Dissecting limiting factors of the Protein synthesis Using Recombinant Elements (PURE) system

Jun Li, Chi Zhang, Poyi Huang, Erkin Kuru, Eliot T. C. Forster-Benson, Taibo Li, and George M. Church

ABSTRACT
Reconstituted cell-free protein synthesis systems such as the Protein synthesis Using Recombinant Elements (PURE) system give high-throughput and controlled access to in vitro protein synthesis. Here we show that compared with the commercial S30 crude extract based RTS 100 E. coli HY system, the PURE system has less mRNA degradation and produces up to ~6-fold full-length proteins. However the majority of polypeptides PURE produces are partially translated or inactive since the signal from firefly luciferase (Fluc) translated in PURE is only ~2/3rd of that measured using the RTS 100 E. coli HY S30 system. Both of the 2 batch systems suffer from low ribosome recycling efficiency when translating proteins from 82 kD to 224 kD. A systematic fed-batch analysis of PURE shows replenishment of 6 small molecule substrates individually or in combination before energy depletion increased Fluc protein yield by ~1.5 to ~2-fold, while creatine phosphate and magnesium have synergistic effects when added to the PURE system. Additionally, while adding EF-P to PURE reduced full-length protein translated, it increased the fraction of functional protein and reduced partially translated protein probably by slowing down the translation process. Finally, ArfA, rather than YaeJ or PrfH, helped reduce ribosome stalling when translating Fluc and improved system productivity in a template-dependent fashion.

Introduction
Cell-free protein synthesis (CFPS) enables rapid expression of functional proteins from their genes. It is well applied in display technologies, unnatural amino acids incorporation into proteins, and in vitro protein evolution and provides a platform to study protein translation and folding. Conventional CFPS systems use cell lysates from E. coli cells, wheat germ or rabbit reticulocytes to supply the protein translation machinery and a recombinant T7 RNA polymerase to couple transcription to translation. As a more stripped-down and defined CFPS system, “Protein synthesis Using Recombinant Elements” (PURE) system is reconstituted solely from elements essential to E. coli translation. For example, as opposed to the E. coli extract-based S30 system, the PURE system is devoid of cellular contaminants such as proteases and nuclease inhibitory to protein synthesis. With controllable system components and conditions, the PURE system is an ideal starting platform to construct the minimal self-replicating system.

Conventional batch format PURE experiment has the capacity for a yield of over 100 µg/mL, while batch format S30 system, such as RTS 100 E. coli HY system, claims a yield range from 40–400 µg/mL. Note that yields here count both functional and unfunctional proteins in. However, the batch PURE reaction often lasts for a shorter period of time than the batch format S30 system. Depletion of primary and secondary energy substrate, accumulation of inhibitory byproducts, and shortage of nutrients such as tRNAs, amino acids and other small molecule substrates could be possible explanations for the reaction termination. Also, S30 system contains additional identified and unidentified cellular factors capable of enhancing protein translation and folding.
while PURE system doesn’t. For example, supplementing the PURE system with fractions from a size-exclusion separation of the S30 extract significantly increases the product activity to a level close to that of the S30 system.26 Although the exact composition of these PURE-enhancing fractions are still not clear, it likely contains at least a chaperone as adding the GroEL/ES chaperone system to the PURE system alone can increase the functional firefly luciferase yield by 60%.27

Additionally, the PURE system produces a significant amount of partially translated products most likely due to ribosome stalling on mRNA.27,28 In E. coli, ribosome rescue is accomplished through 3 systems: 1- trans-translation mediated by transfer mRNA (tmRNA) and a partner protein, SmpB,29 2- ArfA with release factor-2 (RF2)30,31 which can release stalled ribosomes from mRNAs with no in-frame stop codons and 3- YaeJ (ArfB).32 In addition, the elongation factor P (EF-P) alleviates ribosome stalling by enhancing translocation of proline codon stretches,33,34 and elongation factor 4 (EF4) stimulates protein synthesis under high Mg$^{2+}$ concentration.35 This list is still growing. For example, PrfH, a pseudogene from E. coli K12, which contains a highly conserved GGQ motif associated with peptidyl hydrolysis activity, has been proposed to respond to signals other than conventional stop codons.36 However, no experimental evidence has been reported on the function of PrfH.

To dissect the limiting factors of PURE protein synthesis, we first characterized PURE protein synthesis by measuring reaction duration, functional, fully and partially translated products. Then we compared PURE with the RTS 100 E. coli HY S30 system in terms of mRNA stabilities, functional, fully and partially translated product yields and ribosome recycling efficiencies. To zoom in to see which substrate depletion limits PURE protein synthesis, we supplemented the initial reaction with higher concentrations of tRNAs, amino acids (AA), the reducing agent dithiothreitol (DTT), 10-Formyl-5,6,7,8-tetrahydrofolic acid (10-CHO-THF), NTPs, the secondary energy substrate creatine phosphate and Mg$^{2+}$ individually or in combination to the PURE system. In parallel, we monitored the accumulation of the inhibitory byproduct, inorganic phosphate, using a chemical assay. Besides, we supplemented the PURE system with purified missing protein factors present in cells and S30 but absent in PURE to study how these factors affect translation. We show adding EF-P increased the final yield of functional product while decreased fully and partially translated product yields even with a template with no poly-proline stretches, possibly by slowing down translation initiation or elongation. EF4 partially rescued functional, fully and partially translated product yields in PURE with elevated Mg$^{2+}$ concentrations. ArfA, but not YaeJ or PrfH, helped increase yield when linear template with no stop codon encoding large protein was expressed in the PURE system.

**Results**

**Measurement of protein synthesis duration in PURE system**

The PURE system is reconstituted from T7 RNA polymerase, ribosomes, initiation factors (IF1, IF2, and IF3), elongation factors (EF-G, EF-Tu, and EF-Ts), release factors (RF1, RF2, and RF3), ribosome recycling factor (RRF), 20 aminoacyl-tRNA synthetases (ARSs), methionyl-tRNA transformylase (MTF), 10-CHO-THF, 46 tRNAs, 20 amino acids, NTPs, DTT, creatine phosphate, creatine kinase, myokinase, nucleoside-diphosphate kinase, and pyrophosphatase. To begin, we assessed the duration of PURE protein synthesis with 3 plasmid templates encoding proteins with different lengths: Namely firefly luciferase with an N-terminal tetra-cysteine tag (TC-Fluc, 63 kD), a fusion protein of HaloTag, mCherry and TolA (HT-Ch-TolA, 82 kD) and a fusion protein of HaloTag, EntF, mCherry and TolA (HT-EntF-Ch-TolA, 224 kD). TolA is a C-terminal 171-amino-acid α-helical spacer excised from E. coli TolA domain II. EntF is a multidomain enzyme from E. coli enterobactin biosynthetic pathway. Templates with completely different sequences, i.e. TC-Fluc compared with HT-Ch-TolA and HT-EntF-Ch-TolA, and similar sequences, i.e., HT-Ch-TolA compared with HT-EntF-Ch-TolA, were used to examine whether the duration of protein synthesis is sequence dependent.

