **3**, 263–265.
9. Young,T.S., Ahmad,I., Yin,J.A. and Schultz,P.G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.*, **395**, 361–374.
10. Wang,L., Zhang,Z., Brock,A. and Schultz,P.G. (2003) Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **100**, 56–61.
11. Zhang,Z., Smith,B.A.C., Wang,L., Brock,A., Cho,C. and Schultz,P.G. (2003) A new strategy for the site-specific modification of proteins *in vivo*. *Biochemistry*, **42**, 6735–6746.
12. Xie,J., Wang,L., Wu,N., Brock,A., Spraggon,G. and Schultz,P.G. (2004) The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotechnol.*, **22**, 1297–1301.
13. Wang,L., Xie,J., Deniz,A.A. and Schultz,P.G. (2003) Unnatural amino acid mutagenesis of green fluorescent protein. *J. Organic Chem.*, **68**, 174–176.
14. Chin,J.W., Martin,A.B., King,D.S., Wang,L. and Schultz,P.G. (2002) Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **99**, 11020–11024.
15. Wang,L., Brock,A., Herberich,B. and Schultz,P.G. (2001) Expanding the genetic code of *Escherichia coli*. *Science*, **292**, 498–500.
16. Santoro,S.W., Wang,L., Herberich,B., King,D.S. and Schultz,P.G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nature Biotechnol.*, **20**, 1044–1048.
17. Chin,J.W., Santoro,S.W., Martin,A.B., King,D.S., Wang,L. and Schultz,P.G. (2002) Addition of p-Azido- L -phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.*, **124**, 9026–9027.
18. Deiters,A. and Schultz,P.G. (2005) *In vivo* incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg. Med. Chem. Lett.*, **15**, 1521–1524.
19. Nguyen,D.P., Lusic,H., Neumann,H., Kapadnis,P.B., Deiters,A. and Chin,J.W. (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.
20. Nguyen,D.P., Elliott,T., Holt,M., Muir,T.W. and Chin,J.W. (2011) Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J. Am. Chem. Soc.*, **133**, 11418–11421.
21. Plass,T., Milles,S., Koehler,C., Schultz,C. and Lemke,E.A. (2011) Genetically encoded copper-free click chemistry. *Angew. Chem. Int. Ed. Engl.*, **50**, 3878–3881.
22. Wan,W., Huang,Y., Wang,Z., Russell,W.K., Pai,P.J., Russell,D.H. and Liu,W.R. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew. Chem. Int. Ed. Engl*, **49**, 3211–3214.
23. Wang,Y.S., Fang,X., Wallace,A.L., Wu,B. and Liu,W.R. (2012) A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J. Am. Chem. Soc.*, **134**, 2950–2953.
24. Yanagisawa,T., Ishii,R., Fukunaga,R., Kobayashi,T., Sakamoto,K. and Yokoyama,S. (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem. Biol.*, **15**, 1187–1197.
25. Wang,Y.S., Fang,X., Chen,H.-Y., Wu,B., Wang,Z.U., Hilty,C. and Liu,W.R. (2013) Genetic incorporation of twelve meta-substituted phenylalanine derivatives using a single pyrrolysyl-tRNA synthetase mutant. *ACS Chem. Biol.*, **8**, 405–415.
26. Boyer,M., Wang,C. and Swartz,J. (2006) Simultaneous expression and maturation of the iron-sulfur protein ferredoxin in a cell-free system. *Biotechnol. Bioeng.*, **94**, 128–138.
27. Boyer,M.E., Stapleton,J.A., Kuchenreuther,J.M., Wang,C. and Swartz,J.R. (2008) Cell-free synthesis and maturation of [FeFe] hydrogenases. *Biotechnol. Bioeng.*, **99**, 59–67.
28. Bundy,B.C., Franciszkowicz,M.J. and Swartz,J.R. (2008) *Escherichia coli*-based cell-free synthesis of virus-like particles. *Biotechnol. Bioeng.*, **100**, 28–37.
29. Hovijitra,N.T., Wuu,J.J., Peaker,B. and Swartz,J.R. (2009) Cell-free synthesis of functional aquaporin Z in synthetic liposomes. *Biotechnol. Bioeng.*, **104**, 40–49.
