ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

Reconstitution of Multi-Protein Complexes through Ribozyme-Assisted Polycistronic Co-Expression

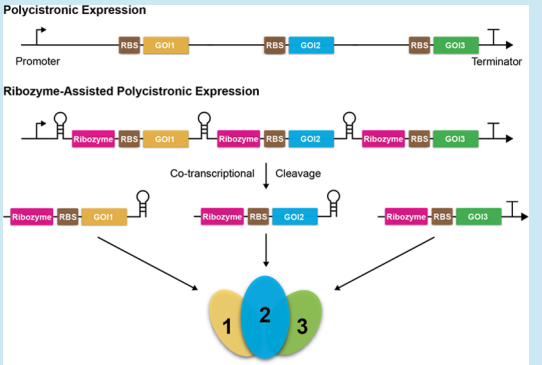
Yan Liu, Zihan Wu, Damu Wu, Ning Gao, and Jinzhong Lin*

Cite This: *ACS Synth. Biol.* 2023, 12, 136−143

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In living cells, proteins often exert their functions by interacting with other proteins forming protein complexes. Obtaining homogeneous samples of protein complexes with correct fold and stoichiometry is critical for its biochemical and biophysical characterization as well as functional investigation. Here, we developed a Ribozyme-Assisted Polycistronic co-expression system (pRAP) for heterologous co-production and in vivo assembly of multi-subunit complexes. In the pRAP system, a polycistronic mRNA transcript is co-transcriptionally converted into individual monocistrons in vivo. Each cistron can initiate translation with comparable efficiency, resulting in balanced production for all subunits, thus permitting faithful protein complex assembly. With pRAP polycistronic co-expression, we have successfully reconstituted large functional multi-subunit complexes involved in mammalian translation initiation. Our invention provides a valuable tool for studying the molecular mechanisms of biological processes.



KEYWORDS: *co-expression, polycistronic, pRAP, reconstitution, multi-protein complex*

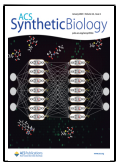
INTRODUCTION

In vitro reconstitution of biological processes, such as mRNA translation, can be critical to aid our research toward understanding these systems. The first step to reconstitute a cellular process is to have access to its functional constituents, which, in many cases, are multi-subunit protein complexes working as nano-machines. However, obtaining a sufficient amount of a multi-protein complex is a non-trivial task and remains challenging. Direct extraction from the native source is often infeasible as a protein complex usually exists at low abundance.[1] Researchers have turned to recombinant production in heterologous host cells, and many strategies to reconstitute a hetero-oligomeric protein complex have been developed.[2] One straightforward approach is to individually express and purify each component and assemble it in vitro into protein complexes. This approach suffers from many drawbacks. For example, protein subunits tend to fold improperly and aggregate when expressed alone.[3]

In vivo reconstitution by co-expression of multiple subunits has gained particular attention as it mimics endogenous protein complex assembly and has a higher chance of obtaining a functional sample.[4] Co-expression is often achieved by co-transforming two or more plasmids, each carrying the gene of one subunit and a different selection marker, into the heterologous expression host.[5] However, the number of plasmids used is not unlimited, as it is challenging to transform and maintain more than four plasmids simultaneously. Furthermore, the uneven copy numbers of each plasmid make protein expression unpredictable. Alternatively, many efforts have been made to assemble multiple genes into one plasmid, in which genes are transcribed from either multiple promotors[6] or one single promotor, creating a long polycistronic mRNA.[7] In the multi-promoter approach, the expression of each gene is governed by its own transcriptional and translational control. The level of protein synthesis can be tuned at the transcriptional level by using a combination of promoters of different strengths.[8] In the one-promoter approach, all genes are organized into one operon transcribed into a polycistronic mRNA. The open reading frame (ORF) for each protein is preceded by an RNA element serving for the ribosome binding site (RBS), to which the small ribosomal subunit binds and initiates translation. In bacteria, the RBS is a short motif called the Shine−Dalgarno (SD) sequence,[9] whereas in mammals, an internal ribosome entry site (IRES) derived from viruses is used to initiate the translation of its downstream ORF.[10]

Compared to the multi-promoter strategy, the polycistronic co-expression has many unique advantages. It requires minimal cargo DNA to deliver multiple genes and thus is frequently