We first measured functional TC-Fluc produced in the PURE system over time with a luciferase assay by converting relative luminescence units to the actual amount of TC-Fluc via a premeasured standard curve. The amount of functional TC-Fluc produced increased as the reaction went on and reached a maximum of ~114 nmol/L (7.2 μg/mL) at 90 min (Fig. S1A), indicating the reaction stopped after ~90 min. This maximum yield is similar to what
Jewett et al. claimed in 37. For the expression of HT-Ch-TolA and HT-EntF-Ch-TolA, we observed a protein-specific, unique fluorescence pattern for both HT-Ch-TolA and HT-EntF-Ch-TolA (Fig. S1B). This strongly suggests that besides full-length products, partially translated polypeptides were also abundant, possibly due to premature termination or pausing during translation.28,38 Finally, the intensity of bands of full-length and partial products reached a maximum ~2 h (Fig. S1B) as quantified in pixels (Table S1), indicative of the stopping point for the PURE reaction under these conditions. Therefore, PURE system protein synthesis terminated within 2 h, regardless of which template/sequence was used.

**Comparing PURE system transcription and translation with RTS 100 E. coli HY S30 system**

To compare PURE system transcription and translation with cell extract based S30 system, we chose TC-Fluc as a model protein. Note that RTS 100 E. coli HY system was chosen as a sample S30 system for this study and yields of S30 system from other suppliers or different preparation methods may vary. Circular plasmid encoding TC-Fluc was added to PURE system and RTS 100 E. coli HY S30 system to the same final concentration of 10 ng/μl. The PURE system reaction was incubated at 37°C for 2 h, while the S30 system reaction was incubated at 30°C for 6 h. Analysis of total RNA extracted from the PURE system reaction revealed 23S, 16S, 5S rRNA (rRNA), tRNAs and the 1.9 kb T7 TC-Fluc mRNA when compared with the native 23S, 16S, 5S rRNA and tRNAs as markers (Fig. 1A–B). Bands larger than the 2.9 kb 23S rRNA (indicated by asterisk *) also appeared on the gel, probably due to the inefficient termination of T7 polymerase.39 No degraded mRNA template was found in the PURE reaction, suggesting that the poorer sustainability of the PURE system is not incurred by ribosome stalling due to truncated mRNA templates (Fig. 1A). In contrast, the total RNA from RTS 100 E. coli HY S30 system reaction contained not only T7 TC-Fluc mRNA appearing at approximately 1.9 kb, but also a smear of different bands smaller than the 16S band indicating partial degradation of the RNA template (Fig. 1B).

To compare the efficiency of translation, TC-Fluc was synthesized in PURE and RTS 100 E. coli HY S30 systems and the yields were quantified by luciferase assay with equal volumes of PURE and S30 reactions. RTS 100 E. coli HY S30 system produced ~138 nmol/L (~9 μg/mL) functional TC-Fluc, which was ~1.5-fold of that produced by PURE system (Fig. 1C). We also quantified yields of full-length proteins, including functional and unfunctional fractions, by mixing equal volumes of PURE and S30 reactions with FLAsH-EDT₂, a biarsenical labeling reagent to label proteins with a TC tag, and running the samples on a SDS-PAGE and detecting fluorescence. Although the control reactions with no plasmid template showed some fluorescent bands and smearing indicative of non-specific FLAsH-EDT₂ labeling (lane 1 and 4 in Fig. 1D), the full-length and incomplete TC-Fluc could still be clearly detected in both of the 2 systems (Fig. 1D).

Based on the band intensity, the PURE system produced ~358 μg/mL full-length TC-Fluc, while the RTS 100 E. coli HY S30 system produced ~58 μg/mL full-length TC-Fluc. All of these values are within the claimed yield ranges of the 2 systems. The significantly lower activity of TC-Fluc in PURE (Fig. 1C) suggests that a large portion of the proteins produced by the PURE system is not functional, most likely due to protein misfolding.

Majority of the bands and smearing other than the full-length product in RTS 100 E. coli HY S30 system came from non-specific labeling of FLAsH-EDT₂. Since the PURE system produced also a large amount of incomplete products of different lengths, we wanted to minimize the interference from non-specific labeling. Therefore, we expressed 3 HaloTag fusion proteins HT-Ch-TolA, HT-EntE-Ch-TolA (EntE is also a multidomain enzyme from E. coli enterobactin biosynthetic pathway) and HT-EntF-Ch-TolA in PURE and RTS 100 E. coli HY S30 systems. HT-Ch-TolA, HT-EntE-Ch-TolA and HT-EntF-Ch-TolA were prepared as linear DNA templates with 2 versions: with and without stop codon. We labeled the products with HaloTag TMR ligand, run them on SDS-PAGE and compared the in-gel fluorescence. The PURE system produced ~2-fold more full-length HT-Ch-TolA, HT-EntE-Ch-TolA and HT-EntF-Ch-TolA compared with the RTS 100 E. coli HY S30 system (Fig. 2). Consistent with the TC labeling, HT labeling confirmed that the PURE system produced larger amounts of incomplete polypeptide products than S30 for all of the 3 proteins (Fig. 2), each with characteristic unique patterns. This suggests that PURE system might be lacking functionalities against premature protein
termination perhaps due to ribosome stalling. Templates with and without stop codon didn’t show much difference on the yield of full-length and partially translated products for all of the 3 proteins in both PURE and RTS 100 E. coli HY S30 systems (Fig. 2), possibly due to low ribosome recycling efficiency.