30. Wuu,J.J. and Swartz,J.R. (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochim. Biophys. Acta*, **1778**, 1237–1250.
31. Yang,W.C., Patel,K.G., Lee,J., Ghebremariam,Y.T., Wong,H.E., Cooke,J.P. and Swartz,J.R. (2009) Cell-free production of transducible transcription factors for nuclear reprogramming. *Biotechnol. Bioeng.*, **104**, 1047–1058.
32. Bundy,B.C. and Swartz,J.R. (2010) Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation. *Bioconjug. Chem.*, **21**, 255–263.
33. Goerke,A.R. and Swartz,J.R. (2009) High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. *Biotechnol. Bioeng.*, **102**, 400–416.
34. Patel,K.G., Ng,P.P., Kuo,C.-C., Levy,S., Levy,R. and Swartz,J.R. (2009) Cell-free production of gaussia princeps luciferase–antibody fragment bioconjugates for ex vivo detection of tumor cells. *Biochem. Biophys. Res. Commun.*, **390**, 971–976.
35. Welsh,J.P., Patel,K.G., Manthiram,K. and Swartz,J.R. (2009) Multiply mutated Gaussia luciferases provide prolonged and intense bioluminescence. *Biochem. Biophys. Res. Commun.*, **389**, 563–568.
36. Zawada,J.F., Yin,G., Steiner,A.R., Yang,J., Naresh,A., Roy,S.M., Gold,D.S., Heinsohn,H.G. and Murray,C.J. (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.
37. Calhoun,K.A. and Swartz,J.R. (2006) Total amino acid stabilization during cell-free protein synthesis reactions. *J. Biotechnol.*, **123**, 193–203.
38. Michel-Reydellet,N., Calhoun,K. and Swartz,J. (2004) Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome. *Metab. Eng.*, **6**, 197–203.
39. Jewett,M.C. and Swartz,J.R. (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.*, **86**, 19–26.
40. Jewett,M.C., Calhoun,K.A., Voloshin,A., Wuu,J.J. and Swartz,J.R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.*, **4**, 1–10.

41. 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.
42. Hoesl,M.G. and Budisa,N. (2012) Recent advances in genetic code engineering in *Escherichia coli*. *Curr. Opin. Biotechnol.*, **23**, 751–757.
43. Phillips-Jones,M.K., Watson,F.J. and Martin,R. (1993) The 3' codon context effect on UAG suppressor tRNA is different in *Escherichia coli* and Human Cells. *J. Mol. Biol.*, **233**, 1–6.
44. Stormo,G.D., Schneider,T.D. and Gold,L. (1986) Quantitative analysis of the relationship between nucleotide sequence and functional activity. *Nucleic Acids Res.*, **14**, 6661–6679.
45. Albayrak,C. and Swartz,J.R. (2013) Using *E. coli*-based cell-free protein synthesis to evaluate the kinetic performance of an orthogonal tRNA and aminoacyl-tRNA synthetase pair. *Biochem. Biophys. Res. Commun.*, **431**, 291–295.
46. Fechter,P., Rudinger,J., Giegé,R. and Théobald-Dietrich,A. (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. *FEBS Lett.*, **436**, 99–103.
47. Martick,M. and Scott,W.G. (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell*, **126**, 309–320.
48. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
49. Cabantous,S., Terwilliger,T.C. and Waldo,G.S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.*, **23**, 102–107.
50. Pédelacq,J.D., Cabantous,S., Tran,T., Terwilliger,T.C. and Waldo,G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.*, **24**, 79–88.
51. Uemori,T., Ishino,Y., Toh,H., Asada,K. and Kato,I. (1993) Organization and nucleotide sequence of the DNA polymerase gene from the archaeon Pyrococcus furiosus. *Nucleic Acids Res.*, **21**, 259–265.
52. Dabrowski,S. and Kur,J. (1998) Cloning and expression in *Escherichia coli* of the recombinant his-tagged DNA polymerases from *Pyrococcus furiosus* and *Pyrococcus woesei*. *Protein Expr. Purif.*, **14**, 131–138.
53. Goerke,A.R. and Swartz,J.R. (2008) Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechn. Bioeng.*, **99**, 351–367.
54. Deiters,A., Cropp,T.A., Mukherji,M., Chin,J.W., Anderson,J.C. and Schultz,P.G. (2003) Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae. *J. Am. Chem. Soc.*, **125**, 11782–11783.