Received: August 1, 2022
Published: December 13, 2022



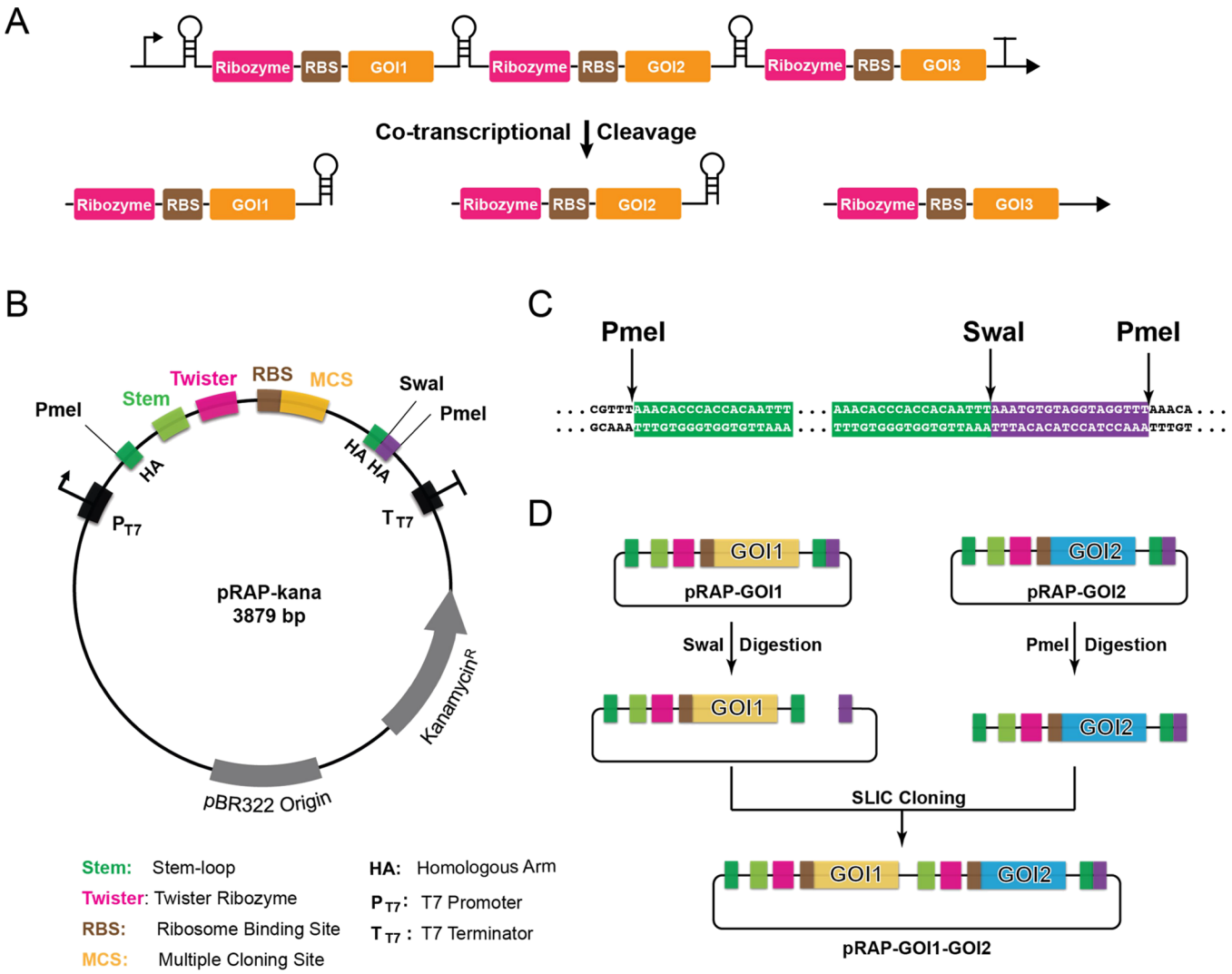


Figure 1. Design of the ribozyme-assisted polycistronic co-expression. (A) Scheme of the expression unit. A long polycistronic mRNA is co-transcriptionally cleaved into individual mRNAs for each cistron, a process mediated by the cis-acting twister ribozyme. (B) Vector map of a pRAP plasmid. The pRAP-kana plasmid with a kanamycin selection marker is shown. The pRAP-amp plasmid carries an ampicillin selection marker instead. (C) Partial sequence from the plasmid highlighting the *SwaI-PmeI* cloning module. Cleavage sites by the *SwaI* and *PmeI* restriction enzymes are indicated. (D) Assembly of multiple genes into pRAP for polycistronic co-expression. To do so, each GOI is first cloned into a pRAP construct. As illustrated, two such constructs, pRAP-GOI1 and pRAP-GOI2, are digested by *SwaI* and *PmeI*, respectively. pRAP-GOI1 cut by *SwaI* serves as the acceptor plasmid, and the donor gene GOI2 comes from *PmeI* digestion of the pRAP-GOI1 construct. A SLIC reaction between the donor and acceptor creates a new co-expression construct pRAP-GOI1-GOI2.

used in gene therapy to deliver two or more therapeutical genes simultaneously.[11] In addition, since all genes are transcriptionally coupled, polycistronic expression provides reproducible ratios of the synthesized proteins, mitigating the effect on protein expression of the fluctuating transcription associated with multiple promoters.[12] However, polycistronic expression suffers from a major problem: translations of cistrons are severely unbalanced. The order of cistrons in the polycistronic mRNA has an unpredictable consequence on the levels of protein expression.[11] Progressively lower expression was reported with more downstream cistrons in the mRNA.[13]

In this work, we have developed a novel expression system for in vivo reconstitution of multi-protein supercomplexes by coupling polycistronic transcription with post-transcriptional processing mediated by self-cleaving ribozymes. In this Ribozyme-Assisted Polycistronic expression system (pRAP), multiple genes are concatenated into a polycistronic form with intervening cis-acting ribozymes and placed under the control of one single promoter. The ribozyme enables co-transcriptional processing through its cleavage activity, generating monocistronic mRNAs for each gene that can undergo independent translation. This new system generates equal amounts of mRNAs for each gene with a comparable translation initiation rate to ensure balanced expression for all protein subunits. A set of pRAP plasmids was constructed with different selection markers so that two or more such plasmids could be used jointly to express a virtually unlimited number of proteins. We employed a standardized cloning procedure to simplify the assembly of genes into the vector. To prove the usefulness of the pRAP system, we have used it to co-express and reconstitute two human translation factors, the heterotrimeric eEF1B complex and the 9-subunit core of the eIF3 complex.