Figure 1. Compare PURE system transcription and translation with RTS 100 E. coli HY S30 system. (A) RNA denaturing gel of PURE reaction transcribing TC-Fluc mRNA. PURE system with no plasmid template and plasmid encoding TC-Fluc were incubated at 37° C for 2 h. Total RNA were purified by phenol/chloroform extraction and isopropanol precipitation from equal volumes of the 2 reactions and then run on lane 2 and 3, respectively. (B) RNA denaturing gel of RTS 100 E. coli HY S30 system transcribing TC-Fluc mRNA. RTS reactions with no plasmid template and plasmid encoding TC-Fluc were incubated at 30° C for 6 h. Total RNAs were purified by phenol/chloroform extraction and isopropanol precipitation from equal volumes of the 2 reactions and then run on lane 2 and 3, respectively. (C) Assessment of functional TC-Fluc translated in PURE and RTS 100 E. coli HY S30 systems. Equal volume aliquots were taken for luciferase assay to measure the actual amount of functional TC-Fluc produced. Values represent averages and error bars are ± standard deviations, with n = 3. (D) Assessment of full-length TC-Fluc produced in PURE and RTS 100 E. coli HY S30 systems. TC-Fluc synthesized in (C) were incubated with FlAsH-EDT₂ biarsenical labeling reagent and analyzed on a SDS-PAGE. The gel was scanned by a typhoon scanner with filter set (508nmEx/528nmEm). Asterisk * indicates full-length product.
Previously we showed that optimization of IFs, EFs, RFs and RRF concentrations can increase the yields of functional Fluc in the PURE system to various extents. Historically, nonproductive energy consumption, secondary energy source depletion, and amino acid exhaustion have been identified as the primary reasons for early termination of CFPS systems that use plasmid DNA templates. To study whether small molecule substrate availability restricted the productivity of the PURE system, we performed fed-batch experiments by replenishing tRNAs, AA, DTT, 10-CHO-THF, NTPs, the secondary energy substrate creatine phosphate and Mg$^{2+}$ individually or in combination at 15 min, 30 min, 45 min, 60 min or 75 min after the reaction started and measured the amount of functional TC-Fluc produced at 90 min, the time-point at which the TC-Fluc synthesis stopped (Fig. S1A).

Feeding 0.3 mM AA or 27 OD$_{260}$/ml tRNAs once at different time points only slightly increased functional TC-Fluc yield (Fig. 3A–B). But when combined together, it increased the TC-Fluc signal maximum ~80% above that of the control (water supplement only) when fed at 75 min (Fig. 3C). The system uses DTT as reducing agent to prevent intramolecular and intermolecular disulfide bonds from forming between cysteine residues of proteins. We replenished 1 mM DTT to the system and found that feeding at 15 min, 30 min, 45 min, 60 min only slightly improved the yield of functional TC-Fluc, while feeding at 75 min gave an increase of ~50% compared with the control (Fig. 3D), indicating that DTT might have been consumed as a reduction agent after the reaction started.

In the PURE system, the enzyme MTF formylates Met-tRNA$^f_{Met}$ using 10-CHO-THF to form fMet-tRNA$^f_{Met}$ for translation initiation. Replenishing 10 mg/ml 10-CHO-THF boosted functional TC-Fluc yield at all time points with a maximum of ~50% at 60 min or 75 min (Fig. 3E).

Not surprisingly, the depletion of creatine phosphate as the secondary energy source was coupled with a rapid increase in inorganic phosphate which sequesters magnesium present in the system necessary for operation of the translation apparatus. We observed that inorganic phosphate concentrations reached ~20 mM after incubating the reaction at 37°C for 3 h (Fig. S2). Thus we fed 2 mM magnesium acetate to offset the deleterious effects of magnesium sequestration and observed a boost on yield at all time points. A maximum of ~50% increase in functional TC-Fluc yield was observed when feeding at 75 min (Fig. 3F). Feeding 4.8 mM magnesium acetate had an inhibitory effect on TC-Fluc translation (Fig. S3A).

Feeding 1.6 mM ATP, 1.6 mM GTP, 0.8 mM CTP and 0.8 mM UTP inhibited translation of TC-Fluc (Fig. 3G) regardless of feeding time points, probably due to NTPs’ chelation with magnesium at a 1:1 molar ratio. Consistently, feeding NTPs at later time points had less inhibitory effect than feeding at earlier time points, that could be explained by NTPs’ being consumed as the reaction went on (Fig. 3G). To offset this chelating effect, we also supplemented 4.8 mM magnesium acetate together with 1.6 mM ATP, 1.6 mM GTP, 0.8 mM CTP and 0.8 mM UTP to the system and observed no effect on functional TC-Fluc yield.
Figure 3. Replenishing small molecule substrates in the PURE system protein synthesis at different time points. PURE system reactions synthesising TC-Fluc were set up and incubated at 37°C. Small molecule substrates as indicated were fed to the reaction at 15 min, 30 min, 45 min, 60 min or 75 min. Control reactions were fed equivalent volumes of water at 0 min. Overall functional TC-Fluc produced was assessed by luciferase assay at 90 min. Values represent averages and error bars are ± standard deviations, with n = 3. PURE system feed supplements of (A) 0.3 mM amino acids; (B) 27 OD$_{260}$/ml tRNAs; (C) 0.3 mM amino acids and 27 OD$_{260}$/ml tRNAs in combination; (D) 1 mM DTT; (E) 10 μg/ml 10-CHO-THF; (F) 2 mM magnesium acetate; (G) 1.6 mM ATP, 1.6 mM GTP, 0.8 mM CTP and 0.8 mM UTP; (H) 4.8 mM magnesium acetate together with 1.6 mM ATP, 1.6 mM GTP, 0.8 mM CTP and 0.8 mM UTP; (I) 20 mM creatine; and (J) 20 mM creatine phosphate and 2 mM magnesium acetate in combination.
suggesting that the availability of NTPs is not a key limiting factor for TC-Fluc synthesis in the PURE system (Fig. 3H). Creatine phosphate was fed at its initial concentration of 20 mM and reached maximum increase in functional TC-Fluc yield of ~60% when fed at 75 min (Fig. 3I). When creatine phosphate is combined with 2 mM magnesium acetate, the yield increased to ~2-fold of that of the control at maximum (Fig. 3J), although higher magnesium acetate concentrations (4.8 mM) did not significantly change the yield observed (Fig. S3B).

Based on these observations on the effects of individual components and time points, we replenished the reaction at the feeding time 75 min with a combination of AA, tRNAs, DTT, 10-CHO-THF, creatine phosphate and magnesium acetate in 3 different concentration sets and monitored the yield of functional TC-Fluc at regular intervals thereafter by luciferase assay. While TC-Fluc synthesis in the control reaction without feeding stopped at 90 min, the replenished reactions continued for an additional 30 min. More importantly a ~2-fold increase in the maximum yield of functional TC-Fluc compared with the control reaction, was reached at 120 min when 0.3 mM AA, 27 OD260/ml tRNAs, 0.1 mM DTT, 10 μg/ml 10-CHO-THF, 20 mM creatine phosphate and 2 mM magnesium acetate were fed (Fig. S4).

