

# Optimizing Protein Production in the One-Pot PURE System: Insights into Reaction Composition and Expression Efficiency

Yan Zhang,\* Matas Deveikis, Yanping Qiu, Lovisa Björn, Zachary A. Martinez, Tsui-Fen Chou, Paul S. Freemont, and Richard M. Murray



Cite This: ACS Synth. Biol. 2025, 14, 1496–1508



Read Online

ACCESS |

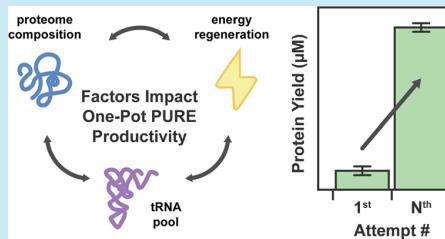
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** The One-Pot PURE (Protein synthesis Using Recombinant Elements) system simplifies the preparation of traditional PURE systems by coculturing and purifying 36 essential proteins for gene expression in a single step, enhancing accessibility and affordability for widespread laboratory adoption and customization. However, replicating this protocol to match the productivity of traditional PURE systems can take considerable time and effort due to uncharacterized variability. In this work, we observed unstable PURE protein expression in the original One-Pot PURE strains, *E. coli* M15/pREP4 and BL21(DE3), and addressed this issue using glucose-mediated catabolite repression to minimize burdensome background expression. We also identified several limitations making the M15/pREP4 strain unsuitable for PURE protein expression, including coculture incompatibility with BL21(DE3) and uncharacterized proteolytic activity. We showed that consolidating all expression vectors into a protease-deficient BL21(DE3) strain minimized proteolysis, led to more uniform coculture cell growth at the time of induction, and improved the stoichiometry of critical translation initiation factors in the final PURE mixture for efficient cell-free protein production. In addition to optimizing the One-Pot PURE protein composition, we found that variations in commercial energy solution formulations could compensate for suboptimal PURE protein stoichiometry. Notably, altering the source of *E. coli* tRNAs in the energy solution alone led to significant differences in the expression capacity of cell-free reactions, highlighting the importance of tRNA codon usage in influencing protein expression yield. Taken together, this work systematically investigates the proteome and biochemical factors influencing the One-Pot PURE system productivity, offering insights to enhance its robustness and adaptability across laboratories.

**KEYWORDS:** PURE system, cell-free protein system, reconstituted *in vitro* transcription-translation, codon pool optimization, cell-free proteome optimization, protein synthesis tools



## INTRODUCTION

Cell-free protein expression harnesses the core gene expression machinery in living cells to enable transcription and translation in a test tube reaction. The open nature of these expression platforms allows for direct manipulation of the reaction environment. Applications of these platforms span high-throughput screening,<sup>1–3</sup> novel protein modifications,<sup>4,5</sup> interfacing biomolecules with synthetic materials,<sup>6–9</sup> on-demand biosensing<sup>10–12</sup> and biomanufacturing,<sup>13,14</sup> and understanding the rules of life by building biological systems from scratch.<sup>15–18</sup>

Nearly all cell-free protein expression platforms can be categorized into two classes—crude lysate-based and reconstituted systems. Crude lysate-based expression systems use extracts from lysed cells, which retain most of the cell's proteome and metabolic pathways.<sup>19,20</sup> They can support gene expression from a wide range of promoters and enable higher protein expression due to effective energy regeneration.<sup>21</sup> They can also be exceptionally affordable, costing approximately \$0.03 per microliter of reaction.<sup>22</sup> However, the active endogenous metabolism and partially intact proteome can

divert critical energy sources and metabolites to side reactions, interfering with the central reactions implemented in these systems.<sup>20</sup>

In contrast, the PURE (Protein synthesis Using Recombinant Elements) system isolates the necessary transcription, translation, and energy regeneration proteins to support gene expression by protein purification and reconstitution. This approach creates a more biochemically defined reaction environment, which enables the construction of predictive models for protein expression,<sup>23,24</sup> allows facile removal and replacement of reaction enzymes for customization,<sup>5</sup> and maintains a minimal reaction proteome toward a self-regenerating synthetic cell.<sup>16</sup> However, traditional PURE systems are expensive to purchase (\$0.64–1.00 per microliter

Received: November 9, 2024

Revised: March 25, 2025

Accepted: March 25, 2025

Published: April 10, 2025



of reaction) and laborious to prepare in-house, posing a significant bottleneck in the broad adoption of PURE systems.

To address these limitations, Lavickova and Maerkel developed the One-Pot PURE system, which offers a simplified alternative.<sup>25</sup> Instead of growing and purifying each of the 36 PURE proteins, they grew a coculture of the 36 *E. coli* strains, each expressing one of the PURE proteins, and purified the proteins in a single preparation. This approach significantly streamlines the PURE system preparation and delivers protein expression yields comparable to commercial PURE systems but at a fraction of the cost (\$0.10 per microliter of reaction).

Despite the promise of this method, replicating the One-Pot PURE protocol to achieve productivity comparable to that of conventional PURE systems can still take considerable time and effort. In this work, we identify sources of variability, leading to reduced productivity in One-Pot PURE preparations. These include cell growth and expression instability to obtain PURE proteins, growth incompatibility between the *E. coli* M15/pREP4 and BL21(DE3) strains during coculture which created suboptimal PURE protein stoichiometry for cell-free protein expression, and uncharacterized protease activity in the M15/pREP4 strain during PURE protein purification. Moreover, we demonstrate that variations in the reaction's biochemical composition can also significantly impact the protein production rate and the final protein yield, emphasizing the need for more precise system optimization. Taken together, our findings enhance the current understanding of factors influencing protein expression capacity in the One-Pot PURE system, providing strategies to enable robust and widespread implementation across diverse laboratories.

## RESULTS AND DISCUSSION

**Enhancing Protein Expression Stability for the One-Pot PURE System through Glucose-Mediated Catabolite Repression.** We began by replicating the One-Pot PURE system using plasmids deposited on Addgene and transforming them into the respective *E. coli* expression strains M15/pREP4 and BL21(DE3). Plasmids containing PURE proteins encoded by pQE30 and pQE60 expression vectors were transformed into M15/pREP4 competent cells with the pREP4 plasmid coding for constitutive LacI expression. Plasmids containing PURE proteins encoded by pET21 and pET15 expression vectors were transformed into BL21(DE3) competent cells (Figure 1A).

In our initial assessment of protein expression in monoculture assays, we revealed that several protein expressions were unstable. Despite confirming plasmid sequences through whole plasmid sequencing, we observed a loss of expression in ten proteins after growing and inducing protein expression from saved glycerol stocks (Figure 1B,C). No growth defect was observed for these nonexpressing cell cultures (Figure SI.III 1). This phenomenon was also observed in several expression strains directly obtained from our collaborator, with the loss of protein expression appearing in other strains. This indicates that expression vector instability is a stochastic yet frequent event in PURE protein expression (Figure SI.III 2).

The expression vector instability is problematic because previously validated PURE protein expression strains can spontaneously lose protein expression after being saved to and grown from glycerol stocks. Without addressing this issue, subsequent preparations of the One-Pot PURE system may exhibit protein dropouts, leading to unproductive batches.

Alternatively, each One-Pot PURE preparation would require fresh transformations of the 36 protein expression vectors, which could be labor-intensive and inefficient. Stabilizing the growth and expression of PURE proteins from glycerol stocks could significantly simplify the preparation workflow and facilitate the dissemination of these strains to other laboratories.

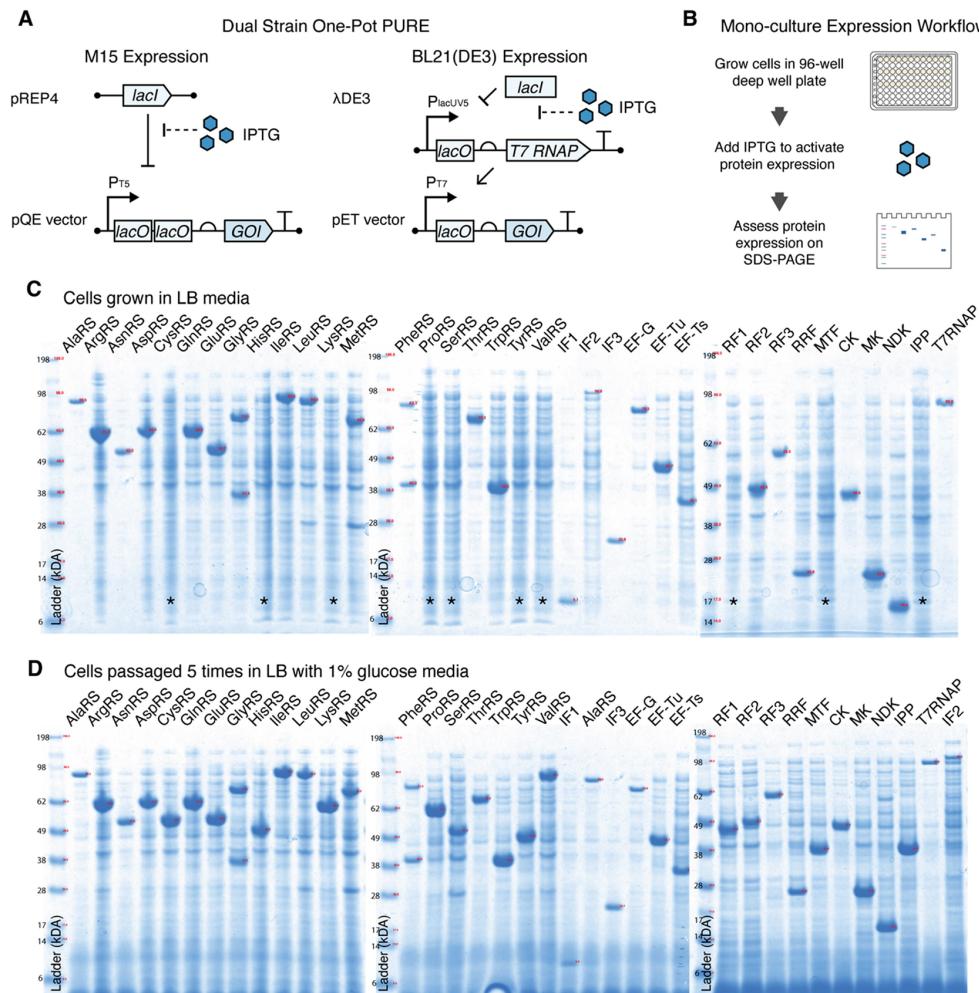
We hypothesize that leaky background protein expression from the P<sub>T7-lacO</sub> promoter in BL21(DE3) and the P<sub>T5-lacO-lacO</sub> promoters in M15/pREP4 strains can create growth burdens, leading to the loss of protein expression. This phenomenon has been observed by others, in which background protein expression burden led to promoter mutations, resulting in the loss of target protein expression.<sup>5</sup> To reduce background protein expression, we investigated whether catabolite repression—through the addition of a preferred carbon source (glucose)—could upregulate the intracellular concentration of repressors (LacI).<sup>26</sup> This upregulation of the LacI repressor could reduce background expression from P<sub>T7-lacO</sub> and P<sub>T5-lacO-lacO</sub> promoters.