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143

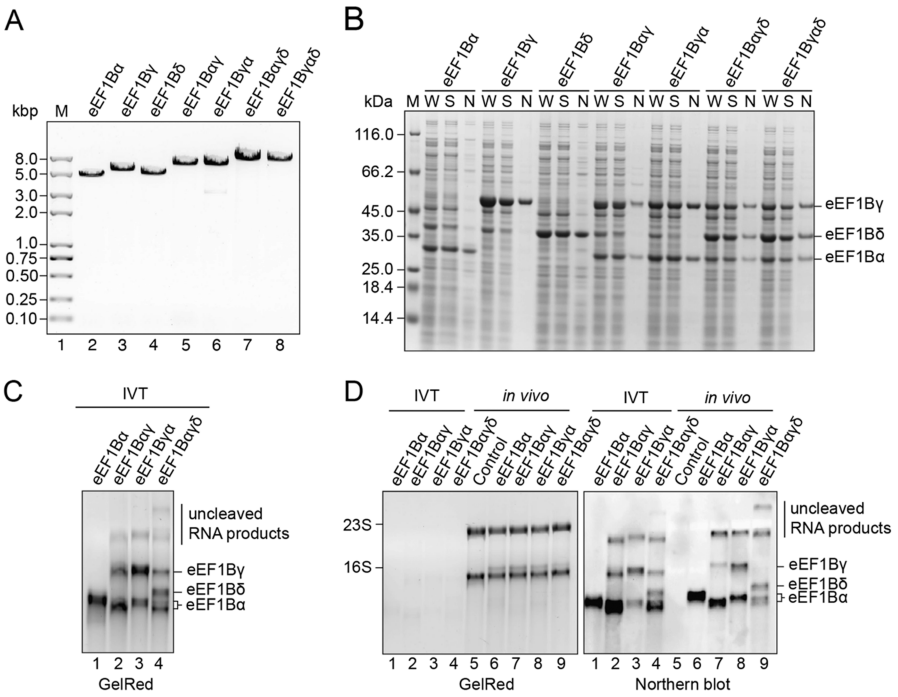


Figure 2. Proof of principle of co-expression mediated by pRAP constructs. (A) Agarose gel analysis of pRAP constructs expressing eEF1B subunits (eEF1Bα, eEF1Bγ, and eEF1Bδ), subcomplexes (eEF1B$\alpha\gamma$ and eEF1B$\gamma\alpha$), and full complexes (eEF1B$\alpha\gamma\delta$ and eEF1B$\gamma\alpha\delta$). All constructs were linearized by *SwaI* digestion, followed by gel electrophoresis and staining with GelRed. DNA ladders (lane M) are shown to the left. (B) Coomassie blue-stained SDS-PAGE of samples of the whole-cell lysate (W), soluble fraction after cell lysis and clarification (S), and elute from nickel-chelating beads (N), following protein overexpression by pRAP constructs in *E. coli* cells. (C) Denaturing agarose gel analysis of RNA cistrons synthesized by T7-mediated in vitro transcription (IVT) using pRAP constructs as templates. (D) Northern blot analysis of IVT (lanes 1−4) and in vivo (lanes 6−9) RNAs. In vivo total RNAs were isolated from *E. coli* cells transformed with eEF1B pRAP constructs and induced with IPTG. Left: GelRed-stained agarose gel before membrane transfer. Right: Northern blot image detected with a biotin-labeled DNA probe complementary to the twister ribozyme. Control: total RNAs isolated from uninduced cells. Note: Due to the extra sequences carried from the transcription terminator, the eEF1Bα cistron is slightly larger when it is located at the end.

RESULTS

Vector Design. To create a self-cleavable polycistronic transcript, we placed a sequence encoding a self-cleaving ribozyme in-between each gene. Following transcription, ribozyme cleavage would liberate each gene from the transcript into independent monocistronic mRNAs, each with the same 5′ UTR (untranslated region) containing an RBS in front of the ORF (Figure 1A and Supplementary Figure S1). We screened several ribozymes and chose the recently discovered twister ribozyme as it delivers the best activity both in vitro and in vivo. We placed twister in front of the RBS, and since twister cleaves itself at its 5′ terminus, all liberated cistrons begin with a twister ribozyme followed by the RBS and ORF. We also introduced a stem−loop before the ribozyme. As a result, all cistrons end with a stem−loop structure, except for the last one, which carries a native stem−loop derived from the transcription termination. The stem structure plays a critical role in stabilizing the monocistronic mRNA (see the Discussion section).

The final vector map is shown in Figure 1B; we made two empty pRAP constructs in this study, pRAP-kana and pRAP-amp, carrying the kanamycin and ampicillin selection markers, respectively. The two constructs can be used in combination, if needed, to reconstitute protein supercomplexes containing many subunits.

To simplify the cloning procedure, we designed a *SwaI-PmeI* module for the sequential assembly of multiple genes into the pRAP construct (Figure 1C). To create a polycistronic pRAP expression vector encoding a multi-component protein complex, the ORF from each gene of interest (GOI) was first cloned into the multiple clone site of pRAP to create a standalone expression construct. For assembly of them into one construct, pRAP-GOI1 was linearized by the *SwaI* restriction enzyme, while the other, pRAP-GOI2, was digested with *PmeI* restriction enzyme to liberate the expression unit for GOI2, and homologous recombination between linearized pRAP-GOI1 and GOI2 would create a new di-cistronic vector for two genes while restoring the *SwaI-PmeI* module (Figure 1D). A repeat of this procedure would introduce more genes into one construct.

Proof of Concept of pRAP. To put pRAP to the test, we used it to reconstitute the human eEF1B complex. eEF1B is a eukaryotic translation factor responsible for exchanging guanine nucleotides on eEF1A. eEF1B consists of three subunits: α, γ, and δ. Previous studies using size exclusion chromatography (SEC) have shown that eEF1Bδ forms high molecular weight aggregates when expressed alone, which could be alleviated by incubating it with purified eEF1B$\alpha\gamma$ subcomplex.[14,15] The exact physiological stoichiometry and architecture of eEF1B remain elusive.[16] To reconstitute eEF1B in vivo, we first cloned each gene into the pRAP plasmid to create pRAP-eEF1Bα, pRAP-eEF1Bγ, and pRAP-eEF1Bδ, respectively. We then constructed two di-cistronic vectors, pRAP-eEF1B$\alpha\gamma$ and pRAP-eEF1B$\gamma\alpha$, to express the $\alpha\gamma$ subcomplex. These two vectors differ in the order of arrangement of α and γ genes in the transcript, which would

