ACS Synthetic Biology

pubs.acs.org/synthbio

Research Article

Clonal Amplification-Enhanced Gene Expression in Synthetic Vesicles

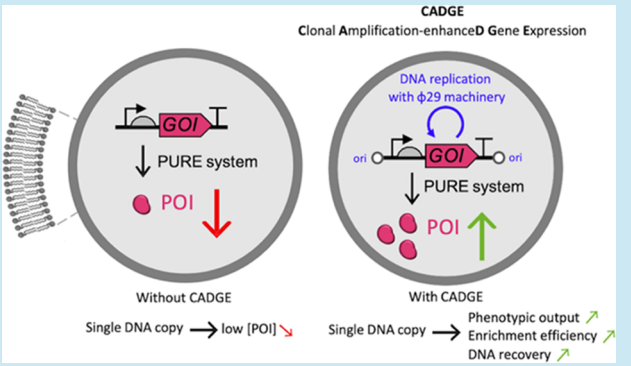
Zhanar Abil,[§] Ana Maria Restrepo Sierra,[§] and Christophe Danelon*

Cite This: *ACS Synth. Biol.* 2023, 12, 1187−1203

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In cell-free gene expression, low input DNA concentration severely limits the phenotypic output, which may impair in vitro protein evolution efforts. We address this challenge by developing CADGE, a strategy that is based on clonal isothermal amplification of a linear gene-encoding dsDNA template by the minimal Φ29 replication machinery and in situ transcription-translation. We demonstrate the utility of CADGE in bulk and in clonal liposome microcompartments to boost up the phenotypic output of soluble and membrane-associated proteins, as well as to facilitate the recovery of encapsulated DNA. Moreover, we report that CADGE enables the enrichment of a DNA variant from a mock gene library via either a positive feedback loop-based selection or high-throughput screening. This new biological tool can be implemented for cell-free protein engineering and the construction of a synthetic cell.



KEYWORDS: *synthetic cell, cell-free gene expression, directed evolution, DNA amplification*

INTRODUCTION

Inspired by natural selection, directed evolution has become a powerful tool in synthetic biology. This engineering approach encompasses cycles of genetic diversification and enrichment of rare desired variants, allowing for accelerated protein evolution even with limited a priori knowledge about the structure−function relationships.[1−3] Directed evolution enabled engineering of a plethora of proteins, genetic pathways, and even genomes to generate variants with improved or tailor-made properties.[4−11] Incorporation of directed evolution principles into the construction framework of a synthetic cell has recently been proposed.[12] Compartmentalized gene expression in liposomes[13−15] has gained considerable momentum in the last few years, with methodological advances that have improved the yield of functional vesicles,[16,17] enabling the reconstitution of complex biological functions, such as DNA replication,[18] phospholipid synthesis,[19] membrane deformation processes,[20−22] and light-triggered ATP synthesis.[23] Moving forward to optimizing and integrating cellular modules may require a system-level evolutionary approach.[12]

Over the past decades, numerous in vivo and cell-free methodologies for gene expression of the targeted phenotypes have been developed. In vivo methodologies have been the most common since a suitable host organism could provide low-cost gene expression with a reliable yield.[24] However, cell-free systems have emerged as an alternative and attractive platform due to the higher degree of controllability and freedom from the constrains related to cell survival.[25,26] Cell-free protein synthesis enabled engineering of a number of proteins, including membrane or cytotoxic proteins[27,28] as well as peptides and proteins that incorporate unnatural amino acids.[29−31] Cell-free protein expression can be accomplished using either cell lysates[32] or in a reconstituted transcription-translation system such as the PURE system.[33]

A prerequisite for directed evolution is a genotype-phenotype link. In cell-free systems, this link is often implemented through ribosomal, mRNA, or other cell-free macromolecular display technologies,[34] although these techniques are often limited to evolution of peptide and protein binding affinities. For evolution of an enzyme's catalytic turnover, however, compartmentalization in emulsion droplets[35] or liposomes[36] is more appropriate. Such biomimetic compartments are also often used as the chassis for engineering toward construction of an artificial cell.[37] Finally, liposomes are exceptionally suited for evolution of membrane proteins requiring a lipid bilayer for solubility and activity.[27]