Indeed, our results showed that by supplementing growth media with 1% w/v glucose, PURE protein expression was stably maintained even after five serial passages from the seed glycerol stock (Figure 1D). For strains grown in catabolite repression media, we also observed improved cell growth and higher protein expression levels in ten PURE proteins previously found to lose expression ability (Figure SI.III 3). Based on these findings, all subsequent One-Pot PURE batches were grown with LB media supplemented with 1% w/v glucose.

**Offsets in the Dual-Strain One-Pot PURE Protein Composition Contribute to Low Cell-Free Protein Production.** After resolving the instability in protein expression, we proceeded to replicate the original One-Pot PURE system with BL21(DE3) and M15/pREP4 expression strains (referred to as Dual-Strain One-Pot PURE) and the exact energy solution formulation.<sup>27</sup> To assess the system's transcription and translation capacity, we used a reporter plasmid with a malachite green aptamer (MGA) to approximate mRNA concentration immediately upstream of the ribosomal binding site of the green fluorescent protein (deGFP) to measure protein yield (Figure 2A), consistent with established designs.<sup>28,29</sup>

Surprisingly, across two batches of the Dual-Strain One-Pot PURE system we sought to replicate, the protein expression was capped at approximately 1.5  $\mu$ M (Figure 2B), while up to 6  $\mu$ M mRNA transcripts were produced. This protein yield is significantly lower than the yield we observed from commercial PURE systems (4–6  $\mu$ M, Figures 3B and D). We also note that the high initial malachite green mRNA concentration at the start of the reaction may have come from the rapid mRNA transcription that had begun before the microwell plate was placed into the plate reader (Figure 5).

Given that we had confirmed all strains were expressing the One-Pot PURE proteins in monoculture expression assays, we hypothesized that there might be differences in cell growth or protein expression in the coculture conditions not captured by the monoculture expression assays. However, using gel electrophoresis to resolve the purified One-Pot PURE protein mixture, as done in the monoculture assays, would be challenging due to the similarity in molecular weight of many proteins. We, therefore, used a targeted proteomics



**Figure 1.** (A) Schematic of protein expression and regulation in *E. coli* strains M15/pREP4 and BL21(DE3) for Dual-Strain One-Pot PURE systems. Each gene of interest is expressed under the control of either the  $P_{T5}$  or  $P_{T7}$  promoter followed by lacO operator site(s). In the M15/pREP4 expression strain, the pREP4 plasmid constitutively expresses the lac inhibitor LacI, which represses protein expression by binding to the lacO sites. In BL21(DE3), endogenous LacI represses the production of T7 RNA polymerase (RNAP) and gene expression by binding to the lacO sites. Addition of the IPTG inducer derepresses LacI inhibition and activates gene expression. (B) Schematic of the monoculture protein expression workflow. Cells were grown in 96-well deep-well plates, and protein expression was activated by adding a 0.1 mM IPTG inducer. Following cell growth, protein expression was assessed via SDS-PAGE. (C) Monoculture protein expression assessment from cells grown in LB media. Ten strains (designated by an asterisk) were found to carry the protein expression plasmid but lacked the protein expression capability. (D) Stable protein expression from cells grown in LB with 1% w/v glucose after five serial passages with all PURE proteins visibly expressed.

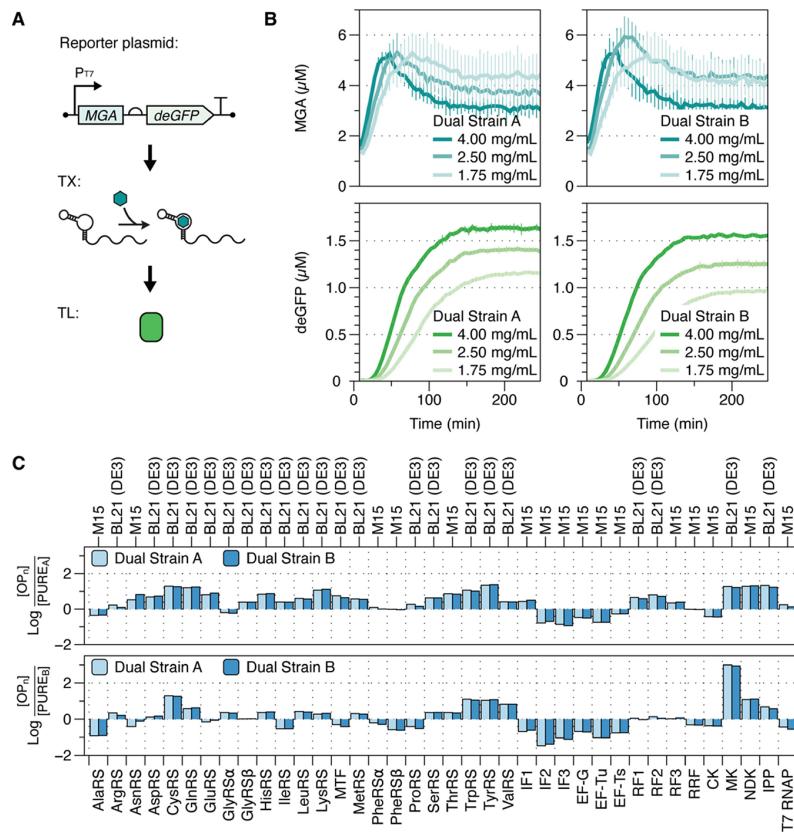
approach to characterize the relative abundance of these proteins.

Using liquid chromatography with tandem mass spectrometry (LC-MS/MS), we mapped the detected peptides to the protein abundances. After scaling for their concentrations in reactions, we compared the relative abundance of individual proteins in the Dual-Strain One-Pot PURE mixture to those in commercial PURE systems (Figure 2C). Our analysis revealed a consistent deficiency in seven proteins in Dual-Strain One-Pot PURE systems compared with the two commercial PURE systems (PURE<sub>A</sub> and PURE<sub>B</sub>). These proteins are alanyl-tRNA synthetase (AlaRS), initiation factor 2 (IF2), initiation factor 3 (IF3), elongation factor-G (EF-G), elongation factor-Tu (EF-Tu), elongation factor-Ts (EF-Ts), and creatine kinase (CK).

**Bypassing Dual-Strain One-Pot Induction Challenge in Coculture with 36-Pot PURE.** Since all seven consistently

deficient proteins were expressed in the *E. coli* M15/pREP4 expression strain, we investigated whether these cells had growth deficiencies. Monoculture growth data revealed that M15/pREP4 cells exhibited slower growth rates than BL21-(DE3) cells, and most of them stopped growing after inducing protein expression (Figure SI.III 1,2). This mismatch in growth rates created an incompatible induction window when coculturing BL21(DE3) and M15/pREP4 cells for protein expression. To address this induction incompatibility, we explored growing the strains independently, followed by pooling them together at harvest for purification in a single preparation. We note here that a similar approach has been pursued as part of an open-source PURE initiative in the synthetic cell community.<sup>30</sup>

This strategy, named the 36-Pot PURE, involved growing and expressing each of the 36 PURE proteins separately while



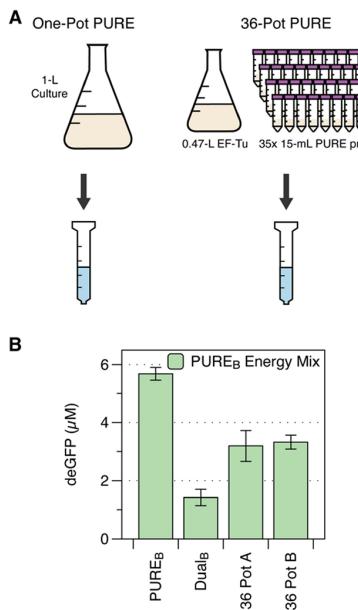
**Figure 2.** (A) Schematic of the reporter plasmid used to assess transcription and translation activity. The plasmid contains an RNA malachite green aptamer (MGA) to assess mRNA transcription and a green fluorescent protein reporter (deGFP) to measure the protein expression. Consistent with established designs, the malachite green aptamer is placed immediately upstream of the ribosomal binding site without an intermediate spacer.<sup>28,29</sup> (B) Productivity of two batches of Dual-Strain One-Pot PURE systems (Dual Strain A and Dual Strain B) with homemade energy buffer in transcription and translation, measured via MGA and deGFP, respectively. The One-Pot PURE protein concentration in the reactions was varied from 1.75 to 4 mg/mL to optimize the cell-free protein yield. Plots represent the averages of reaction triplicates, with error bars indicating the standard deviations of the reaction triplicates. (C) Comparison of protein abundances in two batches of Dual-Strain One-Pot PURE system to those in two commercial PURE systems after scaling by concentrations in the reaction.

maintaining the original inoculum composition for the One-Pot PURE coculture. For a 1 L scale preparation, the original One-Pot PURE method requires a 10 mL inoculum consisting of 47% v/v for elongation factor-thermal unstable (EF-Tu) culture and 1.5% v/v of each of the other proteins from saturated overnight cultures. In 36-Pot PURE, thirty-six cultures were grown: one 470 mL for EF-Tu growth culture and thirty-five 15 mL cultures for all other proteins (Figure 3A). Each culture was inoculated using a 1:100 dilution of a saturated overnight culture, grown to an induction OD above 0.6, induced for protein expression, and incubated for an additional 3 h for protein expression.