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143

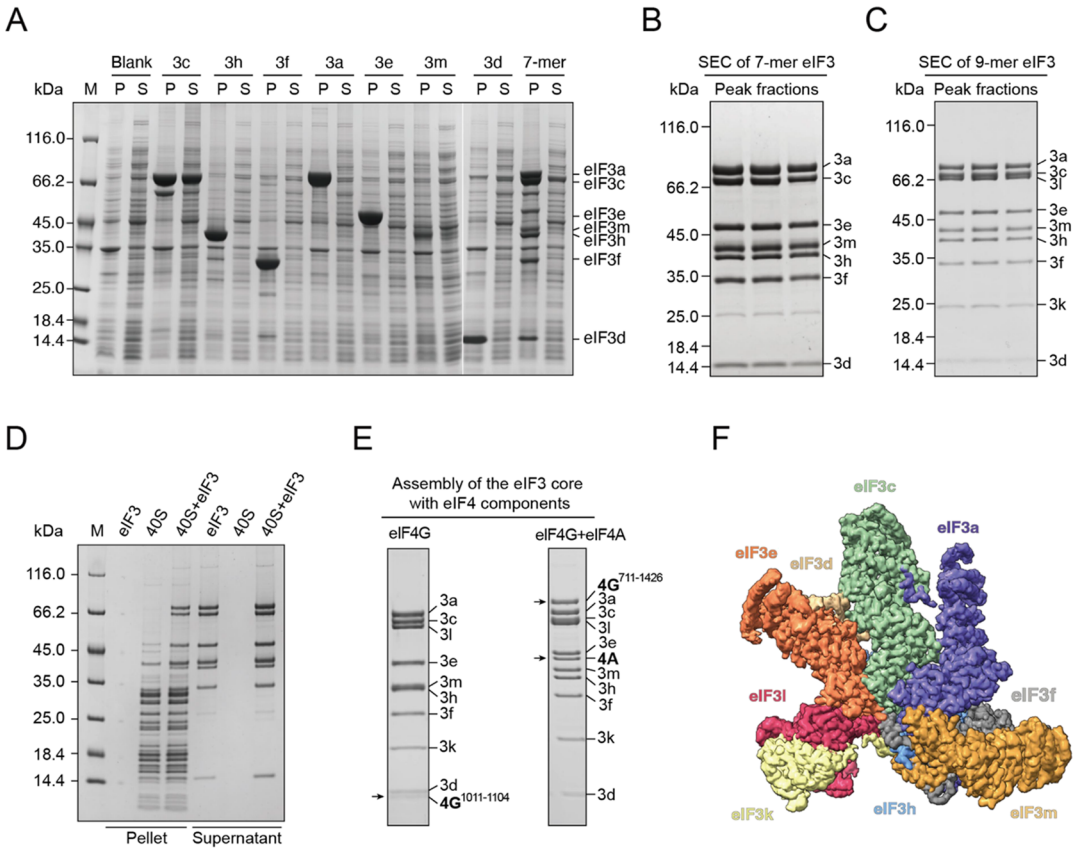


Figure 3. Reconstitution and functional characterization of the human eIF3 complex expressed from pRAP constructs. (A) Coomassie blue-stained SDS-PAGE of recombinant protein production from pRAP constructs expressing or co-expressing eIF3 subunits. Following induction of overexpression, the *E. coli* cells were lysed and clarified by centrifugation. The pellet (P) and supernatant (S) were taken for analysis. Protein molecular markers (M) are shown to the left. (B) Co-expression and purification of the 7-mer eIF3 subcomplex by size exclusion chromatography (SEC). Shown is the SDS-PAGE analysis of the peak fractions containing all seven subunits of eIF3. The complete chromatography profiles are provided in Supplemental Figure S3. (C) Reconstitution of the 9-mer eIF3 core complex. The 7-mer eIF3 subcomplex was mixed with the eIF3kl dimer before separation by SEC for the full 9-mer eIF3 core. Shown is the SDS-PAGE analysis of the SEC peak fractions. (D) Reconstituted eIF3 core binds to the human 40S ribosomal subunit. Binding reactions were pelleted through a sucrose cushion and analyzed by SDS-PAGE with Coomassie blue staining. (E) Interactions of the eIF3 core complex with eIF4F components. Shown on the left is a reconstituted complex of eIF3 with the eIF4G fragment (1011−1104). Right is the reconstituted complex eIF3 with the longer eIF4G (711−1426) and eIF4A. (F) Single-particle cryo-EM reconstruction of the 9-mer eIF3 core complex.

allow us to assess whether the gene position on the transcript affects their expression. Finally, we created pRAP-eEF1B$\alpha\gamma\delta$ and pRAP-eEF1B$\gamma\alpha\delta$ to express the entire complex in which an 8× His tag was added to the N-terminus of δ for downstream purification (Figure 2A).

All constructs (monocistronic and polycistronic) were tested for protein expression in *Escherichia coli (E. coli)*. Proteins were expressed and affinity-purified using nickel-chelating beads. Samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) stained with Coomassie blue, as shown in Figure 2B. All subunits showed overexpression either in their standalone constructs or in the polycistronic ones. Notably, the expression levels of eEF1Bδ in the polycistronic constructs were comparable to that in the standalone construct, while eEF1Bα and eEF1Bγ were expressed at a lower level in the polycistronic constructs (Supplementary Figure S2). This indicates that the genes in the polycistronic construct are translated as efficiently as in the monocistronic one. pRAP-eEF1B$\alpha\gamma$ and pRAP-eEF1B$\gamma\alpha$ produce an equivalent amount of $\alpha\gamma$ subcomplex, suggesting that the position of genes does not affect the protein expression. The final complex pulled down through the His-tag indicates that eEF1B α, γ, and δ subunits form a ternary complex.[17]

To verify that genes encoding eEF1B subunits were processed in vivo into monocistronic forms by ribozyme cleavage, we performed Northern blotting to visualize gene transcripts in cells. As a control, in vitro transcribed mRNA from these expression constructs was efficiently cleaved into monocistronic scripts (Figure 2C) corresponding to the size of each gene. The same cleavage pattern was observed with mRNAs extracted from cells transformed with polycistronic constructs (Figure 2D), suggesting that the twister ribozyme is functional in vivo in our invention.

pRAP Enables the Functional Reconstitution of the Human eIF3 Complex. Following the successful reconstitution of the eEF1B complex through pRAP co-expression, we next applied pRAP in a more difficult case, which is to reconstitute one of the most complex translation factors, the human eIF3 complex. Human eIF3 is a translation initiation factor consisting of up to 13 subunits (eIF3a−eIF3m). It directly interacts with the small 40S ribosomal subunit and 5′-UTR of cellular mRNAs, participating in the cap-dependent translation initiation.[18] It can also initiate the translation of

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143

some viral RNAs during infection and some cellular mRNAs under stress conditions through a cap-independent mechanism.[19,20] Deciphering the function of eIF3 has been central to understanding various mechanisms of translation initiation in eukaryotes. So far, the sample acquirement of eIF3 is mainly through extraction from endogenous sources such as human cells[21] and rabbit reticulocytes,[22] which is tedious and requires a great number of cells to obtain a meaningful amount of sample. Endogenous extraction also excludes the possibility of obtaining subcomplexes of eIF3, which is valuable for dissecting the function of individual eIF3 subunits. The main part of eIF3 constitutes a core, related to the proteasomal lid and COP9 signalosome through its PCI (Proteasome/CSN/eIF3)/MPN (Mpr1/Pad1 N-terminal) domain-containing subunits: eIF3a, eIF3c, eIF3e, eIF3f, eIF3h, eIF3k, eIF3l, and eIF3m.[23] The core also includes the N-terminal stretch of eIF3d, while the C-terminal domain of eIF3d and the rest subunits (eIF3b, eIF3i, eIF3g, and eIF3j) constitute the flexible peripheral part of eIF3. A previous study has shown that eIF3k and eIF3l form a dimer, which is not essential for the eIF3 integrity and cell viability.[24]