**Supplementing PURE system with missing protein factors isolated from in vivo**

Some additional beneficial factors present in vivo and in cell lysates but not in the PURE system also contribute to the sustainability of protein synthesis. Structural and functional studies suggested that EF-P promotes formation of the first peptide bond during proteins synthesis.42 More recent studies showed that EF-P could also act during elongation by facilitating the synthesis of proteins containing stretches of consecutive proline residues.33,34 To test if EF-P would improve translation efficiencies also on templates without any poly-proline stretches, we supplemented EF-P to TC-Fluc synthesis in the PURE system. Indeed, 4 μM EF-P (optimized by a pilot concentration gradient experiment, Fig. S5), increased the yield of the functional TC-Fluc translation -measured by luciferase - as much as ~2-fold (Fig. 4A). Interestingly, the reaction with EF-P produced, first, significantly less functional TC-Fluc before 30 min, about equal amount of functional TC-Fluc at 30 min, and more functional TC-Fluc after 30 min than the reaction without EF-P, eventually resulting in ~2-fold increase in activity by 120 min (Fig. 4A). Paradoxically, the yield of full-length TC-Fluc determined by TC-tag labeling assay was lower in the presence of EF-P at all time points before 120 min and got close to the yield of the reaction with no EF-P at 120 min (Fig. 4B). Although EF-P slows down translation by either slowing initiation or elongation rates, it is conceivable that a slower translation rate allows nascent peptide to fold properly and efficiently, therefore resulting in a larger fraction of functional product.43 EF4 recognizes ribosomes after a defective translocation reaction and induces a back-translocation, thus giving EF-G a second chance to translocate the tRNAs correctly.44
protein synthesis in living cells \(^{35}\) and the cell extract system even with elevated Mg\(^{2+}\) concentrations \(^{44}\) as well as in the classical PURE system. \(^{27}\) Consistent with previous results, EF4 rescued high Mg\(^{2+}\) caused stress in a more defined system such as PURE (Fig S6): EF4 (maximally with a EF4:ribosome molar ratio of 0.7:1) doubled the translation efficiency of functional Fluc produced in the presence of high Mg\(^{2+}\) (+7 mM) stress (Fig. S6A). Similarly, the amount of full-length HaloTag AcGFP (HT-AcGFP) produced in the PURE system significantly decreased when magnesium acetate concentration increased by 7mM. Adding EF4 with a EF4:ribosome molar ratio of 0.7:1 partially rescued the yields of full-length and partially translated HT-AcGFP (Fig. S6B). Yield quantification in number of pixels was provided in Table S2.

In addition to the tmRNA-dependent trans-translation system, 2 other ribosome rescue systems YaeJ and ArfA/RF2 can release stalled ribosomes from mRNAs lacking in-frame stop codons. However, previous studies only showed YaeJ and ArfA/RF2 performed this function when a short peptide of 7 amino acids was synthesized \(^{30}\) and role of these 2 systems on long peptides is still unknown. To examine the activities of YaeJ and ArfA/RF2 systems during the translation of a full-length protein in the PURE system, we tested the addition of YaeJ and ArfA (RF2 already present in the PURE system) with 9 nM of 3 DNA templates: linear TC-Fluc with a stop codon; linear TC-Fluc lacking a stop codon; circular pIVEX 2.3d-TC-Fluc. A pilot of titrations of ArfA or YaeJ was first tested with linear TC-Fluc template with no stop codon to find the best performance concentration (Fig. 5SB–C). The addition of ArfA or YaeJ didn’t have any significant effect on the yield of TC-Fluc when linear TC-Fluc template with a stop codon and circular template pIVEX 2.3d-TC-Fluc were used. However, only ArfA improved TC-Fluc yield when linear TC-Fluc template with no stop codon was used while neither YaeJ (Fig. 5A–B) nor PrfH (Fig. S7) showed any effect. Significantly higher amount of TC-Fluc was produced with the circular template than the linear templates (Fig. 5A–B). Quantifying fully and partially translated TC-Fluc with the TC tag showed that a larger fraction of ribosomes stalled on the mRNA transcribed from the linear templates than the circular one when translation was nearly accomplished (Fig. S8). One possible explanation is that the 3’ sequence after stop codon encoding T7 terminator in the circular template helped mRNA form a secondary structure alleviating ribosome stalling, therefore resulting in a higher yield. Slightly higher amount of TC-Fluc was produced when linear TC-Fluc without stop codon was used than with stop codon (Fig. 5A–B). This result is contradictory to Shimizu et al.’s when they used a template encoding 7 amino acids, \(^{30}\) indicating that stop codon might not be the main limiting factor of ribosome recycling in our case. Instead, for a long construct, the mRNA secondary structure seems to also play an important role in ribosome recycling efficiency and the final product yield. This is because a long mRNA with a complicated secondary structure can trap the majority of the ribosomes in the middle of the mRNA, therefore hindering ribosome recycling.

**Figure 5.** Supplementing the PURE system with ribosome rescuing systems. (A) Effect of ArfA presence to the PURE system TC-Fluc synthesis. PURE system was assembled with 9 nM different templates encoding TC-Fluc as indicated, with or without 8 \(\mu\)M ArfA protein. TC-Fluc activity was measured after 2 h incubation at 37°C. Values represent averages and error bars are ± standard deviations, with \(n = 3\). (***: \(p < 0.005\); N.S: not significant) (B) Effect of YaeJ presence to the PURE system TC-Fluc synthesis. PURE system was assembled with different templates encoding TC-Fluc as indicated, with or without 6 \(\mu\)M YaeJ protein. TC-Fluc activity was measured after 2 h incubation at 37°C. Values represent averages and error bars are ± standard deviations, with \(n = 3\). (N.S: not significant).
Discussion

CFPS systems have already been used for several applications such as directed evolution of proteins, site-specific incorporation of non-standard amino acids, and has the potential for studying individual components of protein translation. Toward building a self-replicating system, the utilization of already existing biologic machineries promise to be a sensible strategy. A CFPS environment that consists of the optimized concentrations of proteins, energy source, monomer substrates, cofactors and DNA with desired genes, in theory, could constitute such a minimal, self-replicating dynamic system. Previous models propose that ~41,673 peptide bonds per ribosomes need to be synthesized to produce all the components of a self-replicating system. For such a daunting task, beyond other cell-extract dependent CFPSs, the PURE system is the ideal starting point because PURE allows controlled access to each of these components/conditions without being limited by contaminants such as endo/exo nucleases. In this work, we tend to dissect the limiting factors of PURE protein synthesis by looking into individual components such as small molecule substrates and missing function modules involved in proteins synthesis.

We first compared PURE reaction with RTS 100 E. coli HY S30 reaction in terms of transcription and translation. Although PURE resulted full-length protein yields up to ~6 times more than the RTS system; the formation of the functional protein by RTS was up to ~1.5-fold higher (Fig. 1C–D), which suggests presence of additional factors in S30 perhaps improving translation fidelity and facilitating proper protein folding and/or post-translational modifications. Furthermore, while a comparison of the expression profiles of reporter proteins in these 2 systems clearly indicated non-specific RNA degradation in RTS 100 E. coli HY S30 system (probably due to nuclease contamination) that was absent in PURE (Fig. 1A–B). However, PURE too showed a substantial, characteristic smear of partially translated reporter proteins suggestive of premature protein translation termination due to the lack of rescue factors and mRNA chaperones that appear to be present in S30 (Fig. 1C–D). We tested the expression of different HaloTag reporter proteins with sizes ranging 82 kDa – 224 kDa in the PURE and RTS 100 E. coli HY S30 systems. We found yields of full-length protein tend to decline as protein size increased and both of the 2 systems showed low ribosome recycling efficiency (Fig. 2).