Using this approach, the 36-Pot PURE system showed a 2-fold increase in deGFP yield compared to the Dual-Strain One-Pot PURE system (Figure 3B). The improvement may be attributed to the mitigation of growth competition, as growing the strains separately prevented the faster-growing BL21(DE3) strains from outcompeting the M15/pREP4 cells in coculture. The 36-Pot format also provided more flexibility in selecting optimal induction times for each protein, thus improving the final PURE protein composition. However, despite these improvements, the gene expression capacity still does not match the levels achieved by commercial PURE systems.

**M15/pREP4 Expression Strain Exhibits Unreported Proteolytic Activity.** In a separate attempt to purify T7 RNAP from the M15/pREP4 strain, we discovered protease activity in this expression strain that had not been reported previously. We found that, despite adding a broad-spectrum protease inhibitor (cCOMPLETE) in the lysis buffer, proteolysis of T7 RNAP into 80 and 20 kDa fragments was observed after protein purification, suggesting that proteolysis occurred during cell lysis (Figure 4). This protease activity was not detected in earlier protein gel analyses, where cultures were heat-denatured immediately after protein expression (Figure 1C,D).

Since M15/pREP4 is not a protease-deficient strain like BL21(DE3), we hypothesize that outer membrane omptin family proteases present in M15/pREP4 were responsible for cleaving T7 RNAP into truncated fragments. This is consistent with literature reports that omptins recognize and cleave T7 RNAP into 80- and 20-kDa fragments.<sup>31</sup> Because omptins are aspartyl proteases unaffected by common protease inhibitors,<sup>32</sup> no inhibition of protease activity was observed when lysis buffer was supplemented with Pepstatin A, an aspartyl protease inhibitor. This is consistent with reports of weak inhibition of omptins.<sup>33</sup> In contrast, no proteolysis was observed when T7



**Figure 3.** (A) Schematic showing the differences in preparation between the Dual-Strain One-Pot PURE and 36-Pot PURE system. (B) PURE<sub>B</sub> is a commercial PURE reaction assembled according to the manufacturer's protocol. Dual-Strain One-Pot PURE and 36-Pot PURE reactions were assembled with 5 mg/mL PURE proteins. All reactions were prepared using a commercial energy solution (supplier B) and contained 5 nM P<sub>T7</sub>-UTR1-deGFP plasmid. The bar graphs represent the average terminal protein yield of reaction triplicates, with error bars indicating the standard deviations of reaction triplicates.

RNAP was expressed in the BL21/pREP4 strain, which lacks one variant of the omptin family proteases, OmpT (Figure 4A,B).

It is also possible for omptins to cleave other PURE proteins with similar recognition motifs.<sup>34</sup> Using established omptin cleavage motifs,<sup>35</sup> we identified 12 possible omptin cleavage sites in eight PURE proteins (Figure 4C). The susceptible peptides identified for T7 RNAP and creatine kinase (CK) matched literature findings.<sup>34,36</sup> We could not find published results confirming omptin-mediated proteolysis for the other six PURE proteins: myokinase (MK), elongation factor-thermal stable (EF-Ts), and threonyl- (ThrRS), valyl- (ValRS), glycyl- (GlyRS), and isoleucyl-tRNA synthetases (IleRS). Using AlphaFold 3 simulations,<sup>37</sup> we found that all but two peptides—one for ValRS and one for IleRS—are located on the external surface of the proteins and are accessible for protease binding (Figure SI.III 4).

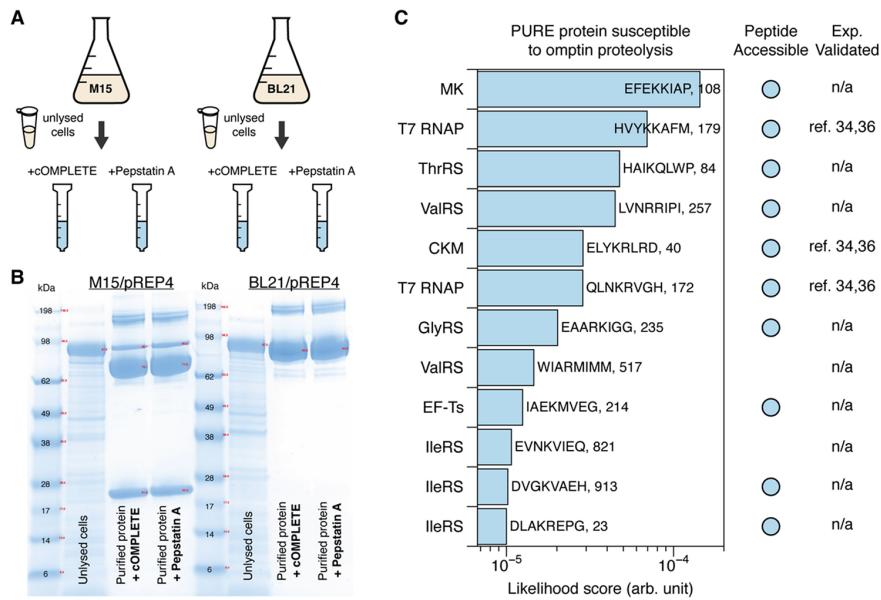
These simulations suggest that proteolysis of PURE proteins caused by omptins may contribute to the lower gene expression capacity observed during our replication of the original Dual-Strain One-Pot PURE system. Since omptins can cleave full-length proteins into truncated fragments, even if these fragments can reassemble and be recovered from affinity column purification, their activities are likely to be reduced, as reported in the case of T7 RNAP.<sup>38</sup> Notably, truncated fragments of T7 RNAP are less efficient in transcription initiation and exhibit a higher frequency of premature transcription termination.<sup>39</sup>

However, we expect the extent of proteolysis to be less significant in the One-Pot PURE and 36-Pot PURE preparations compared to what was observed in this T7 RNAP purification. This expectation arises from the high mRNA transcription activity observed with the Dual-Strain One-Pot PURE system in an earlier experiment (Figure 2A). We hypothesize that the apparent impact of omptin proteolysis on the protein expression capacity of the Dual-Strain One-Pot PURE and 36-Pot systems may be reduced by extra copies of T7 RNAP produced by BL21(DE3) cells and a lower abundance of omptins in the cell mixture with BL21(DE3) cells.

We did not pursue in-house experimental validation of the omptin family protease-mediated proteolysis on the other six PURE proteins. Verifying proteolysis would require purifications for each of the six proteins and incubation in a culture with *E. coli* M15 cells carrying the omptin family outer membrane protease to observe proteolysis. We believe that the simulation results, literature validation of T7 RNAP and CK lysis, and the fact that the original PURE system was developed in BL21/pREP4 and BL21(DE3) strains lacking OmpT protease<sup>40</sup> are sufficient to justify M15/pREP4, used in the development of One-Pot PURE,<sup>25</sup> is an unsuitable strain for PURE protein expression. As a result, we find it is critical to transfer all PURE protein expressions to a protease-deficient strain to ensure the production of full-length, active PURE proteins.

**Enhancing One-Pot PURE Protein Composition and Expression Capacity with a Single-Strain One-Pot PURE System.** To prevent unwanted proteolysis and attain more uniform cell growth in the coculture environment, we transferred all PURE proteins encoded in the pQE30 and pQE60 vectors for M15/pREP4 expression to pET21 vectors for BL21(DE3) expression. However, a caveat of this Single-Strain One-Pot PURE system is that the T7 RNAP needs to be removed from the One-Pot PURE coculture due to the genetic instability of having a P<sub>T7</sub> promoter expressing T7 RNAP. As a result, we opted to purify and add T7 RNAP into One-Pot PURE reactions separately (see the Materials and Methods section). Alternatively, T7 RNAP could be cotransformed with the pREP4 plasmid carrying constitutive LacI expression into BL21 cells and grown in the One-Pot PURE coculture.

Our results showed that five out of the 16 proteins previously expressed in M15/pREP4 cells exhibited at least a 30% improvement in expression levels when grown in BL21(DE3) cells, as determined by monoculture growth and SDS-PAGE analysis (Figure SI.III 5). This enhancement is also evident in the Single-Strain One-Pot PURE coculture. Compared to the Dual-Strain system, the purified PURE protein mixture from the Single-Strain system has a slightly higher composition of translation factors (24 and 26% for Dual Strain and Single Strain, respectively; Figure 5B). Using mass spectrometry to resolve individual protein composition, we found that translation initiation factors 2 and 3 (IF2 and IF3) and creatine kinase (CK) involved in energy regeneration (Figure 5A) are present at significantly higher concentrations in Single-Strain One-Pot PURE. Fold changes in individual protein concentrations in the Single-Strain One-Pot PURE compared to the original Dual-Strain mixture are tabulated in Table SI.III 1. An individual protein composition comparison of the Single-Strain One-Pot PURE to commercial PURE systems is also provided in Figure SI.III 6.



**Figure 4.** (A) Schematic of T7 RNAP preparation to assess the outer membrane omptin family protease activity present in M15/pREP4 and BL21/pREP4 strains. A single culture of cells was grown, induced for T7 RNAP expression, and harvested, following a procedure similar to One-Pot PURE coculture preparations. An unlysed aliquot of cells expressing T7 RNAP was saved for SDS-PAGE analysis. The cultures were then split into two lysis and purification groups, each treated with a different protease inhibitor (cCOMPLETE and Pepstatin A). (B) SDS-PAGE gel result showed T7 RNAP proteolysis when grown and purified in M15/pREP4 strains. Omptin proteases cleaved full-length T7 RNAP into truncated 80 and 20 kDa fragments. Proteolysis of T7 RNAP is absent in the BL21/pREP4 strain, which is deficient in the OmpT protease. (C) Simulated likelihood score of PURE proteins with peptide subsequences susceptible to omptin proteolysis. The identified peptide sequences were assessed for their accessibility for omptin proteolysis (Figure SI.III 4) and compared with existing literature. Simulation results for T7 RNAP and creatine kinase (CK) are confirmed by literature findings.<sup>34,36</sup> The notation of n/a indicates data not available.