We began to reconstitute the 9-mer eIF3 core by cloning genes for each subunit into the pRAP-kana or the pRAP-amp vectors, each carrying a kanamycin and ampicillin selection marker, respectively. eIF3a (1−606), eIF3c (321−913), eIF3h (34−352), eIF3f (84−357), and eIF3k were individually cloned into pRAP-kana vector. eIF3e, eIF3m, and eIF3l and the N-terminal stretch of eIF3d (1−86) were each cloned into the pRAP-amp vector. The non-structural flexible parts of eIF3a, eIF3c, eIF3h, and eIF3f were not included. All these plasmids were tested for protein expression in *E. coli*. As shown in Figure 3A, all proteins were overexpressed. However, most proteins aggregated into inclusion bodies as expected. We then assembled eIF3a, eIF3c, eIF3h, and eIF3f into one polycistronic configuration and obtained the pRAP-eIF3acfh co-expression construct. eIF3e, eIF3m, and eIF3d were congregated into the pRAP-eIF3emd construct. We then coexpressed the 7-mer core by co-transforming pRAP-eIF3acfh and pRAP-eIF3emd. A significant portion of all subunits was detected in the supernatant of clarified cell lysates (Figure 3A, last two lanes). A stable complex was purified into homogeneity containing all seven subunits (Figure 3B and Supplementary Figure S3). Notably, we did not introduce fusion protein tags to help solubilize these eIF3 subunits as was done in previous reconstitution experiments.[25] The eIF3kl dimer was obtained separately by co-expression in *E. coli* with the two pRAP-eIF3k and pRAP-eIF3l constructs. A soluble eIF3kl was purified to homogeneity. The complete 9-mer eIF3 core was obtained by mixing purified 7-mer and the eIF3kl dimer, followed by separation with SEC (Figure 3C and Supplementary Figure S4).

To validate the function of the constituted eIF3 core complex, we performed several experiments. First, the purified eIF3 core was efficiently pelleted by the human ribosomal 40S subunit through a sucrose cushion, while eIF3 only could not be precipitated (Figure 3D). During translation initiation, eIF3 is recruited by the eIF4F complex consisting of three subunits, the cap-binding eIF4E, the helicase eIF4A, and the structural scaffold eIF4G. eIF3 interacts with the eIF4F complex through eIF4G, and the minimal part of eIF4G responsible for binding to eIF3 has been determined to locate between residues 1011 and 1104.[26] We purified two fragments of eIF4G, eIF4G (1101−1104), and eIF4G (711−1426); the latter contains the eIF4A-binding domain, in addition to the eIF3-binding region. SEC confirmed that eIF4G (1101−1104) could form a stable complex with our eIF3 core. Moreover, eIF4G (711−1426) was reconstituted into a super complex with eIF3 and eIF4A (Figure 3E and Supplementary Figure S3). Finally, we characterized the structure of the reconstituted eIF3 core using single-particle cryo-EM reconstruction. The overall structure from the density map was similar to the previously reported eIF3 in the context of the 43S preinitiation complex. However, the resolution was significantly improved owing to the homogeneity of the protein sample (Figure 3F). Notably, this represents the first atomic structure of eIF3 prior to ribosome binding.

DISCUSSION

Co-expression of proteins has been an indispensable technique in molecular and cellular biology, as well as in gene therapy, to deliver two or more functionally related genes simultaneously. Co-expression requires the co-delivery of multiple genes into host cells. These genes can either be independently transcribed into monocistronic mRNAs or share one transcriptional unit but have independent translation units sequentially arranged in one polycistronic mRNA. Each cistron contains a site for translation initiation, which can be an SD sequence in prokaryotes or an IRES in eukaryotes. One major problem associated with the polycistronic design is the uneven expression of genes at the translational level. The positional effect is evident as downstream cistrons produce considerably lower levels of protein than upstream ones.[13] The RBS in the internal cistron may be blocked by the secondary structure of the surrounding RNA, preventing the binding of the ribosome. A recent study showed that mRNAs from bacterial polycistronic operons form modular ORF-centric structures, with the sequence between ORFs being more exposed. This is suggested to be a result of the evolutionary selection of polycistronic operons to ensure the balanced expression of each gene.[27] However, it remains challenging to design a polycistronic mRNA sequence to include this feature computationally.

Here, we have overcome the obstacle of polycistronic co-expression by converting a polycistronic transcript into monocistronic mRNAs in vivo. This was achieved by placing self-cleaving ribozymes in between cistrons. These cis-acting ribozymes induce co-transcriptional cleavages, producing free-standing transcripts for each cistron. The resultant monocistronic mRNAs all have an RBS exposed at the 5′ end, the same as the first cistron in the polycistronic mRNA, and thus are freely accessible to the ribosome. In our examples of using pRAP to reconstitute two protein complexes, each tested protein was overexpressed to the same level in the polycistronic pRAP vector as in the standalone format, suggesting that the positional effect no longer exists.

In the pPAP system, we introduced a stem−loop structure at the end of each cistron. This design was critical as we found that, without the stem structure, only the last cistron was overexpressed while the preceding ones showed poor or even undetectable expression (data not shown). The last cistron differs from the preceding ones in its 3′ UTR after ribozyme cleavage, which contains additional sequences derived from the transcription terminator of RNA polymerase. These extra sequences form an RNA hairpin at the 3′ terminus of the final cistron. In contrast, the preceding cistrons ended with a short single-stranded fragment resulting from the ribozyme cleavage.

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143

RNA devoid of a 3′ terminal hairpin was reported to undergo rapid degradation.[28] This prompted us to redesign the construct by placing a stable stem−loop upstream of the cleavage site of the ribozyme. In doing so, all released cistrons have a stem−loop structure at the 3′ end, which indeed could restore protein expression for all genes.