The lower yields of functional proteins with classical PURE are likely a result of low translation efficiency of the PURE ribosome and misfolding of the nascent polypeptides, suggesting that PURE might be lacking several protein components such as chaperones, translation facilitating factors and efficient recycling modules resulting in the depletion of energy and monomers or accumulation of inhibitory byproducts.

Expression profile of these proteins in PURE over time confirmed that the new protein synthesis stopped ~2 h (Fig. S1). Interestingly, individual fed-batch replenishment of the PURE components at later time points (~75 min), in general, resulted in higher reaction yields within a given time point before the stop of translation (90 min) (Fig. 3) suggesting that this treatment can boost the rate of protein translation. Indeed, replenishment with a combination of these components at later points not only extended the reaction duration to about ~120 min with a ~2-fold yield increase, but also boosted the rate of functional protein translation (Fig. S4). This increase in yields and rates could be explained by countering substrate/energy depletion and/or time-dependent degradation of some of the components (e.g. DTT).

When the system was replenished with some components in earlier time points, this boost was not seen, probably because of a variety of inhibition modes by the increased substrate concentrations. For example, higher NTP concentrations at earlier time points were inhibitory (Fig. 3G) likely because of the chelation of the necessary Mg$^{2+}$ cations by excess NTPs. Indeed, supplying the system with magnesium acetate had a beneficial effect on yields at any time-point (Fig. 3F), probably addressing Mg$^{2+}$ cation chelation by the accumulating inorganic phosphate created by the creatine phosphate shuttle.

The PURE system still lacks some crucial recycling components and chaperone systems relaxing mRNA secondary structures. A close inspection of other factors known to affect protein translation but absent in the classical PURE system, brought up EF-P as a candidate which is known to reverse protein translation stalling on poly-proline stretches. Indeed, EF-P increased the yield of functional protein 2-fold at the latest time points (~120 min) even though the full-length protein yields were lower at all time-points.
(Fig. 4). This is consistent with EF-P being known to slow down protein translation at the stage of initiation\textsuperscript{48} and/or elongation, which might facilitate a more efficient folding of the nascent peptide into a functional protein as well as by lowering premature protein termination events. EF4, just like in the cell extract system\textsuperscript{44} and living cells,\textsuperscript{35} boosts protein translation not only in the classical PURE system\textsuperscript{27} but also in the PURE system with an elevated Mg\textsuperscript{2+} concentration (Fig. S6). Unlike translating short peptides, YaeJ or PrfH didn’t help on ribosome recycling when translating the large protein TC-Fluc with a linear DNA template with no stop codon, although ArfA still performed the function in this case (Fig. 5 and Fig. S7). It indicates other factors may be needed by YaeJ or PrfH to recycle ribosomes when translating large proteins. Besides, our observation also suggests mRNA secondary structure is another factor leading to ribosomes stalling and therefore influences full-length product yield. Adding mRNA chaperones could possibly increase system productivity by reducing ribosome stalling.

Toward a CFPS-based self-replication system, utilization of PURE system is particularly attractive due to its modularity.\textsuperscript{51} Achieving the theoretical translation efficiencies required for a self-replicated system with PURE will likely require inclusion of additional proteins/modules that may be inherently crucial for protein translation in cells. Indeed, here we show that this optimization process can involve substrate replenishment, by-product recycling and addition of new protein factors. At the same time, this process can also provide us with new insights and facilitate us to test hypotheses about potentially important but poorly characterized players in protein translation, as it did here about EF-P, EF4, ArfA, YaeJ and PrfH.

Materials and methods

Media, chemicals, reagents and device

Unless specified, all chemicals were obtained from Sigma-Aldrich. For protein purification, liquid cultures of strains were grown in either SB media (24 g/L tryptone, 12 g/L yeast extract, 5 g/L glucose, 2 g/L NaH\textsubscript{2}PO\textsubscript{4}, 16.4 g/L K\textsubscript{2}HPO\textsubscript{4}, 4 mL/L glycerol) or 2YT media (16 g/L Tryptone, 10 g/L yeast extract, 5 g/L NaCl). Purified protein concentrations were determined by standard Bradford assay (Bio-Rad), Phusion High-Fidelity PCR Master Mix, restriction enzymes, Quick Ligation Kit and PURE system were purchased from NEB. RTS 100 E.coli HY Kit (S30 system) was purchased from SPRIME. The Pierce 3.5K MWCO Microdialysis Plate was obtained from Thermo Fisher Scientific (Rockford, IL).

Molecular cloning and protein purification

The Fluc gene and TC-Fluc was cloned into the NcoI and XhoI restriction sites of pHVE-2.3d (5 PRIME). Linear templates encoding Fluc or TC-Fluc starting from T7 promoter and ending right before or at stop codon were generated by PCR for further use in PURE system reactions. Cloning of plasmids encoding HT-Ch-TolA, HT-EntE-Ch-TolA and HT-EntF-Ch-TolA was described in\textsuperscript{27} The vector pLJ for expressing N-terminal HaloTagged AcGFP was derived from pFN18K (Promega). Specifically AcGFP was inserted after the sequences of an N-terminal HaloTag domain. pHVEX 2.3d-Fluc, pHVEX 2.3d-TC-Fluc, pLJ, linear Fluc and TC-Fluc templates were purified with phenol: chloroform: isoamyl alcohol extraction and ethanol precipitation before being used for in vitro transcription and translation.

EF4 was cloned, overexpressed and purified as described in\textsuperscript{27} The genes encoding YaeJ and ArfA were cloned from E. coli MG1655 genomic DNA (ATCC) with a C-terminal histag and inserted into pET-24b (Merck Millipore) via NdeI and XhoI. The 12 C-terminal amino acid residues of ArfA were deleted because these residues lead to poor expression and their deletion does not affect ArfA activity.\textsuperscript{30,49} PET-24b YaeJ and ArfA were transformed to BL21 (DE3) strain (NEB). Cells were grown in LB with 50 µg/mL kanamycin at 37 °C, induced with 1 mM IPTG when OD600 reached 0.6, and further incubated at 37 °C to an A600 of 1.4 –1.8. Cells were recovered by centrifugation and pellets were washed by binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 40 mM imidazole, 6 mM β-ME pH 7.4) 3 times and lysed by Bug-Buster Protein Extraction Reagent (Novagen) with 100 µL Halt-protease inhibitor (Thermo-Fisher), 6 mM β-ME in binding buffer. Cell lysates were centrifuged at 15,000 g for 25 min. The supernatant was applied to GE histrap-Gravity column and washed with binding buffer. His-tagged enzymes were eluted by elution buffer (20 mM
sodium phosphate, 0.5 M NaCl, 500 mM imidazole, 6 mM β-ME pH 7.4), pooled, concentrated with Amicon-Ultra-4 concentrator with 3K MWCO, dialyzed to a buffer containing 50 mM Tris–HCl pH 7.6, 100 mM KCl and 1 mM DTT for 6 h twice and then a storage buffer containing 50 mM Tris with 30% glycerol. Purified YaeJ and ArfA were stored at -80°C.