However, coupling the gene expression and enrichment steps in a cell-free system within a microcompartment is often limited by the low yield of synthesized proteins from a single DNA

Received: December 17, 2022
Published: April 4, 2023

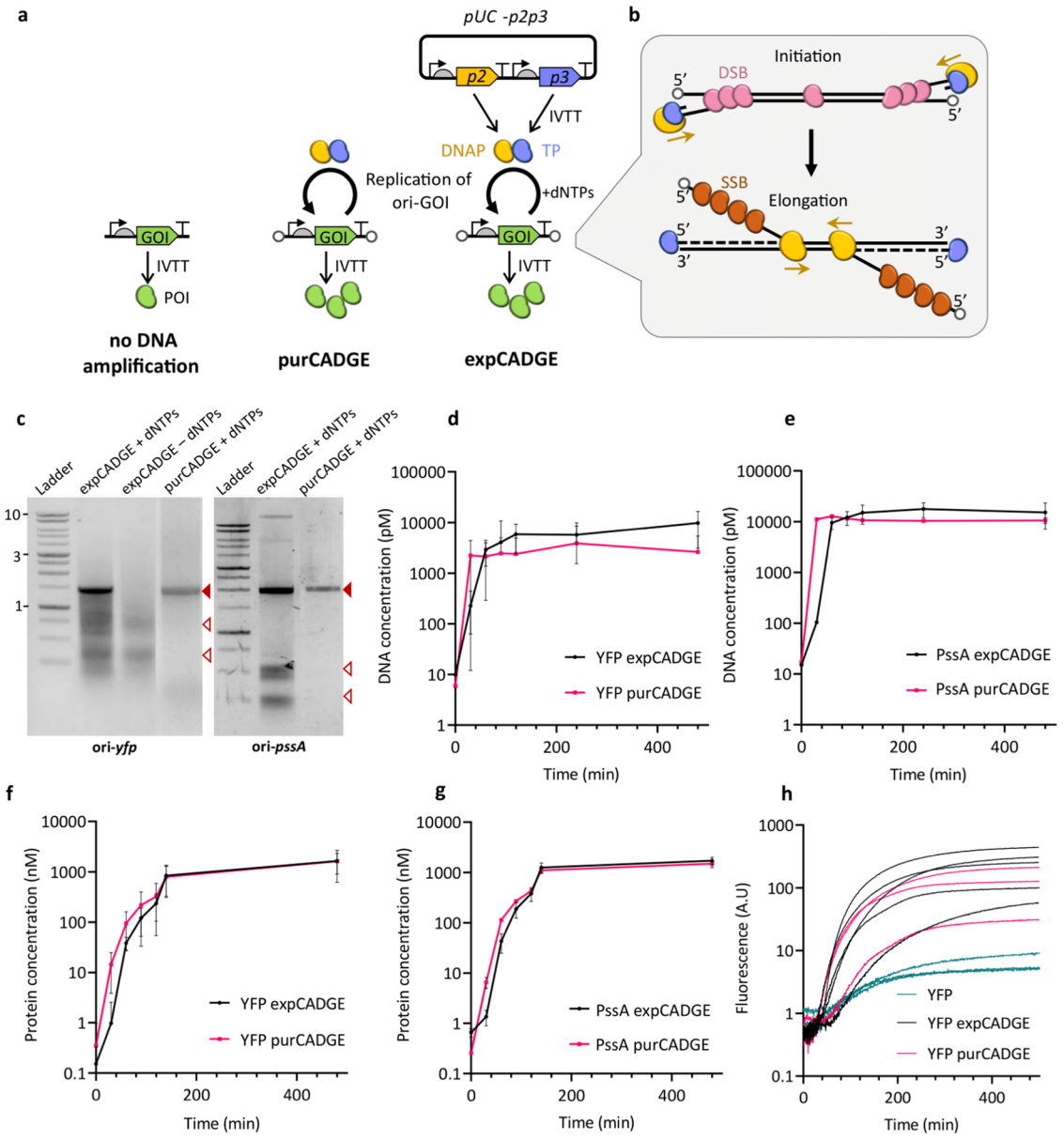


Figure 1. Principles and validation of the CADGE strategy in bulk reactions. (a) Schematics of the different gene expression configurations used in this study. (b) Schematic of linear DNA replication by the Φ29 minimal DNA replication machinery. (c) Visualization of amplified DNA in CADGE samples via agarose gel electrophoresis. Filled red arrowheads indicate expected product size, and empty red arrowheads indicate unfinished smaller products. (d, e) Time-course analysis of ori-*yfp* (d) and ori-*pssA* (e) DNA amplification in bulk CADGE reactions via absolute quantitative polymerase chain reaction (qPCR). (f, g) Time-course analysis of YFP and PssA expression in bulk CADGE reactions via absolute liquid chromatography and tandem mass spectrometry (LC-MS/MS) quantitation. (h) Kinetic measurements of YFP fluorescence from gene expression without ("YFP") and with CADGE. Different curves of same color are biological replicates.

template. Although detectable activity of cell-free expressed proteins arising from a single gene copy has been demonstrated in some experimental conditions,[27,38−40] it is hardly surprising that below a certain threshold, template DNA concentration is a limiting factor for in vitro protein expression.[15,41−44] In fact, production of full-length proteins in reconstituted systems ceases before NTPs and amino acids get depleted, and efforts to increase the amount of protein from low DNA concentrations remain frustrated.[45] Therefore, clonal amplification of expression templates is a generic solution to enhance protein yield and activity readout, as well as the recovery of selected DNA variants.

A major challenge in cell-free directed evolution is the coupling of DNA amplification from single template copies with gene expression and quantitation of the activity of the protein of interest for fitness assignment in *one* environment. For example, rolling circle amplification (RCA) based on the Φ29 DNA polymerase and replication cycle reaction (RCR) based on a reconstituted *Escherichia coli* replisome are compatible with droplet microcompartments.[28,46] However, RCR has not been combined with in vitro transcription-translation (IVTT) in a one-pot reaction yet, and polymerase chain reaction (PCR) requires heating steps that are incompatible with IVTT in one-pot reactions. On the other hand, combination of RCA with gene expression is only possible after optimization of some components for transcription and translation to minimize cross-inhibition effects,[47−50] proscribing the use of standard commercial kits for IVTT. Thus, so far, DNA templates cannot