Theoretically, an unlimited number of genes can be assembled into one pRAP construct for reliable co-expression. If necessary, two or more pRAP constructs can be used in concert to deliver more genes for reconstituting protein supercomplexes, as in the case of the eIF3 complex. In this study, we tried pRAP in bacteria, which has a limited capacity for post-translational modifications (PTMs). To reconstitute a eukaryotic complex with necessary PTMs, one can adapt pRAP by substituting the SD sequence with an IRES compatible with the eukaryotic ribosome to co-express proteins in mammalian cells. On the other hand, recombinant eukaryotic proteins obtained from bacteria could help delineate the functions of some important PTMs, for example, the methylation in eEF1A[29] and the diphthamide modification in eEF2.[30] Mammalian eIF3 also contains many PTMs at sub-stoichiometric levels whose physiological significance is unclear.[31] With PTM-free eIF3, we can investigate the functions of these PTMs by re-introducing them in a controlled manner.

Multi-protein complexes generally have a defined subunit stoichiometry. When co-expressed, the ability to tune the expression level of the subunit to match the stoichiometry is sometimes beneficial for the successful reconstitution of the complex. Previously, control of the protein synthesis could be achieved at the transcriptional level using different promotors of varied strengths.[8] In our pRAP system, all genes are transcriptionally coupled; expression can be fine-tuned at the translational level by modulating the translation initiation rate. In bacteria, the correlation of SD sequence variants with translation initiation has been comprehensively studied.[32] Libraries of SD sequences with predictable performance are available.[33] Correspondingly, more IRES elements, including mutants of varied strengths of translation initiation, are being established.[34] These RBSs can be adapted to pRAP to control the stoichiometry of co-expressed proteins.

MATERIALS AND METHODS

Sequence and Ligation-Independent Cloning (SLIC). The SLIC experiments were performed as previously described with some optimizations.[35] First, a 5X T4 DNA polymerase mix dedicated to SLIC was made by diluting 10 μL (10 U/μL) of T4 DNA polymerase (New England Labs) and 40 μL of 10× CutSmart buffer (New England Labs) in 30 μL of distilled water. For each SLIC experiment, a 2.5 μL reaction mixture was prepared by mixing 0.5 μL of linearized vector (50 ng), 1.5 μL of PCR insert (100−300 ng), and 0.5 μL of 5× T4 DNA polymerase mix, and placed at room temperature for 3 min before being incubated on ice for 10 min. The whole 2.5 μL reaction was then used to transform 50−100 μL of chemically competent DH5α cells to select the correct clones. DNA primers used in this study are listed in Supplementary Table S1.

Generation of pRAP Empty Constructs. The sequence of the expression unit containing the T7 promoter, stem−loop, twister ribozyme, *PmeI*/*SwaI* cloning module and the transcription terminator was synthesized commercially (Supplementary Figure S1). The pRAP-kana empty construct was derived from the pET-28b vector by substituting the original expression unit with the synthetic one. In short, the synthetic sequence was PCR amplified (primers: syn-RAP-F and syn-RAP-R), and the pET28 backbone was linearized by PCR amplification (primers: pET28b-backbone-F and pET28b-backbone-R). The two PCR products were gel-purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and combined using SLIC to create the pRAP-kana empty vector. The pRAP-amp vector was obtained by exchanging the kanamycin resistance cassette for an ampicillin resistance cassette. Briefly, the pRAP-kana vector was linearized by PCR amplification using primers pRAP-backbone-F and pRAP-backbone-R, and the ampicillin cassette was obtained from the pETDuet plasmid using primers Amp-F and Amp-R. The two fragments were ligated by SLIC to create the pRAP-amp vector.

Generation of pRAP Expression and Co-Expression Constructs. Protein coding sequences used in this study were obtained by either PCR amplification from the human cDNA library (for eEF1B subunits) or commercial whole-gene synthesis after codon optimization (for eIF3 and eIF4F subunits). The pRAP expression constructs for each gene were created by the standard SLIC method. To create a bi-cistronic pRAP expression construct, two pRAP expression constructs harboring each gene were treated separately. One pRAP construct was digested with the *SwaI* restriction enzyme. This linearized vector was ready to accept a new cistron, which was taken from the second construct either by *PmeI* digestion or PCR amplification with primers *PmeI*-F and *PmeI*-R. The linearized vector and insert were combined by the SLIC method to generate the co-expression construct. The same principle can be applied to introduce more cistrons into the pRAP construct.

Northern Blotting of In Vitro Transcription (IVT) or In Vivo RNAs. For IVT RNAs, the expression unit in the pRAP expression construct, from the T7 promoter to the terminator, was PCR amplified and used as a template for IVT according to the standard protocol. For in vivo RNAs, total RNAs were extracted from *E. coli* cells with the FastPure Total RNA Isolation Kit (Vazyme). The cells were transformed with the pRAP co-expression vector and induced with IPTG for at least 2 h before being harvested for RNA isolation. For Northern blotting, total RNAs (2 μg per lane) or IVT RNAs (0.1 μg per lane) were subjected to 1.2% formaldehyde agarose gel electrophoresis. The RNAs were then transferred to the nylon membrane and cross-linked to the membrane using a UV transilluminator (254 nm). The membrane with immobilized RNAs was then incubated in a 20 mL hybridization buffer containing 0.25 pmol/mL biotin-labeled DNA probes (5′-biotin-GCTATTTTTGCGGGCTTG-TAACCGCCCTCGGC-3′) that targets the twister ribozyme. The biotin-labeled probes were detected using a Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Scientific, USA). The results were visualized using the ChemiDoc Touch imaging system (Bio-Rad).

Protein Overexpression and Purification. Overexpression of protein or protein complexes was carried out in HI-Control BL21(DE3) *E. coli* cells. The transformed cells were grown in 50 mL of LB medium in the presence of appropriate antibiotics at 37 °C to an OD_{600} of 0.6 and then shifted to 18 °C prior to induction with 0.2 mM IPTG at 18 °C for 14 h. The cells were harvested and lysed by sonication. Cell lysates were clarified by centrifugation, and the pellet and supernatant

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143

were saved for later analysis. For purification, the His-tagged proteins in the supernatant were first immobilized on Ni-NTA agarose (GE Healthcare), washed with buffer A (20 mM HEPES pH 7.5, 300 mM KCl, 20 mM imidazole, 2 mM β-mercaptoethanol, and 5% glycerol), and then eluted with buffer B (20 mM HEPES pH 7.5, 300 mM KCl, 500 mM imidazole, 2 mM β-mercaptoethanol, and 5% glycerol). The 7-mer eIF3 was further purified by SEC.