We next compared the productivity of the Dual-Strain and Single-Strain versions of the One-Pot PURE system. Using only the energy solution from commercial vendors and replacing commercial PURE proteins with One-Pot PURE proteins, we observed consistently higher protein production with the Single-Strain One-Pot PURE system (Figure 5C,D). Since both commercial vendors' energy solutions contained DTT as the reducing agent, which has unfavorable interactions with malachite green dye, preventing accurate quantification of mRNA transcripts using malachite green aptamer (MGA) fluorescence as a proxy, we switched the reporter plasmid from P<sub>T7</sub>-MGA-UTR1-deGFP used in Figure 2 to P<sub>T7</sub>-UTR1-deGFP.<sup>28</sup> We conducted control experiments to rule out the possibility that the improvement in gene expression resulted from switching reporter plasmids (Figure SI.III 7).

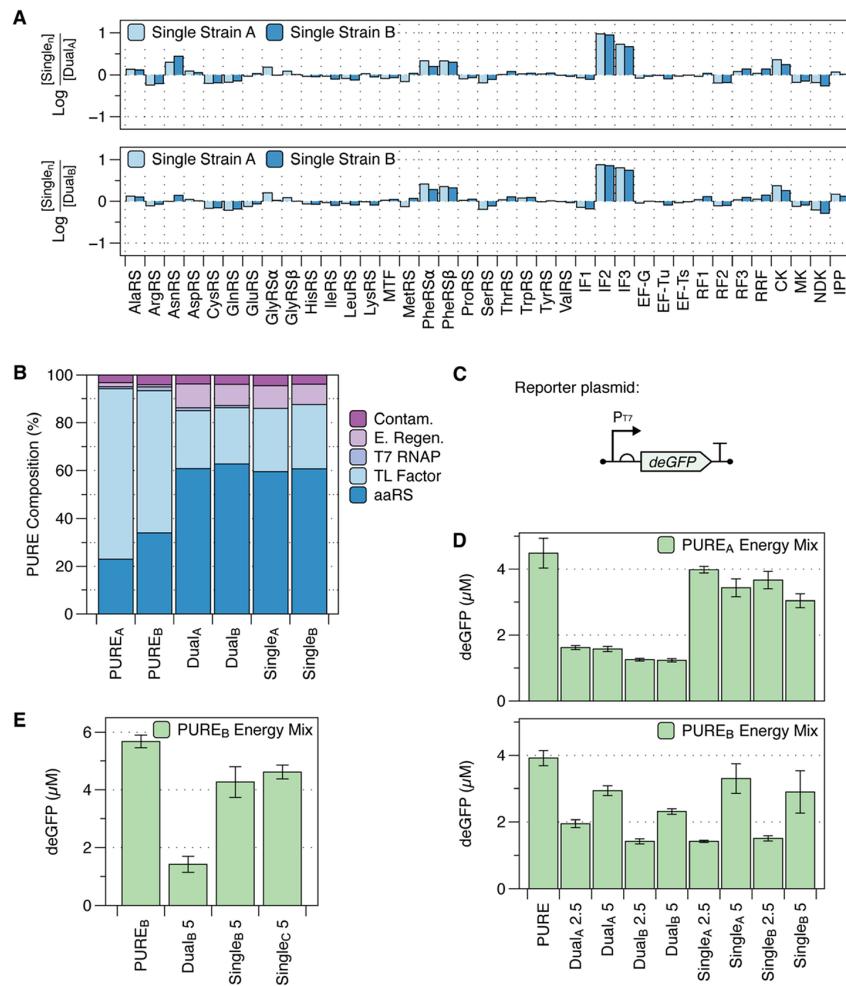
The robustness of the Single-Strain system was further validated through a reproducibility study conducted in our collaborator's lab. A new batch of Single-Strain One-Pot PURE system was prepared following the same protocol (referred to as Single<sub>C</sub>), and its productivity was compared to shipped aliquots of previous batches, Single<sub>B</sub> and Dual<sub>B</sub>. Even in a different laboratory setting, using separate equipment and reaction consumables, a higher yield of the Single-Strain system was reproduced on the first attempt. This was observed both with commercial energy systems (Figure 5E) and with homemade energy solutions (Figure SI.III 8,9).

**High-Yield Protein Expression from One-Pot PURE Requires a Balance between the Protein Production Rate and Energy Solution Formulation.** The differences in productivity between the Dual-Strain One-Pot PURE system when using two commercial PURE energy solutions, referred

to as PURE<sub>A</sub> and PURE<sub>B</sub>, are particularly interesting. It appears that higher amounts of One-Pot PURE proteins can be added in reactions with PURE<sub>B</sub> energy solution as a strategy to compensate for a suboptimal stoichiometry of translation factors in the purified protein mixture (Figure 5D). In reactions with the PURE<sub>A</sub> and homemade energy solution, increasing the concentration of One-Pot PURE proteins from 2.5 to 5 mg/mL decreased protein production.

Although we do not know the exact formulation of these energy solutions, comparing the protein production rates (coupled transcription-translation rates) of One-Pot PURE systems to commercial PURE systems provides valuable insights (Figures 6A and SI.III 10,11). In the PURE<sub>A</sub> energy solution, the protein production rates of the Dual-Strain and Single-Strain One-Pot PURE systems reflected the anticipated PURE protein differences based on their translation factor content (Figure 6B). Deficiencies in translation factors in the Dual-Strain One-Pot PURE system contributed to a lower production rate over the course of the reaction (Figure 6B, subpanels i,ii). With more protein translation machinery, the Single-Strain One-Pot PURE system could match or even exceed the production rate of commercial PURE<sub>A</sub> proteins (Figure 6B, subpanels iii,iv). However, higher production rate peaks led to steeper rate declines shortly after, as observed for the Single-Strain One-Pot PURE system with the PURE<sub>A</sub> energy solution.

In the PURE<sub>B</sub> energy solution, when both the Dual-Strain and Single-Strain One-Pot PURE proteins are added to 5 mg/mL, both systems exhibited comparable protein expression levels despite the deficiency in translation factors found in the Dual-Strain One-Pot PURE protein mixture (Figure 6B,

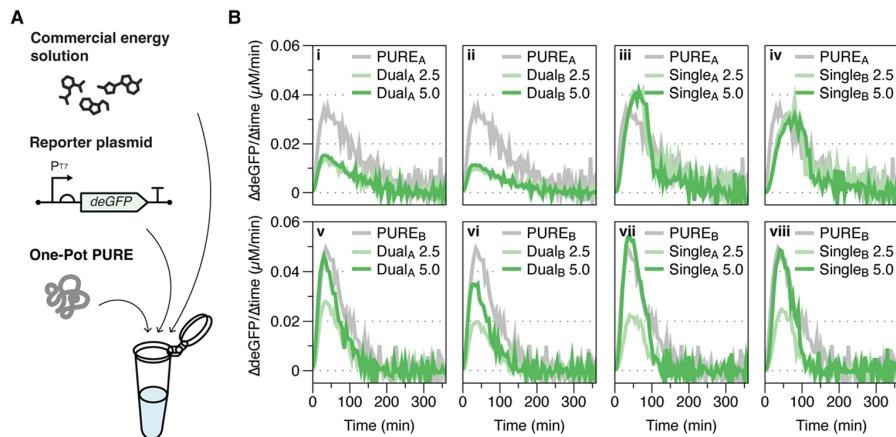


**Figure 5.** (A) Comparison of the relative abundances of individual One-Pot PURE proteins in the Single-Strain and Dual-Strain systems. Initiation factors 2 and 3 (IF2, IF3) show the most notable increase in composition in the Single-Strain One-Pot PURE system. Previously underexpressed aminoacyl-tRNA synthetases (AlaRS and PheRS) and creatine kinase (CK) also increased in composition in the Single-Strain One-Pot PURE system. (B) Composition of commercial PURE proteins, Dual-Strain One-Pot PURE proteins, and Single-Strain One-Pot PURE proteins grouped by protein classes: aminoacyl-tRNA synthetase (aaRS), translation factors (TL factors), T7 RNAP, and energy regeneration (E Regen.). Detected proteins not part of the PURE proteins are categorized as contaminants (Contam.). Ribosomal proteins are excluded from the analysis. (C) Schematic of reporter plasmid  $P_{T7}$ -UTR1-deGFP expressed in PURE reactions to assess system productivity. (D) Productivity of Dual-Strain and Single-Strain One-Pot PURE reactions assembled with two commercial energy solutions. PURE<sub>A</sub> and PURE<sub>B</sub> are control reactions with 5 nM  $P_{T7}$ -UTR1-deGFP plasmids assembled according to the manufacturer's protocol. “Dual<sub>n</sub> 2.5” and “Dual<sub>n</sub> 5” designate varying batches of Dual-Strain One-Pot PURE proteins added at 2.5 and 5 mg/mL. Likewise, “Single<sub>n</sub> 2.5” and “Single<sub>n</sub> 5” designate varying batches of Single-Strain One-Pot PURE proteins added at 2.5 and 5 mg/mL. The bar graphs represent the average terminal protein yield of reaction triplicates, with error bars indicating the standard deviations of reaction triplicates. (E) Productivity comparison of the collaborator's batch of Single-Strain One-Pot PURE (Single<sub>C</sub>). PURE<sub>B</sub> is the control reaction assembled according to the manufacturer's protocol. Single-Strain and Dual-Strain One-Pot PURE reactions were assembled using 5 mg/mL PURE protein, using the commercial energy solution from supplier B and 5 nM  $P_{T7}$ -UTR1-deGFP plasmid. The bar graphs represent the average terminal protein yield of reaction triplicates, with error bars indicating the standard deviations of reaction triplicates.

subpanels v–viii). It is important to note that the contribution to these differences in translation rate comes from the energy solution alone (amino acids, NTPs, tRNA, creatine phosphate, and other cofactors). All reactions were assembled with the same reporter plasmids and ribosomes. This result suggests that an optimal preparation of the energy solution could compensate for an otherwise suboptimal One-Pot PURE protein stoichiometry in reactions.