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

be amplified efficiently in the same solution where the cell-free system is performed. Therefore, multiple-step workflows have been implemented, which require droplet-based microfluidic handling[28,51−53] or bead display.[54−58]

In this study, we simplify the in vitro evolution methodology by a single isothermal, clonal amplification-enhanced gene expression, or CADGE. The strategy relies on the protein-primed replication machinery of the *Bacillus subtilis* bacteriophage Φ29[59] consisting of DNA polymerase (DNAP, encoded by gene *p2*), terminal protein (TP, encoded by gene *p3*), double-stranded DNA-binding protein (DSB, encoded by gene *p6*), and single-stranded DNA-binding protein (SSB, encoded by gene *p5*), and requires prior flanking of the gene of interest (GOI) with Φ29 origins of replication (ori) using a standard recombinant DNA technique of choice. The Φ29 DNAP is chosen largely due to its strand-displacement activity, a relatively rare property for a family B DNA polymerase.[60,61] This activity enables it to displace the nontemplate DNA strand at ambient temperatures, thus ensuring compatibility with cell-free transcription-translation. In addition, Φ29 DNAP has an excellent processivity,[60,62] which could be useful for efficiently replicating long and multigene DNA templates. To initiate the replication, DNAP forms a complex with TP,[63] and the heterodimer is recruited to replication origins, a process that is facilitated by DSB.[64] DSB activates the replication initiation by forming a multimeric nucleoprotein complex at the origins of replication,[65] whereas TP primes the DNA synthesis at each end, remaining covalently attached to the 5′-end of the daughter strand.[66] After successful priming, DNAP dissociates from the complex and continues the polymerization activity.[67] SSB is another auxiliary protein, which assists in the replication by stabilizing the displaced strand.[68] Using this system, we previously realized transcription-translation-coupled self-replication of a two-gene construct.[18] Herein, we demonstrate that transcription-translation-coupled amplification of orthogonal genes can be achieved in bulk and in liposome compartments, improving the expression level of a GOI. As a proof of concept, we show the enrichment of an ori-*GOI* from a mock library encapsulated in liposomes, a key step toward cell-free protein evolution. Moreover, we apply CADGE to enable the screening of protein functions that are relevant in the field of synthetic cell construction.