Reconstitution and Functional Characterization of the eIF3 Core. The 9-mer eIF3 core complex was assembled by incubating the purified 7-mer eIF3 subcomplex and the eIF3kl dimer at a molar ratio of 1:2 for 30 min on ice. The 9-mer core was then separated by SEC on a Superose 6 Increase 10/300 GL column equilibrated with SEC buffer (20 mM HEPES-KOH, pH 8.0, 100 mM KCl, 2 mM $MgCl_2$). The sample was concentrated, flash-frozen in liquid nitrogen, and stored at −80 °C. The interaction assay of the eIF3 core with eIF4F components was done analogously on the same size exclusion column. Interaction between eIF3 and the 40S ribosomal subunit was assayed by the 40S pelleting experiment. Purified 40S was incubated with a fourfold molar excess of eIF3 for 30 min at room temperature. Mixtures were then loaded on a 30% sucrose cushion (20 mM HEPES-KOH, pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 4 mM DTT) and centrifuged at 50,000 rpm for 4 h using a TLA-55 rotor (Beckman) at 4 °C. Supernatants were carefully removed and saved, and the pellets were resuspended. Both the supernatant and pellet were analyzed by SDS-PAGE. The cryo-EM structure of the eIF3 core was determined as previously described.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.2c00416.

Oligonucleotides used in this study; sequence of the pRAP expression unit; expression analysis of eEF1B subunits in the pRAP constructs; size exclusion chromatography from the reconstitution experiments of eIF3; and mass spectrometry analysis of reconstituted eIF3 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jinzhong Lin − *State Key Laboratory of Genetic Engineering, School of Life Sciences, Zhongshan Hospital, Fudan University, Shanghai 200438, China;* orcid.org/0000-0001-5666-8526; Phone: +86-21-31246764; Email: linjinzhong@fudan.edu.cn

Authors

Yan Liu − *State Key Laboratory of Genetic Engineering, School of Life Sciences, Zhongshan Hospital, Fudan University, Shanghai 200438, China*

Zihan Wu − *State Key Laboratory of Genetic Engineering, School of Life Sciences, Zhongshan Hospital, Fudan University, Shanghai 200438, China*

Damu Wu − *State Key Laboratory of Membrane Biology, Peking-Tsinghua Joint Center for Life Sciences, School of Life Sciences, Peking University, Beijing 100871, China*

Ning Gao − *State Key Laboratory of Membrane Biology, Peking-Tsinghua Joint Center for Life Sciences, School of Life Sciences, Peking University, Beijing 100871, China;* orcid.org/0000-0003-3067-9993

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.2c00416

Author Contributions

J.L. and Y.L. conceived and designed the experiments. Y.L. and D.W. prepared the materials and performed the experiments. L.Y., Z.W., and D.W. analyzed the data. J.L. and Y.L. wrote the manuscript. All authors provided the final approval of the manuscript.

Funding

National Natural Science Foundation of China (32130063 to J.L.; 32200987 to Z.W.); China Postdoctoral Science Foundation (2022M720800 to Z.W.); Starry Night Science Fund at Shanghai Institute for Advanced Study of Zhejiang University (SNZJU-SIAS-009 to J.L.); Zhangjiang mRNA innovation and translation center, Shanghai; Innovation Program of Shanghai Municipal Education Commission (2021-01-07-0007-E00074 to J.L.)

Notes

The authors declare the following competing financial interest(s): J.L. has a patent pending for the pRAP co-expression system.

■ ACKNOWLEDGMENTS

We thank the Cryo-Electron Microscopy Facilities in the Multiscale Research Institute of Complex Systems and the School of Life Sciences at Fudan University.

■ REFERENCES

(1) Havugimana, P. C.; Hart, G. T.; Nepusz, T.; Yang, H.; Turinsky, A. L.; Li, Z.; Wang, P. I.; Boutz, D. R.; Fong, V.; Phanse, S.; et al. A census of human soluble protein complexes. *Cell* **2012**, *150*, 1068−1081.

(2) Romier, C.; Ben Jelloul, M.; Albeck, S.; Buchwald, G.; Busso, D.; Celie, P. H.; Christodoulou, E.; De Marco, V.; van Gerwen, S.; Knipscheer, P.; et al. Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 1232−1242.

(3) Li, C.; Schwabe, J. W. R.; Banayo, E.; Evans, R. M. Coexpression of nuclear receptor partners increases their solubility and biological activities. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 2278−2283.

(4) Kerrigan, J. J.; Xie, Q.; Ames, R. S.; Lu, Q. Production of protein complexes via co-expression. *Protein Expr. Purif.* **2011**, *75*, 1−14.

(5) Kholod, N.; Mustelin, T. Novel vectors for co-expression of two proteins in E. coli. *Biotechniques* **2001**, *31*, 326−328.

(6) Scheich, C.; Kummel, D.; Soumailakakis, D.; Heinemann, U.; Bussow, K. Vectors for co-expression of an unrestricted number of proteins. *Nucleic Acids Res.* **2007**, *35*, No. e43.

(7) Tan, S.; Kern, R. C.; Selleck, W. The pST44 polycistronic expression system for producing protein complexes in Escherichia coli. *Protein Expr. Purif.* **2005**, *40*, 385−395.

(8) Patel, Y. D.; Brown, A. J.; Zhu, J.; Rosignoli, G.; Gibson, S. J.; Hatton, D.; James, D. C. Control of Multigene Expression Stoichiometry in Mammalian Cells Using Synthetic Promoters. *ACS Synth. Biol.* **2021**, *10*, 1155−1165.

(9) Shine, J.; Dalgarno, L. Determinant of cistron specificity in bacterial ribosomes. *Nature* **1975**, *254*, 34−38.

(10) Pelletier, J.; Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **1988**, *334*, 320−325.