We found it intriguing that when the production rate from the One-Pot PURE system exceeded that of commercial PURE

proteins, a steeper drop in the production rate soon followed (Figure 6B, subpanels iii, iv, vii, and viii). We first attempted to explain the phenomenon from the perspective of the system's energy regeneration capability. While fast protein production can occur, it can also significantly deplete the available energy molecules (ATP and GTP) critical for powering translation initiation and elongation reactions. A fast production rate surpassing the system's energy regeneration capacity can negatively impact the final protein yield by stalling protein translation, potentially resulting in premature translation



**Figure 6.** (A) Schematic of the One-Pot PURE reaction composition using Dual-Strain and Single-Strain One-Pot PURE proteins with commercial energy solutions, assembled according to manufacturers' protocols. All reactions used 5 nM  $P_{T7}$ -UTR1-deGFP reporter plasmid. In the case of Single-Strain One-Pot PURE reactions, 0.5  $\mu$ M purified T7 RNAP was added. Chemical structures and protein icons are obtained from BioRender. (B) Comparison of the protein translation rate for various preparations of One-Pot PURE proteins (added to the reaction at 2.5 or 5.0 mg/mL) using two commercial energy solutions. Plots represent the average of reaction triplicates. Time traces of deGFP production are provided in the Supplementary Figures SI.III 8,9.

termination and lower apparent protein production. This corroborates the finding that supplementing an additional energy regeneration system to the PURE reaction can help attain a higher maximum protein production rate.<sup>41</sup>

However, the energy regeneration bottleneck can only explain the protein production behavior observed in reactions with PURE<sub>A</sub> energy solution and Single-Strain One-Pot PURE proteins (Figure 6B, subpanels iii,iv). If energy regeneration was the only bottleneck in protein production, then reactions with suboptimal PURE protein stoichiometry (Dual-Strain, One-Pot PURE) should have a lower yet longer-lasting maximal protein production rate in the PURE<sub>A</sub> energy solution. Similarly, energy regeneration alone cannot explain the discrepancy between the Dual-Strain and Single-Strain One-Pot PURE expression behaviors in PURE<sub>A</sub> and PURE<sub>B</sub> energy solutions.

These observations led us to consider additional factors contributing to the differences in protein expression rates. One component often overlooked in optimizing cell-free expression capacity is the tRNA composition and codon usage. Because the exact tRNA composition for each codon in the PURE reaction is not known, and protein translation is influenced by codon usage, it is often challenging to optimize the expressed protein's codon usage to the tRNA composition in the reaction.

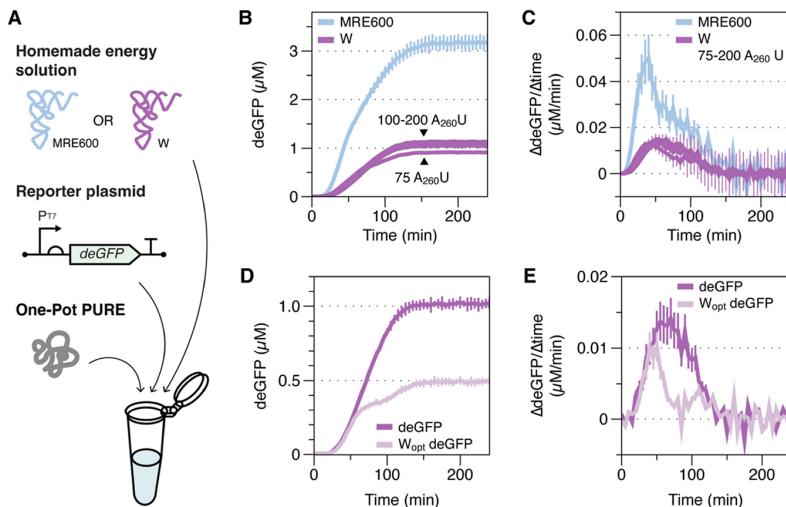
Since the same reporter plasmid was used for Dual-Strain and Single-Strain One-Pot PURE reactions in different PURE energy solutions, the differences in the final yield and maximal production rate may be influenced by the tRNA codon composition in different energy formulations. We hypothesize that a suboptimal tRNA composition in the reaction, not matching well with the codon usage of the reporter protein, may also contribute to ribosome stalling on the mRNA transcript when encountering suboptimal codons.<sup>42</sup> This could be exacerbated by an mRNA transcript with a high translation initiation rate, such as the one used in the reporter plasmid (UTR1 ribosomal binding site),<sup>43</sup> which increases the frequency of ribosomal collisions. Colliding ribosomes could lead to ribosomal falloff, resulting in a low apparent yield of

protein translation. This also corroborates reports of C-terminal truncated protein produced in PURE reactions.<sup>44</sup>

**E. coli** tRNA Composition Critically Influences One-Pot PURE Reaction Productivity. We further demonstrated the influence of suboptimal tRNA composition on protein production by preparing homemade energy solutions using two different sources of commercial *E. coli* tRNA mixtures: one from MRE600 and one from W. The homemade energy solution formulation followed the exact recipe of the One-Pot PURE protocol, with the exception of the source of *E. coli* tRNA.<sup>27</sup> In these experiments, the same concentrations of the Single-Strain One-Pot PURE proteins, reporter plasmids, ribosomes, and T7 RNAP were used. The only difference between the reactions was the source—and possibly the composition—of *E. coli* tRNA in the energy solution (Figure 7A).

It was observed that the energy solution prepared using MRE600 tRNA exhibited a high protein production rate and final protein yield (Figure 7B). When we replaced tRNA in the energy solution with W tRNA, we observed a significant decrease in the protein production rate and yield (Figure 7C). Increasing the concentration of W tRNA did not improve the rate and yield of protein expression, suggesting the possibility of a critical deficiency in one or more tRNA bottlenecks in the translation reaction, as evidenced by the fact that increasing the tRNA concentration did not improve the protein expression rate.

To assess potential codon usage bias associated with switching to the W tRNA pool, we codon-optimized the deGFP to *E. coli* W's tRNA using the CDS Calculator.<sup>45–49</sup> Interestingly, W codon-optimized deGFP exhibited only 0.5-fold of the original deGFP's expression level, possibly due to expanded codon usage compared to the original deGFP design (Figure 7C and Table SI.III 2). Notably, the protein expression rate for W-optimized deGFP was comparable to that of the original deGFP for the first 50 min of the reaction, suggesting that translation initiation and elongation were not limiting factors early in the process. However, after 50 min, expression reached a bottleneck, resulting in a sharp decline in the



**Figure 7.** (A) Schematic of the One-Pot PURE reaction composition using in-house prepared energy solutions with commercial tRNAs from *E. coli* MRE600 and *E. coli* W. Each reaction contained 5 nM P<sub>T7</sub>-UTR1-deGFP to assess protein production. Transfer RNA structures and protein icons used are obtained from BioRender. (B) Comparison of deGFP yield for batch A of the Single-Strain One-Pot PURE proteins (added to the reaction at 2.5 mg/mL) in energy solution with 22 A<sub>260</sub>U/mL MRE600 tRNA or 30–80 A<sub>260</sub>U/mL W tRNA. Error bars represent standard deviations of reaction triplicates. (C) Protein production rates represent the average of the reaction triplicates. Error bars represent standard deviations of reaction triplicates. (D) Comparison of deGFP yield using the original deGFP codon usage compared to W codon-optimized deGFP usage. Error bars represent standard deviations of reaction triplicates. (E) Protein production rates represent the average of reaction triplicates. Error bars represent standard deviations of reaction triplicates.

expression rate. We also note that W tRNA is currently unavailable for purchase, which limits our ability to investigate these effects further.

Nevertheless, these findings are significant because the complexity and composition of tRNAs can be overlooked when preparing energy solutions for PURE reactions. Differences in codon usage among various *E. coli* strains can lead to unexpected deficiency in tRNAs associated with the codons used. Without prior knowledge of these tRNA deficiencies, it can be difficult to codon-optimize the protein for effective expression, resulting in a low apparent protein yield despite an otherwise productive system.

## CONCLUSIONS

In this work, we identified several factors impacting the expression capacity of the One-Pot PURE system. We improved the stability of One-Pot PURE protein expression through glucose-mediated catabolite repression. By switching from the original *E. coli* M15/pREP4 and BL21(DE3) dual-strain PURE protein expression approach to a single BL21(DE3) strain approach, we ensured the uniformity of cell growth at the time of induction for protein expression to improve the PURE protein stoichiometry in the purified mixture. We also mitigated undesirable protease degradation of PURE proteins from M15/pREP4 expression.

Of note, the original PURE system used BL21/pREP4 and BL21(DE3), which do not have the optin family protease, for all PURE protein expression from pQE and pET expression vectors, respectively.<sup>40</sup> Our findings suggest that the M15/pREP4 strain used in the original One-Pot PURE system is unsuitable for PURE protein production. By switching all PURE protein expression to the BL21 family, this new “single-strain” approach also increased the expression levels of critical translation initiation and energy regeneration proteins and, consequently, the productivity of the One-Pot PURE reaction.

Interestingly, our characterizations of both Dual- and Single-Strain One-Pot PURE proteins in two commercial PURE energy solutions revealed that an optimized energy solution formulation could overcome suboptimal PURE protein stoichiometry (Figure 5D). Replicating this phenomenon using our in-house prepared energy solutions with different *E. coli* tRNA sources revealed the critical role of the reaction’s tRNA pool in the translation capacity of PURE reactions.

These findings showcased that the gene expression capacity of the PURE system is influenced by both the PURE protein stoichiometry and biochemical factors. At the protein level, we found that when translation initiation factors (initiation factors 2 and 3, IF2 and IF3) and energy regeneration protein, creatine kinase (CK), are present at higher concentrations in the PURE protein mixture, higher expression capacity is observed (Figure 5D,E). At the biochemical level, we found that the energy solution composition significantly impacts the One-Pot PURE production rate and the final protein yield. Using the PURE<sub>A</sub> energy solution, the reaction proceeded with a slower protein production rate and took over 6 h to stabilize to the final deGFP concentration (Figure SI.III 10). This suggests that there may be a reaction rate-limiting step imposed by the PURE<sub>A</sub> energy solution, and this limitation can make One-Pot PURE proteins with suboptimal protein composition to proceed even more slowly, resulting in lower protein yield. On the other hand, the PURE<sub>B</sub> energy solution supports a faster protein production rate overall, with reactions completed in under 3 h (Figure SI.III 11). This suggests that the rate-limiting step observed in the PURE<sub>A</sub> energy solution may be removed in PURE<sub>B</sub> to deliver a high protein yield with a shorter reaction lifetime. We hypothesize that this rate-limiting step can be the reaction’s tRNA composition. We also note that PURE<sub>A</sub> and PURE<sub>B</sub> energy solutions are proprietary information from commercial vendors. The exact compositions

of these energy solutions are not known to us, which limits the conclusions we can draw.