RESULTS

Design of CADGE. The CADGE strategy involves the following minimal requirements (Figure 1a,b):

(1) A GOI is inserted between the 191-bp-oriL and 194-bp-oriR origins of replication of the Φ29 genome, although the 68-bp minimal origins could potentially be used as well.[69] The DNA template must be linearized with the origins at each end of the molecule, which can be achieved by PCR amplification from ori-containing plasmid DNA. Moreover, the linear DNA has to be phosphorylated at each 5′-end, which can be done by using 5′-phosphorylated primers. One, two,[18] or, in principle, more genes can be encoded on a single ori-flanked DNA template. Hereafter, we refer to such linear constructs as ori-*GOI*.

(2) The PURE*frex2.0* system is chosen for IVTT because of its higher purity and reduced nuclease activity compared to other commercial PURE systems.[45] Thus, the linear DNA construct contains regulatory elements compatible with gene expression in PURE*frex2.0*. These comprise a T7 promoter, g10 leader sequence, *E. coli* ribosome binding site, and a transcription terminator (e.g., T7 and vesicular stomatitis virus terminators).

(3) The system requires four minimal protein components of the Φ29 replication machinery: DNAP, TP, SSB, and DSB, plus dNTPs and ammonium sulfate for the efficient dimerization of the replication initiation complex[70] (Figure 1a,b). DNAP and TP can either be introduced in a purified form (purCADGE) or in situ expressed from a separate DNA construct (expCADGE). In the latter configuration, the two genes *p2* and *p3* are introduced on a single plasmid, self-replication being prohibited by the circular nature of the DNA. Although SSB and DSB can be functionally expressed in the PURE system,[18] we recommend supplying them as purified proteins since they are required at micromolar concentrations and their cell-free expression would create a burden on the transcription-translation apparatus. The linear replication product in CADGE is essentially identical to the parental DNA molecule—except for the fact that TP is covalently bound at the 5′-end of each daughter strand. In the current protocol, the 5′-TP is lost with subsequent PCR amplification during recovery of the total DNA from liposomes. Thus, the resulting recovered DNA is identical in its structure to the original template DNA and does not require any additional processing between rounds of evolution.

Validation of CADGE in Bulk Reactions. We first evaluated the performance of CADGE on amplifying a GOI in bulk reactions. Two ori-GOI fragments encoding either the enhanced yellow fluorescent protein (eYFP) or *Escherichia coli* phosphatidylserine synthase (PssA) (ori-*yfp* and ori-*pssA*, respectively) were constructed. The DNA template was added to PURE*frex2.0* in the presence of the DNA replication machinery, dNTPs, and ammonium sulfate, and the solution was incubated at 30 °C. With both purCADGE and expCADGE, we found that the full-length DNA (Figure 1c) can be amplified to saturation by two to three orders of magnitude from an input template concentration of 10 pM within 2 h, as confirmed by absolute quantification by qPCR (Figure 1d,e). Although full-length replication products are the most abundant DNA species, shorter fragments corresponding to incomplete polymerization products are also visible in the gel, especially with expCADGE (Figure 1c).

To test if template amplification is accompanied with an increase in protein expression levels, we quantified the concentrations of eYFP and PssA by liquid chromatography-coupled mass spectrometry (LC-MS), and confirmed the production of both proteins to up to 1 μM, with no noticeable differences between purCADGE and expCADGE (Figure 1f,g). These amounts of protein expression were comparable to the generally reported cell-free protein synthesis levels,[20,71,72] but with considerably (two to three orders of magnitude) less input of template DNA. The concentrations of YFP and PssA increased by at least 6-fold with expCADGE compared to unamplified gene conditions (Figures S1 and S2). Importantly, fluorescence kinetics measurements show that in the absence of DNA amplification, only a very low level of YFP was expressed even after several hours of incubation (Figure 1h). This finding indicates that the enhanced protein expression is the direct result of gene amplification. Kinetic analysis of protein synthesis gives apparent maximum translation rates (defined as the highest