(11) Renaud-Gabardos, E.; Hantelys, F.; Morfoisse, F.; Chaufour, X.; Garmy-Susini, B.; Prats, A. C. Internal ribosome entry site-based

vectors for combined gene therapy. *World J. Exp. Med.* **2015**, *5*, 11−20.
(12) Sokolenko, S.; George, S.; Wagner, A.; Tuladhar, A.; Andrich, J. M.; Aucoin, M. G. Co-expression vs. co-infection using baculovirus expression vectors in insect cell culture: Benefits and drawbacks. *Biotechnol. Adv.* **2012**, *30*, 766−781.
(13) Kim, K. J.; Kim, H. E.; Lee, K. H.; Han, W.; Yi, M. J.; Jeong, J.; Oh, B. H. Two-promoter vector is highly efficient for overproduction of protein complexes. *Protein Sci.* **2004**, *13*, 1698−1703.
(14) Bec, G.; Kerjan, P.; Waller, J. P. Reconstitution in vitro of the valyl-tRNA synthetase-elongation factor (EF) 1 beta gamma delta complex. Essential roles of the NH2-terminal extension of valyl-tRNA synthetase and of the EF-1 delta subunit in complex formation. *J. Biol. Chem.* **1994**, *269*, 2086−2092.
(15) Sheu, G. T.; Traugh, J. A. Recombinant subunits of mammalian elongation factor 1 expressed in Escherichia coli. Subunit interactions, elongation activity, and phosphorylation by protein kinase CKII. *J. Biol. Chem.* **1997**, *272*, 33290−33297.
(16) Le Sourd, F.; Boulben, S.; Le Bouffant, R.; Cormier, P.; Morales, J.; Belle, R.; Mulner-Lorillon, O. eEF1B: At the dawn of the 21st century. *Biochim. Biophys. Acta* **2006**, *1759*, 13−31.
(17) Bondarchuk, T. V.; Shalak, V. F.; Lozhko, D. M.; Fatalska, A.; Szczepanowski, R. H.; Liudkovska, V.; Tsuvariev, O. Y.; Dadlez, M.; El'skaya, A. V.; Negrutskii, B. S. Quaternary organization of the human eEF1B complex reveals unique multi-GEF domain assembly. *Nucleic Acids Res.* **2022**, *50*, 9490.
(18) Jackson, R. J.; Hellen, C. U. T.; Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 113−127.
(19) Hashem, Y.; des Georges, A.; Dhote, V.; Langlois, R.; Liao, H. Y.; Grassucci, R. A.; Pestova, T. V.; Hellen, C. U.; Frank, J. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **2013**, *503*, 539−543.
(20) Meyer, K. D.; Patil, D. P.; Zhou, J.; Zinoviev, A.; Skabkin, M. A.; Elemento, O.; Pestova, T. V.; Qian, S. B.; Jaffrey, S. R. 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **2015**, *163*, 999−1010.
(21) Pestova, T. V.; Hellen, C. U.; Shatsky, I. N. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell. Biol.* **1996**, *16*, 6859−6869.
(22) Benne, R.; Hershey, J. W. Purification and characterization of initiation factor IF-E3 from rabbit reticulocytes. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 3005−3009.
(23) Hofmann, K.; Bucher, P. The PCI domain: a common theme in three multiprotein complexes. *Trends Biochem. Sci.* **1998**, *23*, 204−205.
(24) Wagner, S.; Herrmannova, A.; Sikrova, D.; Valasek, L. S. Human eIF3b and eIF3a serve as the nucleation core for the assembly of eIF3 into two interconnected modules: the yeast-like core and the octamer. *Nucleic Acids Res.* **2016**, *44*, 10772−10788.
(25) Sun, C.; Todorovic, A.; Querol-Audi, J.; Bai, Y.; Villa, N.; Snyder, M.; Ashchyan, J.; Lewis, C. S.; Hartland, A.; Gradia, S.; et al. Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 20473−20478.
(26) Villa, N.; Do, A.; Hershey, J. W.; Fraser, C. S. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, −d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* **2013**, *288*, 32932−32940.
(27) Burkhardt, D. H.; Rouskin, S.; Zhang, Y.; Li, G. W.; Weissman, J. S.; Gross, C. A. Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. *Elife* **2017**, *6*, No. e22037.
(28) Hajnsdorf, E.; Kaberdin, V. R. RNA polyadenylation and its consequences in prokaryotes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2018**, *373*, No. 20180166.
(29) Jakobsson, M. E.; Malecki, J. M.; Halabelian, L.; Nilges, B. S.; Pinto, R.; Kudithipudi, S.; Munk, S.; Davydova, E.; Zuhairi, F. R.; Arrowsmith, C. H.; et al. The dual methyltransferase METTL13 targets N terminus and Lys55 of eEF1A and modulates codon-specific translation rates. *Nat. Commun.* **2018**, *9*, 3411.
(30) Davydova, E.; Ho, A. Y. Y.; Malecki, J.; Moen, A.; Enserink, J. M.; Jakobsson, M. E.; Loenarz, C.; Falnes, P. O. Identification and characterization of a novel evolutionarily conserved lysine-specific methyltransferase targeting eukaryotic translation elongation factor 2 (eEF2). *J. Biol. Chem.* **2014**, *289*, 30499−30510.
(31) Andaya, A.; Villa, N.; Jia, W.; Fraser, C. S.; Leary, J. A. Phosphorylation stoichiometries of human eukaryotic initiation factors. *Int. J. Mol. Sci.* **2014**, *15*, 11523−11538.
(32) Mutalik, V. K.; Guimaraes, J. C.; Cambray, G.; Lam, C.; Christoffersen, M. J.; Mai, Q. A.; Tran, A. B.; Paull, M.; Keasling, J. D.; Arkin, A. P.; et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **2013**, *10*, 354−360.
(33) Kuo, S. T.; Jahn, R. L.; Cheng, Y. J.; Chen, Y. L.; Lee, Y. J.; Hollfelder, F.; Wen, J. D.; Chou, H. D. Global fitness landscapes of the Shine-Dalgarno sequence. *Genome Res.* **2020**, *30*, 711−723.
(34) Koh, E. Y. C.; Ho, S. C. L.; Mariati; Song, Z.; Bi, X.; Bardor, M.; Yang, Y. An internal ribosome entry site (IRES) mutant library for tuning expression level of multiple genes in mammalian cells. *PLoS One* **2013**, *8*, No. e82100.
(35) Stevenson, J.; Krycer, J. R.; Phan, L.; Brown, A. J. A practical comparison of ligation-independent cloning techniques. *PLoS One* **2013**, *8*, No. e83888.

https://doi.org/10.1021/acssynbio.2c00416
ACS Synth. Biol. 2023, 12, 136−143