Lastly, we show that commercial preparations of *E. coli* tRNA sourced from strains MRE600 and W significantly impact effective protein translation. Specifically, we found that an optimal One-Pot PURE protein mixture could lose more than 3-fold in protein expression by switching to an energy formulation with a different source of *E. coli* tRNA (Figure 7A,B). Similarly, by altering the codon usage coding for the same set of amino acids, the expression capacity in the One-Pot PURE reaction could be reduced by half (Figure 7C,D).

Although MRE600 is a highly preferred source for *E. coli* tRNA due to the strain's low RNase I activity,<sup>50</sup> there is no apparent reason behind tRNA obtained from strain W (a safe and fast-growing strain utilizing sucrose as an industrially preferred carbon source)<sup>51</sup> exhibiting suboptimal translation capacity. However, the low expression rate exhibited by the strain W tRNA demonstrates that at least one of the critical tRNAs used for the reporter protein expression was critically deficient in this system, which bottlenecks the reaction. This highlights a critical limitation and source of variability in most cell-free systems (lysate-based and PURE), which is that *E. coli* tRNA is also a complex mixture and exhibits differential concentration profiles across growth stages and host strains.<sup>52,53</sup> For robust control over both One-Pot PURE reactions and potentially lysate-based cell-free expression systems, a promising approach can be developing an optimal tRNA composition.<sup>54</sup>

Taken together, this work seeks to obtain a comprehensive understanding of the factors influencing the gene expression capacity of the One-Pot PURE system. By addressing both protein and biochemical composition, our findings offer valuable insights into not only optimizing One-Pot PURE systems but also extending them to lysate-based cell-free protein synthesis, paving the way for more efficient and reliable applications in synthetic biology and biotechnology.

## MATERIALS AND METHODS

**Escherichia coli Strains and Plasmids.** *E. coli* BL21-(DE3) (Thermo Scientific) and M15/pREP4 (Qiagen) strains were used for protein expression. Supporting Information I contains the Addgene catalog numbers and plasmid sequences for Dual-Strain One-Pot. To establish the Single-Strain One-Pot, all proteins, except T7 RNAP, were amplified via Gibson Assembly and placed onto the pET21a backbone. All sequences of cloned proteins transplanted onto pET21a are provided in Supporting Information I.

Sequence-verified plasmids (Primordium Laboratories, now Plasmidsaurus) were transformed into their respective BL21-(DE3) and M15/pREP4 expression strains and plated on Luria Broth (LB) Agar plates with 100 µg/mL carbenicillin and 1% glucose. An inoculation loop was used to scrape multiple colonies to inoculate the overnight culture in LB with 100 µg/mL carbenicillin and 1% glucose to obtain the population expression average. To facilitate inoculation of subsequent final cultures, 250 µL of the overnight culture was saved into frozen glycerol stocks with 250 µL of 50% glycerol in 1.3 mL of 96-well plates (NEST Biotech). The glycerol stock plate was sealed with an aluminum film and stored at -80 °C.

Sequences of plasmid DNA used for reporter protein expression are provided in Supporting Information I. All plasmid DNA was prepared by using the NucleoBond Xtra Midi kit (Macherey-Nagel) and eluted in nuclease-free water.

## Monoculture Expression Assay for One-Pot PURE Proteins.

The protocol was adapted from Grasemann et al.<sup>27</sup> with few modifications. An overnight culture of One-Pot PURE protein expression strains was inoculated either from freshly transformed cells grown on LB agar plates with carbenicillin (without or with 1% w/v glucose to exert catabolite repression) or from frozen glycerol stocks using a cryoreplicator (Enzysscreen, CR1000). The overnight culture was grown in 96-well deep-well plates sealed with a Breath-Easier membrane (Diversified Biotech) at 37 °C and 1000 rpm.

The next morning, a fresh starter culture was prepared using 4 µL of overnight cells and 396 µL of LB with carbenicillin. OD600 was monitored at the time of 0.1 mM IPTG induction to initiate protein expression after 2–2.5 h of growth and 3 h after protein expression. 10 µL of cells expressing One-Pot PURE proteins was mixed with Laemmli Buffer (Bio-Rad) and heat-denatured at 95 °C for 10 min, and 10 µL of the mixture was loaded onto Bolt 4–12% Bis-Tris Protein Gels (Invitrogen), along with protein ladders (SeeBlue Plus2 Prestained Protein Standard, Invitrogen). Samples were run at 150 kV and 400 mA for 30 min with NuPAGE MES SDS buffer (Invitrogen).

The gel was stained with SimplyBlue Safe Stain (Invitrogen) for 1 h and destained in water overnight before imaging on a ChemiDoc Imager (Bio-Rad) or a GelDoc Go Imager (Bio-Rad). The obtained protein gel images were analyzed using the accompanying software Image Lab 6.1 for Mac (Bio-Rad).

**One-Pot PURE Reaction Assembly and Data Acquisition.** Detailed methods describing the preparation of One-Pot PURE protein mixture and energy solution are provided in Supporting Information II.

All One-Pot reactions were assembled on ice using low-protein-binding microcentrifuge tubes (Eppendorf). Each reaction is 5 µL in total volume and consists of 2 µM *E. coli* ribosome (NEB), 5 nM reporter protein expression plasmids (either P<sub>T7</sub>-MGA-UTR1-deGFP or P<sub>T7</sub>-UTR1-deGFP), 2.5× Energy Solution (125 mM HEPES-KOH pH 7.6, 250 mM potassium L-glutamate, 5 mM ATP and GTP at pH 7.5, 2.5 mM UTP and CTP at pH 7.5, either 56 A<sub>260</sub>U/mL for *E. coli* MRE600 tRNA or 75–200 A<sub>260</sub>U/mL for *E. coli* W tRNA, 50 mM creatine phosphate, 2.5 mM TCEP, 50 µM folic acid, 5 mM spermidine, and 0.75 mM amino acid mixture at pH 7.5), 1.75–5 mg/mL One-Pot PURE proteins, and 0.5–1 µM T7 RNAP (only for Single-Strain One-Pot). For reactions using malachite green aptamer (MGA) fluorescence as a proxy for mRNA transcripts, 10 µM malachite green dye was added to the reaction.

Following assembly, 5 µL of each reaction was pipetted onto a low-volume 384-well microplate with a nonbinding surface (Corning), and the reaction trajectory was measured every 5 min using the BioTek Synergy Plate Reader at 37 °C. Fluorescence for deGFP was measured at 485/515 nm excitation/emission wavelengths at Gain 61, and fluorescence for malachite green aptamer was measured at 610/650 nm excitation/emission wavelengths at Gain 150. The fluorescence data for MGA and deGFP were calibrated to protein concentration data using synthesized MGA RNA and deGFP protein standards, as provided in Supporting Information III, Figure S1III 12.

**Acquisition and Analysis of Proteomics Data.** A brief summary of the LC–MS/MS workflow is provided below. Detailed methods are provided in Supporting Information II.

LC–MS/MS analyses of the digested peptides were performed on an EASY-nLC 1200 (Thermo Fisher Scientific) coupled to a Q Exactive HF hybrid quadrupole–Orbitrap mass spectrometer (Thermo Fisher Scientific). MS2 fragmentation spectra were searched with Proteome Discoverer SEQUEST (version 2.5; Thermo Scientific, Waltham, MA) against an *in silico* tryptic digested *E. coli* BL21 (DE3) and *E. coli* M15 database. The relative abundance of parental peptides was calculated by integrating the area under the curve of the MS1 peaks.

The proteomics data presented in this paper are from the same batch of LC-MS/MS runs for results to be comparable. The relative protein abundances for each sample are first normalized by scaling the total sample protein abundance with the average total protein abundance for all samples. The proteins involved in the PURE system were extracted from the entire proteome data set and scaled by their composition in PURE reactions (manufacturer's recommended compositions for PURE<sub>A</sub> and PURE<sub>B</sub>, 4 mg/mL for Dual-Strain One-Pot PURE batches A and B in Figure 2, and 2.5 mg/mL for all Dual- and Single-Strain One-Pot PURE batches A and B in Figure 5).

**Characterization of Protease Activity in M15/pREP4 and BL21/pREP4 Strains.** Plasmid-expressing T7 RNAP (Addgene no. 124128) was cotransformed into M15 and BL21 cells with pREP4 plasmid coding for constitutive LacI expression. Transformants were grown in a 5 mL overnight culture with LB with antibiotics and 1% glucose. 1 mL of the overnight culture was used to start the 150 mL final culture with LB and antibiotics. When the culture OD exceeded 0.8, 1 mM IPTG was used to induce T7 RNAP expression, and the culture was grown for an additional 3 h. At harvest, the cultures were divided into 50 mL aliquots and centrifuged at 5000g for 20 min to obtain cell pellets, and the pellets were stored at –80 °C until use. One 100 μL aliquot of the M15 and BL21 cultures was saved as unlysed controls.

To characterize M15 proteolysis during protein purification, each of the cell pellets was resuspended in 1 mL of lysis buffer (50 mM HEPES-KOH pH 7.6, 100 mM NH<sub>4</sub>Cl, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, and 1 mM TCEP). Two variations of the lysis buffer were prepared: one with Pepstatin A (1 μg/mL), and another with cOmplete Mini, EDTA-free protease inhibitor at the recommended working concentration. The resuspended cells were lysed using the freeze–thaw method by alternating between flash freezing in liquid nitrogen and thawing in a room-temperature water bath 8 times. The lysate was centrifuged at 15,000 rpm at 4 °C for 20 min to remove cellular debris before His-tag purification.

A 2 mL column (Thermo Scientific) was prepared with 200 μL of Ni-Sepharose resin (Cytiva) and equilibrated using 15 mL of wash buffer (50 mM HEPES-KOH pH 7.6, 100 mM NH<sub>4</sub>Cl, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM TCEP, and 20 mM imidazole-HCl pH 7). Following equilibration, the lysate–resin mixture was rocked at 4 °C for 1 h to enhance protein binding to the resin, washed with 20 mL of wash buffer, and eluted with 250 μL of elution buffer (50 mM HEPES-KOH pH 7.6, 100 mM NH<sub>4</sub>Cl, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM TCEP, and 500 mM imidazole-HCl pH 7). The unlysed cell culture aliquots, as well as purified proteins from M15 and BL21 strains, were then prepared for SDS-PAGE analysis using the same setting described in the previous sections.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.4c00779>.