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

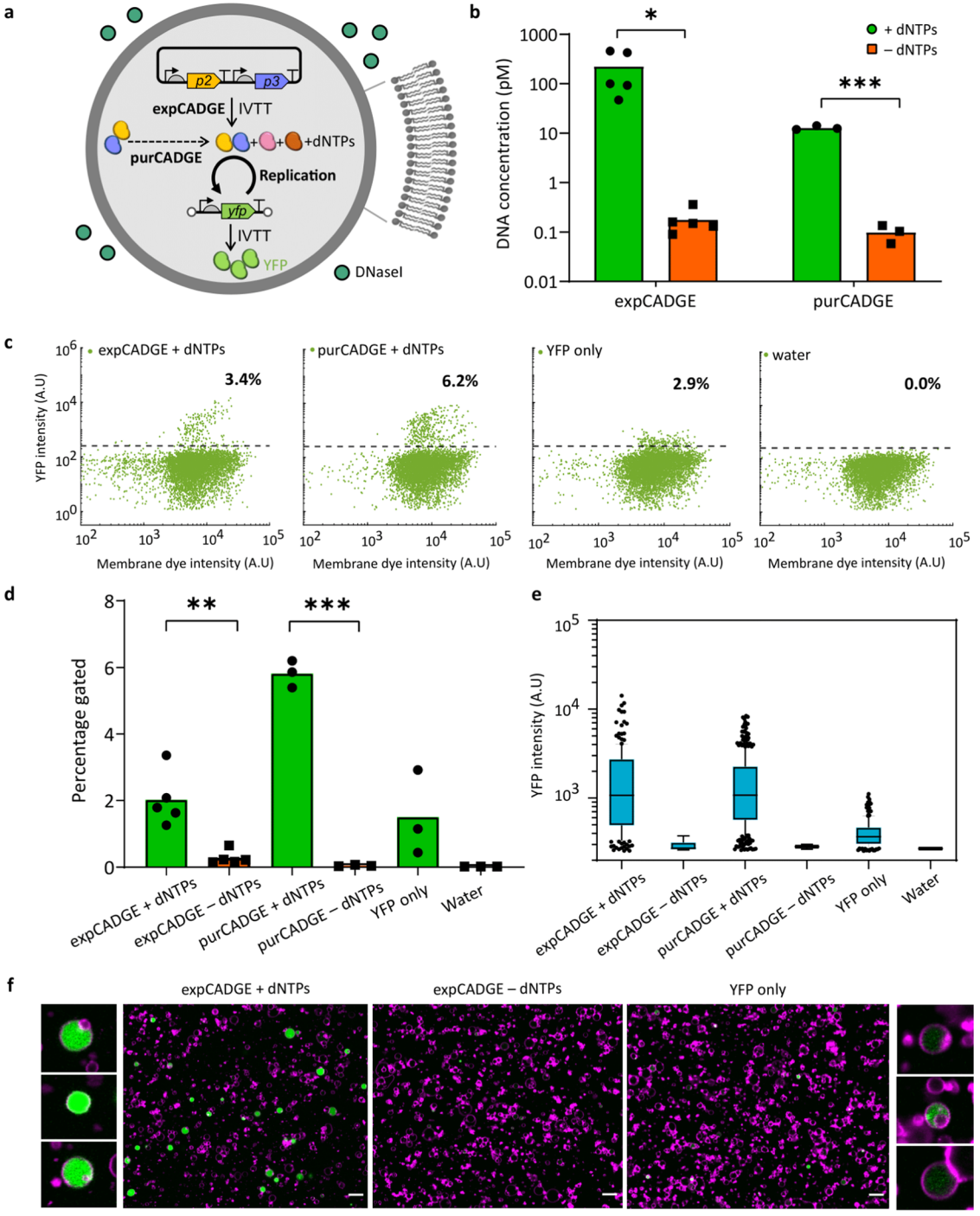


Figure 2. Effect of in-liposome CADGE on the phenotypic output of a reporter gene. (a) Schematic of in-vesiculo reporter gene amplification and expression via CADGE. (b) Absolute quantitation of ori-*yfp* DNA by qPCR in lysed liposomes. Total 10 pM input template DNA concentration was used, which was reduced due to externally supplied DNase I. (c) Populational variation of in-vesiculo YFP fluorescence in CADGE samples measured by flow cytometry. (d) Quantitation of the fraction of liposomes showing above-background YFP fluorescence estimated by the horizontal gate in (c). (e) Box plot representation of flow cytometric data of YFP expression in CADGE liposomes. (f) Confocal microscopy imaging of CADGE and unamplified samples. Magenta: Texas Red (membrane dye). Green: YFP. Scale bar: 5 μm. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$.