For EF-P preparation, plasmid pST39-his-efp-yjeA-yheK, which encoding EF-P and its modification enzymes was kindly provided by Dr. Park Myung Hee at NIH. Overexpression and purification of EF-P was performed by following the protocol from.50 Specifically, the plasmid was transformed into BL21 (DE3) pLysS strain (Agilent) and grown in 5 mL of 2YT media with 100 μg/mL of Carbenicillin overnight. 250 mL media were inoculated with the overnight cultures and protein expression was induced with 1 mM of IPTG at OD600 around 0.6–0.8. The cells were incubated at shaker at 20°C for 16 h and harvest by centrifuge. Cell pellet was then resuspended in Buffer A (50 mM Hepes-KOH pH 7.6, 1 M NH₄Cl, 10 mM MgCl₂ and 7 mM β-ME) and lysed with french pressure cell. The protein purification was performed with ÄKTAlprime (GE Healthcare) equipped with 5 mL HisTrap™ HP column (GE Healthcare). EF-P was eluted into Buffer A plus 200 mM Imidazole and dialyzed into buffer C (50 mM Hepes-KOH pH 7.6, 100 mM KCl, 10 mM MgCl₂ and 7 mM β-ME). EF-P was concentrated to 0.5 mL and stored in -80°C upon adding one volume of 80 % glycerol.

PrfH gene was cloned from E. coli MG1655 genome into pGEX-6p-1 vector and transformed into BL21 (DE3) pLysS strain. A single colony was inoculated and grown in 5 mL of 2YT media with 100 μg/mL of Carbenicillin overnight. 50 mL media were inoculated with the overnight cultures and protein expression was induced with 1 mM of IPTG at room temperature for 3 h before harvesting via centrifugation. Protein purification was performed with Glutathione Sepharose 4B according to manufacturer instruction, and cleaved off the sepharose beads with PreScission Protease at 4°C in cleavage buffer containing 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 1% Triton X-100. PreScission Protease cleavage generated a single protein product corresponding to PrfH (with no affinity tags) based on Coomassie Blue staining.

**Total RNA extraction and electrophoresis**

15 μl of PURE system and 15 μl of S30 system (RTS 100 E.coli HY Kit) reactions were set up according to manufacturer’s instructions with 10 ng/μl (3 nM) TC-Fluc plasmid. PURE reactions were incubated at 37°C for 2 h and S30 reactions were incubated at 30°C for 6 h. Then 10 μl of each reaction was taken for luciferase assay; 1 μl was taken for labeling assay; 2 μl was taken for total RNA extraction. Specifically, 2 μl reaction was mixed with 200 μl buffer TE. Then 200 μl Acid-Phenol:Chloroform, 5:1, pH 4.5 (premixed with IAA (125:24:1), Ambion Cat. No. 9722) was added to the mixture and vortexed for 15 s. Upper aqueous phase was taken to mix with 200 μl Chloroform. Aqueous phase was mixed with 1/10 volume of 3 M Sodium Acetate and 3X volume of ice-cold ethanol to precipitate the RNA. Extracted total RNAs were resuspended with buffer TE. Concentrations of RNAs were measured by Nanodrop. 1μg of each RNA sample was denatured and subjected to electrophoresis in 1% agarose gel, followed by visualization with SYBR Green dye (Life Technologies).

**Luciferase assay**

Fluc activity was measured by Promega Luciferase Assay System kit. Relative luminescence units were read by SpectraMaxM5 plate reader (Molecular Devices, Sunnyvale, CA). The amount of functional TC-Fluc produced was calculated by converting relative luminescence units to the actual amount of TC-Fluc via a premeasured standard curve.

**Labeling of in vitro translated proteins**

To further examine full-length and incomplete products of *in vitro* translation in PURE system, HT-Ch-TolA, HT-EntE-Ch-TolA, HT-EntF-Ch-TolA and HT-AcGFP were *in vitro* translated and labeled with a fluorescent TMR reporter (Promega); TC-Fluc was *in vitro* translated and labeled with FlAsH-EDT₂ biarsenical labeling reagent (Life Technologies). Specifically, reactions expressing HT tagged proteins were diluted by 10 times with PBS and incubated with 5 μM TMR ligand at room temperature for 30 min before SDS-
PAGE analysis. Gels were scanned at 580 nm (excitation at 532 nm) by Typhoon Trio Imager (GE Healthcare) and a fluorescence gel image was analyzed with ImageQuant TL software. For labeling of TC-Fluc, PURE system and S30 system reactions were first diluted 1:10 in PBS and mixed with 1.25 μM FlAsH-EDT₂ biarsenical labeling reagent (diluted in PBS). The mixture was incubated at 25 °C for 30 min, and then subjected to SDS-PAGE analysis. Fluorescent signal was obtained with Typhoon Trio Imager (GE Healthcare) at 528 nm (excitation at 508 nm).

**PURE fed-batch reactions**

PURE batch reactions (10 μL) synthesising TC-Fluc were performed following the manufacture’s instructions. The volumes of water or substrates (individual and in combination, prepared to proper stock concentrations to give desired replenishment) added to PURE were set to 1 μL to avoid dilution effect. We first set up a series of 10 μL PURE reactions and added 1 μL of water at time 0 min, 15 min, 30 min, 45 min, 60 min or 75 min to the reactions. TC-Fluc was quantified at 90 min by luciferase assay to make sure these testing reactions produced similar amount of TC-Fluc (data not shown). Therefore adding 1 μL water at time 0 min can be used as control for the fed-batch tests of substrates. An equivalent volume of feed substrates were added directly into the reactions at given time points to replenish the desired substrate at concentrations specified in the text.

**Inorganic phosphate quantitative assay**

Quantitative analysis of inorganic phosphate was performed using the EnzChek Phosphate Assay Kit (Life Technologies). PURE samples were first quenched by frozen on liquid nitrogen and later thawed and assayed. To measure phosphate concentration, samples were diluted 200-fold with MQ water and compared with a linear phosphate standard ranging from 2 to 150 μM. Absorbance was read at 360 nm following the manufacture’s instructions.