Sequence information for all plasmids used in this work, including their Addgene catalog numbers where applicable ([PDF](#))

Extended methods for One-Pot PURE protein purification, homemade PURE energy solution preparation, T7 RNA polymerase and deGFP purification, and the mass spectrometry workflow for proteomics analysis ([PDF](#))

Supplementary figures and tables referenced throughout the main text, providing additional analysis and characterizations to support the paper's conclusion ([PDF](#))

## ■ AUTHOR INFORMATION

### Corresponding Author

Yan Zhang — Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0003-0719-5456](https://orcid.org/0000-0003-0719-5456); Email: [yz473@caltech.edu](mailto:yz473@caltech.edu)

### Authors

Matas Deveikis — Department of Infectious Diseases, Faculty of Medicine, Imperial College London, London SW7 2AZ, U.K.; [orcid.org/0000-0001-7444-2950](https://orcid.org/0000-0001-7444-2950)

Yanping Qiu — Proteome Exploration Laboratory, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0003-2948-2173](https://orcid.org/0000-0003-2948-2173)

Lovisa Björn — Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0009-0007-6983-3341](https://orcid.org/0009-0007-6983-3341)

Zachary A. Martinez — Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0002-7830-3162](https://orcid.org/0000-0002-7830-3162)

Tsui-Fen Chou — Proteome Exploration Laboratory, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0003-2410-2186](https://orcid.org/0000-0003-2410-2186)

Paul S. Freemont — Department of Infectious Diseases, Faculty of Medicine, Imperial College London, London SW7 2AZ, U.K.; [orcid.org/0000-0002-5658-8486](https://orcid.org/0000-0002-5658-8486)

Richard M. Murray — Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0002-5785-7481](https://orcid.org/0000-0002-5785-7481)

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acssynbio.4c00779>

### Author Contributions

Y.Z., R.M.M., P.S.F., and T.C.: Conceptualization; Y.Z., M.D., Y.Q., L.B., and Z.A.M.: Investigation; Y.Z., M.D., Y.Q., L.B., and Z.A.M.: Analysis; Y.Z., M.D., L.B., and Z.A.M.: Writing—manuscript; Y.Z., M.D., R.M.M., P.S.F., L.B., Z.A.M., Y.Q., and T.C.: Writing—review and editing; Y.Z., M.D., L.B., and Z.A.M.: Visualization; R.M.M., P.S.F., and T.C.: supervision.

### Funding

R.M.M. and Y.Z. are supported by the National Science Foundation, award MCB-2152267, and Schmidt Futures, grant number 2023-3-2-1. Y.Z. is also supported by the Caltech Presidential Postdoctoral Fellowship. P.S.F. and M.D. are

supported by the Engineering and the Physical Sciences Research Council (EPSRC), award EP/T013788/1. M.D. is also supported by the EPSRC Centres for Doctoral Training in Biodesign Engineering at Imperial College London. L.B. is supported by the Caltech Summer Undergraduate Research Fellowship (SURF).

## Notes

The authors declare the following competing financial interest(s): R.M.M. has a financial stake in Tierra Biosciences, a private company that uses bacterial lysate-based cell-free technologies for protein expression and screening.

## ACKNOWLEDGMENTS

We sincerely thank the Murray Lab members, particularly Dr. Zoila Jurado and Miryong (Miki) Yun, who contributed to the project's early setup and discussions. We also thank b.next, a company working on open-source synthetic cell protocols (<https://bnext.bio>), for insightful discussions on the 36-Pot PURE preparation. We also thank Dr. Jurado, Manisha Kapasiawala, and Dr. John Marken for their review and feedback on this manuscript. Dr. Jurado prepared the P<sub>T7</sub>-MGA-UTR1-deGFP plasmid<sup>28</sup> used in this work. We thank Prof. Lulu Qian for access to the ChemiDoc Imager used for protein gel image acquisition. We thank the Caltech Protein Expression Center (PEC) for preparing the purified T7 RNAP used in this work. The chemical, protein, and tRNA icons used in Figures 6A and 7A are obtained from BioRender under the Academic licensing agreement ZF282CDXBI. The authors made all the graphics in Adobe Illustrator.

## REFERENCES

- (1) Hunt, A. C.; Vögeli, B.; Hassan, A. O.; Guerrero, L.; Kightlinger, W.; Yoesep, D. J.; Krüger, A.; DeWinter, M.; Diamond, M. S.; Karim, A. S.; Jewett, M. C. A Rapid Cell-Free Expression and Screening Platform for Antibody Discovery. *Nat. Commun.* **2023**, *14* (1), 3897.
- (2) Swank, Z.; Maerkl, S. J. CFPU: A Cell-Free Processing Unit for High-Throughput, Automated In Vitro Circuit Characterization in Steady-State Conditions. *BioDesign Res.* **2021**, *2021*, No. 2968181.
- (3) Tabuchi, T.; Yokobayashi, Y. High-Throughput Screening of Cell-Free Riboswitches by Fluorescence-Activated Droplet Sorting. *Nucleic Acids Res.* **2022**, *50* (6), 3535–3550.
- (4) Martin, R. W.; Des Soye, B. J.; Kwon, Y.-C.; Kay, J.; Davis, R. G.; Thomas, P. M.; Majewska, N. I.; Chen, C. X.; Marcum, R. D.; Weiss, M. G.; Stoddart, A. E.; Amiram, M.; Ranji Charna, A. K.; Patel, J. R.; Isaacs, F. J.; Kelleher, N. L.; Hong, S. H.; Jewett, M. C. Cell-Free Protein Synthesis from Genomically Recoded Bacteria Enables Multisite Incorporation of Noncanonical Amino Acids. *Nat. Commun.* **2018**, *9* (1), 1203.
- (5) Ngo, P. H. T.; Ishida, S.; Busogi, B. B.; Do, H.; Ledesma, M. A.; Kar, S.; Ellington, A. Changes in Coding and Efficiency through Modular Modifications to a One Pot PURE System for In Vitro Transcription and Translation. *ACS Synth. Biol.* **2023**, *12* (12), 3771–3777.
- (6) Zhang, Y.; Kojima, T.; Kim, G.-A.; McNerney, M. P.; Takayama, S.; Styczynski, M. P. Protocell Arrays for Simultaneous Detection of Diverse Analytes. *Nat. Commun.* **2021**, *12* (1), 5724.
- (7) Manzer, Z. A.; Ghosh, S.; Jacobs, M. L.; Krishnan, S.; Zipfel, W. R.; Piñeros, M.; Kamat, N. P.; Daniel, S. Cell-Free Synthesis of a Transmembrane Mechanosensitive Channel Protein into a Hybrid-Supported Lipid Bilayer. *ACS Appl. Bio Mater.* **2021**, *4* (4), 3101–3112.
- (8) Thakur, M.; Breger, J. C.; Susumu, K.; Oh, E.; Spangler, J. R.; Medintz, I. L.; Walper, S. A.; Ellis, G. A. Self-Assembled Nanoparticle-Enzyme Aggregates Enhance Functional Protein Production in Pure Transcription-Translation Systems. *PLoS One* **2022**, *17* (3), No. e0265274.
- (9) Luo, S.; Adam, D.; Giaveri, S.; Barthel, S.; Cestellos-Blanco, S.; Hege, D.; Paczia, N.; Castañeda-Losada, L.; Klose, M.; Arndt, F.; Heider, J.; Erb, T. J. ATP Production from Electricity with a New-to-Nature Electrobiochemical Module. *Joule* **2023**, *7* (8), 1745–1758.
- (10) Jung, J. K.; Alam, K. K.; Verosloff, M. S.; Capdevila, D. A.; Desmau, M.; Clauer, P. R.; Lee, J. W.; Nguyen, P. Q.; Pastén, P. A.; Matiasek, S. J.; Gaillard, J.-F.; Giedroc, D. P.; Collins, J. J.; Lucks, J. B. Cell-Free Biosensors for Rapid Detection of Water Contaminants. *Nat. Biotechnol.* **2020**, *38* (12), 1451–1459.
- (11) Zhang, Y.; Steppe, P. L.; Kazman, M. W.; Styczynski, M. P. Point-of-Care Analyte Quantification and Digital Readout via Lysate-Based Cell-Free Biosensors Interfaced with Personal Glucose Monitors. *ACS Synth. Biol.* **2021**, *10* (11), 2862–2869.
- (12) Karlukow, M.; da Silva, S. J. R.; Guo, Y.; Cicek, S.; Krokovsky, L.; Homme, P.; Xiong, Y.; Xu, T.; Calderón-Peláez, M.-A.; Camacho-Ortega, S.; Ma, D.; de Magalhães, J. J. F.; Souza, B. N. R. F.; de Albuquerque Cabral, D. G.; Jaenes, K.; Sutyrina, P.; Ferrante, T.; Benítez, A. D.; Nipaz, V.; Ponce, P.; Rackus, D. G.; Collins, J. J.; Paiva, M.; Castellanos, J. E.; Cevallos, V.; Green, A. A.; Ayres, C.; Pena, L.; Pardee, K. Field Validation of the Performance of Paper-Based Tests for the Detection of the Zika and Chikungunya Viruses in Serum Samples. *Nat. Biomed. Eng.* **2022**, *6* (3), 246–256.
- (13) Pardee, K.; Slomovic, S.; Nguyen, P. Q.; Lee, J. W.; Donghia, N.; Burrill, D.; Ferrante, T.; McSorley, F. R.; Furuta, Y.; Vernet, A.; Lewandowski, M.; Boddy, C. N.; Joshi, N. S.; Collins, J. J. Portable, On-Demand Biomolecular Manufacturing. *Cell* **2016**, *167* (1), 248–259.e12.
- (14) Stark, J. C.; Jaroentomeechai, T.; Moeller, T. D.; Herschewitz, J. M.; Warfel, K. F.; Moricz, B. S.; Martini, A. M.; Dubner, R. S.; Hsu, K. J.; Stevenson, T. C.; Jones, B. D.; DeLisa, M. P.; Jewett, M. C. On-Demand Biomanufacturing of Protective Conjugate Vaccines. *Sci. Adv.* **2021**, *7* (6), No. eabe9444.
- (15) Shin, J.; Jardine, P.; Noireaux, V. Genome Replication, Synthesis, and Assembly of the Bacteriophage T7 in a Single Cell-Free Reaction. *ACS Synth. Biol.* **2012**, *1* (9), 408–413.
- (16) Lavickova, B.; Laohakunakorn, N.; Maerkl, S. J. A Partially Self-Regenerating Synthetic Cell. *Nat. Commun.* **2020**, *11* (1), 6340.
- (17) Godino, E.; López, J. N.; Zarguit, I.; Doerr, A.; Jimenez, M.; Rivas, G.; Danelon, C. Cell-Free Biogenesis of Bacterial Division Proto-Rings That Can Constrict Liposomes. *Commun. Biol.* **2020**, *3* (1), 1–11.
- (18) Okauchi, H.; Ichihashi, N. Continuous Cell-Free Replication and Evolution of Artificial Genomic DNA in a Compartmentalized Gene Expression System. *ACS Synth. Biol.* **2021**, *10* (12), 3507–3517.
- (19) Garenne, D.; Beisel, C. L.; Noireaux, V. Characterization of the All-E. Coli Transcription-Translation System myTXTL by Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2019**, *33* (11), 1036–1048.
- (20) Miguez, A. M.; Zhang, Y.; Piorino, F.; Styczynski, M. P. Metabolic Dynamics in Escherichia Coli-Based Cell-Free Systems. *ACS Synth. Biol.* **2021**, *10* (9), 2252–2265.
- (21) Garamella, J.; Marshall, R.; Rustad, M.; Noireaux, V. The All E. Coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS. Synth. Biol.* **2016**, *5* (4), 344–355.
- (22) Sun, Z. Z.; Hayes, C. A.; Shin, J.; Caschera, F.; Murray, R. M.; Noireaux, V. Protocols for Implementing an Escherichia Coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Visualized Exp.* **2013**, *79*, No. e50762.
- (23) Matsuura, T.; Tanimura, N.; Hosoda, K.; Yomo, T.; Shimizu, Y. Reaction Dynamics Analysis of a Reconstituted Escherichia Coli Protein Translation System by Computational Modeling. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (8), E1336–E1344.
- (24) Jurado, Z.; Pandey, A.; Murray, R. M.; A Chemical Reaction Network Model of PURE. *bioRxiv* August 14, **2023**.
- (25) Lavickova, B.; Maerkl, S. J. A Simple, Robust, and Low-Cost Method To Produce the PURE Cell-Free System. *ACS Synth. Biol.* **2019**, *8* (2), 455–462.