slope) comprised between 3.8 and 6.4 nM min^{-1} and a time before saturation of about 300 min (Figure 1f,g, Table S1). These values are consistent with previous data obtained with nanomolar concentrations of DNA template,[17,45] suggesting that CADGE does not significantly delay or slow down protein production. In fact, the apparent YFP production rate increases 25-fold with purCADGE and 40-fold with expCADGE compared to the condition with unamplified ori-*yfp* (Figure 1h).

CADGE Improves Phenotypic Output in Liposomes. We next sought to demonstrate that CADGE is able to increase the number of liposomes with detectable amounts of synthesized protein starting from clonal quantities of DNA molecules (Figure 2a). To this end, the construct ori-*yfp* and the CADGE components were encapsulated in a polydisperse population of liposomes, the bilayer of which is composed of biologically relevant lipids found in the composition of *E. coli*

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

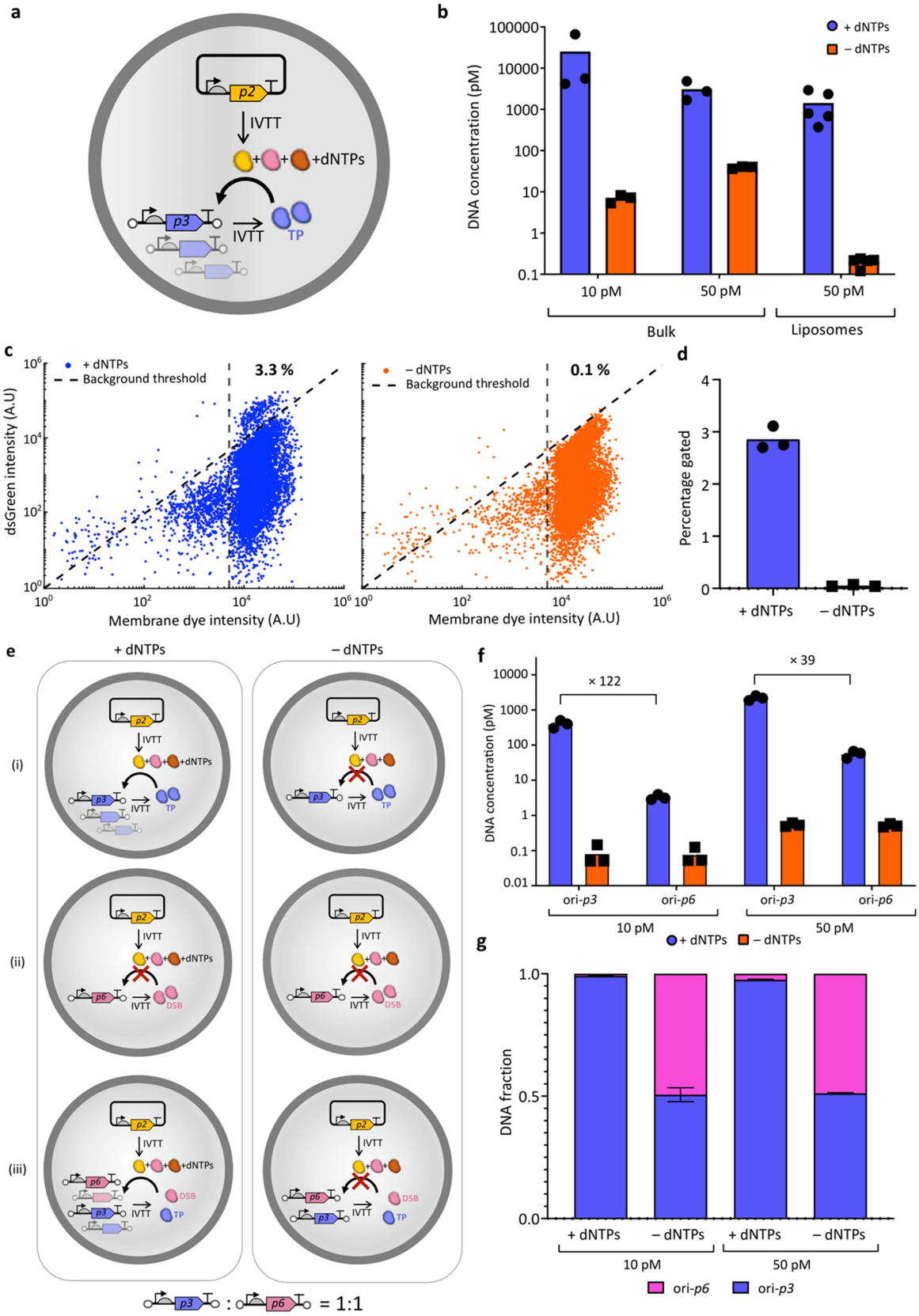


Figure 3. CADGE with a positive feedback loop. (a) Schematic of in-vesiculo ori-*p3* DNA amplification and expression via CADGE. (b) Absolute quantitation of ori-*p3* DNA by qPCR in lysed liposomes. Total 10 pM input template DNA concentration was used, which was reduced due to externally supplied DNase I. (c) Populational variation of dsGreen fluorescence in CADGE liposomes measured by flow cytometry. (d) Quantitation of the fraction of liposomes showing above-background dsGreen fluorescence estimated by the diagonal gate in (c). (e) Schematics of the different gene expression configurations used in (f) and (g). (f) Absolute DNA quantitation of two genes by qPCR of liposome suspensions after a single round

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

Figure 3. continued

of selection. Total 10 or 50 pM input template DNA concentration was used, which was reduced due to externally supplied DNase I. (g) Calculated fractions of the two genes in the mixture after a single round of selection. Error bars indicate standard deviation from three biological replicates.