55. Li,Y., Wang,E. and Wang,Y. (1999) A modified procedure for fast purification of T7 RNA polymerase. *Protein Expr. Purif.*, **16**, 355–358.
56. Budyka,M.F. (2007) Photochemistry of aromatic azides and nitrenes. *High Energy Chem.*, **41**, 176–187.
57. Calhoun,K.A. and Swartz,J.R. (2005) An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates. *Biotechnol. Prog.*, **21**, 1146–1153.
58. Imburgio,D., Rong,M., Ma,K. and McAllister,W.T. (2000) Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry*, **39**, 10419–10430.
59. Miller,J.H. and Albertini,A.M. (1983) Effects of Surrounding Sequence on the Suppression of Nonsense Codons. *J. Mol. Biol.*, **164**, 59–71.
60. Pedersen,W.T. and Curran,J.F. (1991) Effects of the Nucleotide 3′ to an Amber Codon on Ribosomal Selection Rates of Suppressor tRNA and Release Factor-1. *J. Mol. Biol.*, **219**, 231–241.
61. Sitaraman,K., Esposito,D., Klarmann,G., Le Grice,S.F., Hartley,J.L. and Chatterjee,D.K. (2004) A novel cell-free protein synthesis system. *J. Biotechnol.*, **110**, 257–263.
62. Lui,B.H., Cochran,J.R. and Swartz,J.R. (2011) Discovery of improved EGF agonists using a novel *in vitro* screening platform. *J. Mol. Biol.*, **413**, 406–415.
63. Patel,K.G. and Swartz,J.R. (2011) Surface functionalization of virus-like particles by direct conjugation using azide-alkyne click chemistry. *Bioconjug. Chem.*, **22**, 376–387.
64. Young,T.S. and Schultz,P.G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem.*, **285**, 11039–11044.
65. Mottagui-Tabar,S. and Isaksson,L.A. (1997) Only the last amino acids in the nascent peptide influence translation termination in *Escherichia coli* genes. *FEBS Lett.*, **414**, 165–170.
66. Mottagui-Tabar,S., Björnsson,A. and Isaksson,L.A. (1994) The second to last amino acid in the nascent peptide as a codon context determinant. *EMBO J.*, **13**, 249–257.
67. Ayer,D. and Yarus,M. (1986) The context effect does not require a fourth base pair. *Science*, **231**, 393–395.
68. Adamski,F., McCaughan,K., Jorgensen,F., Kurland,C. and Tate,W. (1994) The Concentration of Polypeptide Chain Release Factors 1 and 2 at Different Growth Rates of *Escherichia coli*. *J. Mol. Biol.*, **238**, 302–308.
69. Short,G.F., Golovine,S.Y. and Hecht,S.M. (1999) Effects of release factor 1 on *in vitro* protein translation and the elaboration of proteins containing unnatural amino acids. *Biochemistry*, **38**, 8808–8819.
70. Sando,S., Ogawa,A., Nishi,T., Hayami,M. and Aoyama,Y. (2007) *In vitro* selection of RNA aptamer against *Escherichia coli* release factor 1. *Bioorg. Med. Chem. Lett.*, **17**, 1216–1220.
71. Mukai,T., Hayashi,A., Iraha,F., Sato,A., Ohtake,K., Yokoyama,S. and Sakamoto,K. (2010) Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res.*, **38**, 8188–8195.
72. Johnson,D.B.F., Xu,J., Shen,Z., Takimoto,J.K., Schultz,M.D., Schmitz,R.J., Xiang,Z., Ecker,J.R., Briggs,S.P. and Wang,L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.*, **7**, 779–786.
73. Loscha,K.V., Herlt,A.J., Qi,R., Huber,T., Ozawa,K. and Otting,G. (2012) Multiple-site labeling of proteins with unnatural amino acids. *Angew. Chem. Int. Ed. Engl.*, **51**, 2243–2246.
74. Isaacs,F.J., Carr,P.A., Wang,H.H., Lajoie,M.J., Sterling,B., Kraal,L., Tolonen,A.C., Gianoulis,T.A., Goodman,D.B., Reppas,N.B. *et al.* (2011) Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science*, **333**, 348–353.