- (26) Stölke, J.; Hillen, W. Carbon Catabolite Repression in Bacteria. *Curr. Opin. Microbiol.* **1999**, 2 (2), 195–201.
- (27) Grasemann, L.; Lavickova, B.; Elizondo-Cantú, M. C.; Maerkl, S. J. OnePot PURE Cell-Free System. *J. Visualized Exp.* **2021**, 172, No. e62625.
- (28) Jurado, Z.; Murray, R. M. Impact of Chemical Dynamics of Commercial PURE Systems on Malachite Green Aptamer Fluorescence. *ACS Synth. Biol.* **2024**, 13 (10), 3109–3118.
- (29) Kapasiawala, M.; Murray, R. M. Metabolic Perturbations to an Escherichia Coli-Based Cell-Free System Reveal a Trade-off between Transcription and Translation. *ACS Synth. Biol.* **2024**, 13 (12), 3976–3990.
- (30) *Nucleus*. Nucleus. <https://nucleus.bnnext.bio>.
- (31) Grodberg, J.; Dunn, J. J. *ompT* Encodes the Escherichia Coli Outer Membrane Protease That Cleaves T7 RNA Polymerase during Purification. *J. Bacteriol.* **1988**, 170 (3), 1245–1253.
- (32) Vandepitte-Rutten, L.; Kramer, R. A.; Kroon, J.; Dekker, N.; Egmond, M. R.; Gros, P. Crystal Structure of the Outer Membrane Protease OmpT from Escherichia Coli Suggests a Novel Catalytic Site. *EMBO J.* **2001**, 20 (18), 5033–5039.
- (33) Brannon, J. R.; Burk, D. L.; Leclerc, J.-M.; Thomassin, J.-L.; Portt, A.; Berghuis, A. M.; Gruenheid, S.; Le Moual, H. Inhibition of Outer Membrane Proteases of the OmpT Family by Aprotinin. *Infect. Immun.* **2015**, 83 (6), 2300–2311.
- (34) McCarter, J. D.; Stephens, D.; Shoemaker, K.; Rosenberg, S.; Kirsch, J. F. Substrate Specificity of the Escherichia Coli Outer Membrane Protease OmpT. *J. Bacteriol.* **2004**, 186 (17), 5919–5925.
- (35) MEROPS—the Peptidase Database. <https://www.ebi.ac.uk/merops/cgi-bin/pepsum?id=A26.001&type=P> (accessed 2024-10-20).
- (36) Hwang, B.-Y.; Varadarajan, N.; Li, H.; Rodriguez, S.; Iverson, B. L.; Georgiou, G. Substrate Specificity of the Escherichia Coli Outer Membrane Protease OmpP. *J. Bacteriol.* **2007**, 189 (2), 522.
- (37) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savv, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, 630 (8016), 493–500.
- (38) Muller, D. K.; Martin, C. T.; Coleman, J. E. Processivity of Proteolytically Modified Forms of T7 RNA Polymerase. *Biochemistry* **1988**, 27 (15), 5763–5771.
- (39) Ikeda, R. A.; Richardson, C. C. Enzymatic Properties of a Proteolytically Nicked RNA Polymerase of Bacteriophage T7. *J. Biol. Chem.* **1987**, 262 (8), 3790–3799.
- (40) 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 (8), 751–755.
- (41) Yadav, S.; Perkins, A. J. P.; Liyanagedera, S. B. W.; Bougas, A.; Laohakunakorn, N. ATP Regeneration from Pyruvate in the PURE System. *bioRxiv* September 8, 2024.
- (42) Bicknell, A. A.; Reid, D. W.; Licata, M. C.; Jones, A. K.; Cheng, Y. M.; Li, M.; Hsiao, C. J.; Pepin, C. S.; Metkar, M.; Levinsky, Y.; Fritz, B. R.; Andrianova, E. A.; Jain, R.; Valkov, E.; Köhrer, C.; Moore, M. J. Attenuating Ribosome Load Improves Protein Output from mRNA by Limiting Translation-Dependent mRNA Decay. *Cell Rep.* **2024**, 43 (4), No. 114098.
- (43) Shin, J.; Noireaux, V. Efficient Cell-Free Expression with the Endogenous E. Coli RNA Polymerase and Sigma Factor 70. *J. Biol. Eng.* **2010**, 4 (1), 8.
- (44) Doerr, A.; Foschepoth, D.; Forster, A. C.; Danelon, C. In Vitro Synthesis of 32 Translation-Factor Proteins from a Single Template Reveals Impaired Ribosomal Processivity. *Sci. Rep.* **2021**, 11 (1), 1898.
- (45) Farasat, I.; Kushwaha, M.; Collens, J.; Easterbrook, M.; Guido, M.; Salis, H. M. Efficient Search, Mapping, and Optimization of Multi-protein Genetic Systems in Diverse Bacteria. *Mol. Syst. Biol.* **2014**, 10 (6), 731.
- (46) Hossain, A.; Lopez, E.; Halper, S. M.; Cetnar, D. P.; Reis, A. C.; Strickland, D.; Klavins, E.; Salis, H. M. Automated Design of Thousands of Nonrepetitive Parts for Engineering Stable Genetic Systems. *Nat. Biotechnol.* **2020**, 38 (12), 1466–1475.
- (47) Halper, S. M.; Hossain, A.; Salis, H. M. Synthesis Success Calculator: Predicting the Rapid Synthesis of DNA Fragments with Machine Learning. *ACS Synth. Biol.* **2020**, 9 (7), 1563–1571.
- (48) Cetnar, D. P.; Salis, H. M. Systematic Quantification of Sequence and Structural Determinants Controlling mRNA Stability in Bacterial Operons. *ACS Synth. Biol.* **2021**, 10 (2), 318–332.
- (49) LaFleur, T. L.; Hossain, A.; Salis, H. M. Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria. *Nat. Commun.* **2022**, 13 (1), 5159.
- (50) Kurylo, C. M.; Alexander, N.; Dass, R. A.; Parks, M. M.; Altman, R. A.; Vincent, C. T.; Mason, C. E.; Blanchard, S. C. Genome Sequence and Analysis of Escherichia Coli MRE600, a Colicinogenic, Nonmotile Strain That Lacks RNase I and the Type I Methyltransferase, EcoKI. *Genome Biol. Evol.* **2016**, 8 (3), 742–752.
- (51) Archer, C. T.; Kim, J. F.; Jeong, H.; Park, J. H.; Vickers, C. E.; Lee, S. Y.; Nielsen, L. K. The Genome Sequence of E. Coli W (ATCC 9637): Comparative Genome Analysis and an Improved Genome-Scale Reconstruction of E. Coli. *BMC Genomics* **2011**, 12 (1), 9.
- (52) Dong, H.; Nilsson, L.; Kurland, C. G. Co-Variation of tRNA Abundance and Codon Usage in Escherichia Coli at Different Growth Rates. *J. Mol. Biol.* **1996**, 260 (5), 649–663.
- (53) Wei, Y.; Silke, J. R.; Xia, X. An Improved Estimation of tRNA Expression to Better Elucidate the Coevolution between tRNA Abundance and Codon Usage in Bacteria. *Sci. Rep.* **2019**, 9 (1), 3184.
- (54) Maheshwari, A. J.; Calles, J.; Waterton, S. K.; Endy, D. Engineering tRNA Abundances for Synthetic Cellular Systems. *Nat. Commun.* **2023**, 14 (1), 4594.



CAS BIOFINDER DISCOVERY PLATFORM™

## BRIDGE BIOLOGY AND CHEMISTRY FOR FASTER ANSWERS

Analyze target relationships,  
compound effects, and disease  
pathways

Explore the platform