plasma membrane.[18] Ori-*yfp* was introduced at 10 pM bulk concentration, corresponding to an expected average number of DNA molecules per liposome $\lambda = 0.2$ (Methods section) if one assumes an average liposome radius of 2 μm.[19] To confine the IVTT and replication reactions to the interior of the liposomes, we introduced DNase I to the outer phase of the liposome population, which yielded a concentration of left-over DNA inside vesicles of around 100 fM (Figure 2b). The extent of clonal amplification was assessed by comparing end-point data (typically overnight incubation at 30 °C) with (+) and without (−) dNTPs. To recover the DNA for analysis, we heat-inactivated DNase I and released the DNA from the liposomes by dilution in water. Quantification by qPCR revealed that in both purCADGE and expCADGE, over 100 times more DNA was obtained in the presence of dNTPs than in the absence thereof (Figure 2b), suggesting that DNA was amplified inside the liposomes.

To assess the effect of DNA amplification on gene expression, we analyzed individual liposomes for YFP signal using flow cytometry (Figure 2c−e). We confirmed that under both CADGE configurations, and in the presence of dNTPs, a higher percentage of liposomes exhibited a YFP fluorescence above the background level (Figure 2c,d). This was expected from the strongly reduced protein expression level at low DNA concentrations (Figure S3). Interestingly, the mean intensity of YFP-expressing liposomes was about 5-fold higher in the presence of dNTPs compared to those in the absence but also compared to samples that contained none of the components of the minimal replication machinery, and the range of intensity values expanded across almost two orders of magnitude (Figure 2c,e). Similar observation was made from fluorescence imaging of individual liposomes (Figure 2f). These results suggest that clonal amplification not only boosts gene expression to overpass the threshold for measurable activity but also increases the dynamic range of the phenotypic output. Although the percentage of YFP-expressing liposomes was slightly higher with purCADGE compared to expCADGE (Figure 2d), the intensity profiles were similar (Figure 2e), suggesting that co-expression of *p2* and *p3* genes does not significantly affect the production of protein of interest (POI) in liposomes. Similar conclusion could be reached from bulk reactions (Figure 1f,g).