**PURE transcription and translation with translation factors**

Five μL reactions were performed following the PURE system manual with 2 μL solution A, 1.5 μL solution B, 0.8 U/μL SUPERase IN RNase Inhibitor (Life Technologies), circular or linear DNA template with concentration indicated in the text, and proper amount of supplement factors as indicated in the text. Reactions were incubated at 37 °C for 2 h unless otherwise stated. For analyzing the effect of EF-P on TC-Fluc activity at different time points, samples for prior time points were stored at 4 °C till the last time point was reached, and TC-Fluc activity and/or fluorescent labeling was analyzed for all samples together.

**Disclosure of potential conflicts of interest**

No potential conflicts of interest were disclosed.

**Acknowledgements**

We thank Dr. Anthony Forster for comments and suggestions.

**Funding**

This work was funded by programs from the Department of Energy (Genomes to Life Center) [Grant #DE-FG02-02ER63445] and a gift funding from Origins of Life Initiative.

**Author contributions**

JL designed the study. JL, CZ, PH, TL and EFB performed the experiments. JL and CZ analyzed the data. JL and EK wrote the manuscript. All authors reviewed the manuscript. GMC performed a supervision role.

**ORCID**

Taibo Li http://orcid.org/0000-0002-6624-9293

**References**

1. Villemagne D, Jackson R & Douthwaite JA Highly efficient ribosome display selection by use of purified components for in vitro translation. J Immunol Methods 2006; 313:140-8; PMID:16730021; https://doi.org/10.1016/j.jim.2006.04.001
2. Ohashi H, Shimizu Y, Ying W & Ueda T Efficient protein selection based on ribosome display system with purified components. Biochem Biophys Res Commun 2007; 352:270-6; PMID:17113037; https://doi.org/10.1016/j.bbrc.2006.11.017
3. Josephson K, Hartman MC & Szostak JW Ribosomal synthesis of unnatural peptides. J Am Chem Soc 2005; 127:11727-35; PMID:16104750; https://doi.org/10.1021/ja0515809
4. Forster AC, Tan Z, Nalam MN, Lin H, Qu H, Cornish VW, Blacklow SC. Programming peptidomimetic syntheses by translating genetic codes designed de novo. Proc...
New England BioLabs Inc [Internet]. Advantages of the

Jewett MC, Forster AC. Update on designing and

Towards synthesis of a minimal cell. Methods Enzymol 1983; 101:635-44; PMID:6888279

Anastasina M, Terenin I, Butcher SJ, Kainov DE. A techni-que to increase protein yield in a rabbit reticulocyte lysate translation system. Biotechniques 2014; 56:36-9; PMID:24447137; https://doi.org/10.2144/000114125

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T. Cell-free translation reconstituted with purified components. Nat Biotechnol 2001; 19:751-5; PMID:11479568; https://doi.org/10.1038/90802

Zubay G. In vitro synthesis of protein in microbial systems. Annu Rev Genet 1973; 7:267-87; PMID:4593305; https://doi.org/10.1146/annurev.ge.07.120173.001411

Wang HH, Huang PY, Xu G, Haas W, Marblestone A, Li J, Gygi SP, Forster AC, Jewett MC, Church GM. Multiplexed in vivo His-tagging of enzyme pathways for in vitro single-pot multienzyme catalysis. ACS Synth Biol 2012; 1, 43-52; PMID:22737598; https://doi.org/10.1021/sb3000029

Forster AC, Church GM. Towards synthesis of a minimal cell. Mol Syst Biol 2006; 2:45; PMID:16924266; https://doi.org/10.1038/msb4100090

Jewett MC, Forster AC. Update on designing and building minimal cells. Curr Opin Biotechnol 2010; 21:697-703; PMID:20638265; https://doi.org/10.1016/j.copbio.2010.06.008

New England BioLabs Inc [Internet]. Advantages of the PURE System. Available from: https://www.neb.com/products/protein-expression-and-purification-technologies/cell-free-expression/cell-free-expression/advantages-of-the-pure-system

5 PRIME [Internet]. RTS 100 E. coli HY Kit Manual. [2009]. Available from: http://wolfson.huji.ac.il/expression/products/in-vitro/RTS100Ecoli manuals.pdf

Jackson K, Kanamori T, Ueda T, Fan ZH. Protein synthesis yield increased 72 times in the cell-free PURE system. Integr Biol (Camb) 2014; 6, 781-8; PMID:25008400; https://doi.org/10.1039/C4IB00088A

Carlson ED, Gan R, Hodgman CE, Jewett MC. Cell-free protein synthesis: applications come of age. Biotechnol Adv 2014; 30:1185-94; https://doi.org/10.1016/j.biotechadv.2011.09.016

Kim DM, Swartz JR. Prolonging cell-free protein synthesis with a novel ATP regeneration system. Biotechnol Bioeng 1999; 66:180-8; PMID:10577472; https://doi.org/10.1002/(SICI)1097-0290(1999)66:3%3c180::AID-BIT6%3e3.0.CO;2-S

Kim DM, Swartz JR. Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol Prog 2000; 16:385-90; PMID:10835240; https://doi.org/10.1021/bp000031y

Jewett MC, Swartz JR. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol Bioeng 2004; 86:19-26; PMID:15007837; https://doi.org/10.1002/bit.20026

Calhoun KA, Swartz JR. Energizing cell-free protein synthesis with glucose metabolism. Biotechnol Bioeng 2005; 90:606-13; PMID:15830344; https://doi.org/10.1002/bit.20449

Jewett MC, Swartz JR. Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. Biotechnol Bioeng 2004; 87:465-72; PMID:15286983; https://doi.org/10.1002/bit.20139

Michel-Reydellet N, Calhoun K & Swartz J Amino acid stabilization for cell-free protein synthesis by modification of the Escherichia coli genome. Metab Eng 2004; 6:197-203; PMID:15256209; https://doi.org/10.1016/j.ymben.2004.01.003

Nakano H, Shinbata T, Okumura R, Sekiguchi S, Fujishiro M, Yamane T. Efficient coupled transcription/translation from PCR template by a hollow-fiber membrane bioreactor. Biotechnol Bioeng 1999; 64, 194-9; PMID:10397855; https://doi.org/10.1002/(SICI)1097-0290(19990720)64:2%3c194::AID-BIT8%3e3.0.CO;2-5

Shimizu Y, Kanamori T, Ueda T. Protein synthesis by pure translation systems. Methods 2005; 36:299-304; PMID:16076456; https://doi.org/10.1016/j.jmeth.2005.04.006

Hillebrecht JR, Chong S. A comparative study of protein synthesis in in vitro systems: from the prokaryotic reconstituted to the eukaryotic extract-based. BMC Biotechnol 2008; 8:58; PMID:18664286; https://doi.org/10.1186/1472-6750-8-58

Li J, Gu L, Aach J, Church GM. Improved cell-free RNA and protein synthesis system. PLoS One 2014; 9: e106232; PMID:25180701; https://doi.org/10.1371/journal.pone.0106232

Ramachandiran V, Kramer G, Hardesty B. Expression of different coding sequences in cell-free bacterial and eukaryotic systems indicates translational pausing on Escherichia coli ribosomes. FEBS Lett 2000; 482:185-8; PMID:11479568; https://doi.org/10.1016/S0014-5793(00)10001-5

Haebel PW, Gutmann S, Ban N. Dial tm for rescue: tmRNA engages ribosomes stalled on defective mRNAs. Curr Opin Struct Biol 2004; 14:58-65; PMID:15102450; https://doi.org/10.1016/j.sbi.2004.01.010

Shimizu Y. ArfA recruits RF2 into stalled ribosomes. J Mol Biol 2012; 423:624-31; PMID:22922063; https://doi.org/10.1016/j.jmb.2012.08.007
[31] Chadani Y, Ito K, Kutsukake K & Abo T ArfA recruits release factor 2 to rescue stalled ribosomes by peptidyl-tRNA hydrolysis in Escherichia coli. Mol Microbiol 2012; 86:37-50; PMID:22857598; https://doi.org/10.1111/j.1365-2958.2012.01810.x

[32] Chadani Y, Ono K, Kutsukake K, Abo T. Escherichia coli YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. Mol Microbiol 2011; 80:772-85; PMID:21418110; https://doi.org/10.1111/j.1365-2958.2011.07607.x

[33] Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, Rodnina MV. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science 2013; 339, 85-8; PMID:23239624; https://doi.org/10.1126/science.1229017

[34] Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science 2013; 339:82-5; PMID:23239623; https://doi.org/10.1126/science.1228985

[35] Pech M, Karim Z, Yamamoto H, Kitakawa M, Qin Y, Nierhaus KH. Elongation factor 4 (EF4/LepA) accelerates protein synthesis at increased Mg2+ concentrations. Proc Natl Acad Sci U S A 2011; 108:3199-203; PMID:21300907; https://doi.org/10.1073/pnas.1012994108

[36] Baranov PV, Vestergaard B, Hamelryck T, Gesteland RF, Nyborg J, Atkins JP. Diverse bacterial genomes encode an operon of two genes, one of which is an unusual class-I release factor that potentially recognizes atypical mRNA signals other than normal stop codons. Biol Direct 2006; 1:28; PMID:16970810; https://doi.org/10.1186/1745-6150-1-28

[37] Jewett MC, Fritz BR, Timmerman LE, Church GM. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. Mol Syst Biol 2013; 9:678; PMID:23799452; https://doi.org/10.1038/msb.2013.31

[38] Purvis IJ, Betanny AJ, Santiago TC, Coggins JR, Duncan K, Eason R, Brown AJ. The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis. J Mol Biol 1987; 193:413-7; PMID:21062370; https://doi.org/10.1016/0022-2836(87)90230-0

[39] Du L, Gao R, Forster AC. Engineering multigene expression *in vitro* and in vivo with small terminators for T7 RNA polymerase. Biotechnol Bioeng 2009; 104:1189-96; PMID:19650080; https://doi.org/10.1002/bit.22491

[40] Carlson ED, Gan R, Hodgman CE, Jewett MC. Cell-free protein synthesis: applications come of age. Biotechnol Adv 2012; 30:1185-94; PMID:22008973; https://doi.org/10.1016/j.biotechadv.2011.09.016

[41] Schoborg JA, Hodgman CE, Anderson MJ, Jewett MC. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. Biotechnol J 2014; 9630-40; PMID:24323955; https://doi.org/10.1002/biot.201300383

[42] Blaha G, Stanley RE, Steitz TA. Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science 2009; 325:966-70; PMID:19696344; https://doi.org/10.1126/science.1175800

[43] Zhang G, Hubalewska M, Ignatova Z. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat Struct Mol Biol 2009; 16:274-80; PMID:19198590; https://doi.org/10.1038/nsmb.1554

[44] Qin Y, Polacek N, Vesper O, Staub E, Einfeldt E, Wilson DN, Nierhaus KH. The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. Cell 2006; 127:721-33; PMID:17101332; https://doi.org/10.1016/j.cell.2006.09.037

[45] Kazuta Y, Matsuura T, Ichihashi N, Yomo T. Synthesis of milligram quantities of proteins using a reconstituted *in vitro* protein synthesis system. J Biosci Bioeng 2014; 118:554-7; PMID:24880499; https://doi.org/10.1016/j.jbiosc.2014.04.019

[46] Zea CJ, Camci-Unal G, Pohl NL. Thermodynamics of binding of divalent magnesium and manganese to uridine phosphates: implications for diabetes-related hypoglycaemia and carbohydrate biocatalysis. Chem Cent J 2008; 2:15; PMID:18627619; https://doi.org/10.1186/1752-153X-2-15

[47] Kazuta Y, Adachi J, Matsuura T, Ono N, Mori H, Yomo T. Comprehensive analysis of the effects of Escherichia coli ORFs on protein translation reaction. Mol Cell Proteomics 2008; 7:1530-40; PMID:18453339; https://doi.org/10.1074/mcp.M800051-MCP200

[48] Hersch SJ, Elgamal S, Katz A, Iba M & Navarre WW. Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. J Biol Chem 2014 289:28160-71; PMID:25148683; https://doi.org/10.1074/jbc.M114.593277

[49] Chadani Y, Ono K, Ozawa S, Takahashi Y, Takai K, Nanamiya H, Tozawa Y, Kutsukake K, Abo T. Ribosome rescue by Escherichia coli ArfA (YhdL) in the absence of YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. Mol Syst Biol 2011; 7:808; PMID:21062370; https://doi.org/10.1016/j.molsmb.2011.05.004

[50] Park JH, Johansson HE, Aoki H, Huang BX, Kim HY, Ganoza MC, Park MH. Post-translational modification by beta-lysylation is required for activity of Escherichia coli ORF-P (EF-P). J Biol Chem 2012; 287:2579-90; PMID:22128152; https://doi.org/10.1074/jbc.M112.359670

[51] Li J, Haas W, Jackson K, Kuru E, Jewett MC, Fan HZ, Gygi SP, Church GM. Cogenerating Synthetic Parts toward a Self-Replicating System. ACS Synth Biol 2017; 1-32. Advance online publication. Available from: http://arep.med.harvard.edu/pdf/Li_Church_ACS_2017.pdf