We noticed that the percentage of YFP-expressing liposomes was lower in the −dNTPs sample compared to the condition where replication reagents were omitted (YFP only, Figure 2d). This suggests that some replication components may inhibit transcription-translation. We tested this hypothesis by varying DSB and SSB concentrations in ori-*yfp* bulk reactions and found that reduced amounts of DSB led to higher expression of ori-*yfp*, while changing SSB concentrations had little effect (Figure S4). Considering that DSB is a Φ29 transcription regulator of early and late genes,[73] it is possible that binding of DSB to the DNA template inhibits gene expression in vitro. Therefore, we decided to lower DSB concentration down to either 52.5 or 105 μg/mL, in order to mitigate inhibitory effects without compromising DNA replication.

CADGE with a Positive Feedback Loop. We then implemented expCADGE with a positive feedback loop coupling POI synthesis back to DNA replication. The autocatalytic framework of this selection strategy may offer a more effective and efficient alternative to fluorescence-based screening methods. Ori-*p3* template coding for TP was introduced at 10 pM concentration ($\lambda = 0.2$), supplemented with an excess amount of plasmid encoding solely the DNAP (Figure 3a), and encapsulated in liposomes. We hypothesized that an initial seed expression of TP could kick off the replication of ori-*p3* with the expressed DNAP and yield increasing amounts of ori-*p3*. Quantitative PCR showed that the *p3* gene was amplified inside liposomes by three orders of magnitude in the presence of dNTPs (Figure 3b) compared to the −dNTPs control. The DNA intercalating dye dsGreen was used as a fluorescent marker to assess DNA amplification in single vesicles by flow cytometry. A fraction of liposomes with increased dsGreen fluorescence compared to the background was detected in the presence of dNTPs, which corroborates that self-amplification of DNA took place (Figure 3c,d).

The high amplification of ori-*p3* prompted us to experimentally determine the bulk concentration of DNA template below which the amplification is "clonal." Experimental validation of the $\lambda = 1$ regime is important considering the polydispersity of the liposome population, which differs from our assumption of constant volume (Methods section). Therefore, we performed a mock enrichment experiment by co-encapsulating ori-*p3* and an equimolar amount of unrelated DNA, also flanked with replication origins (here ori-*p6*) (Figure 3e). In this scenario, DNA replication is conditional to the presence of both DNAP and TP. Therefore, ori-*p6* can only be replicated when co-encapsulated with ori-*p3*, i.e., under nonclonal conditions where more than a single molecule of ori-*GOI* is present in a liposome. Conversely, an enrichment of ori-*p3* over ori-*p6* would indicate that amplification is mostly clonal. After a single round of selection, ori-*p3* was enriched 114-fold and 37-fold over ori-*p6* when starting with 10 pM and 50 pM DNA concentrations, respectively (Figure 3f,g). This result confirms that in-liposome amplification of ori-*GOI* is mostly clonal and that CADGE is suitable for in vitro directed evolution purposes.

Amplification of ori-*p6* was however not totally prohibited, even at 10 pM input mixture concentration (Figure 3f). The latter is not unexpected considering that the estimated probability of co-occupancy of the ori-*p3* and ori-*p6* templates is not zero but is $(1 - e^{-\lambda})^2 = 0.15$ with 50 pM input mixture concentration ($\lambda = 0.5$ for each ori-*GOI*). Together, the significant enrichment of ori-*p3* over ori-*p6* experimentally validates that 10 pM (and to some extent 50 pM) concentration is enough to keep a strong genotype-to-phenotype link in our polydisperse liposome population. This experiment also implies that, as long as DNA replication can be coupled to a POI activity, TP or any other POI can be potentially evolved using this selection scheme.

CADGE Improves the Enrichment Efficiency of a GOI Based on High-Throughput Screening. Next, we asked whether CADGE may be beneficial for in vitro protein evolution via fluorescence-based screening. To this end, we performed a mock enrichment experiment at a clonal expression condition with 10 pM ori-*GOI*, i.e., $\lambda = 0.2$. We aimed to enrich the DNA template ori-*yfp* from a mock library containing an excess of the

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

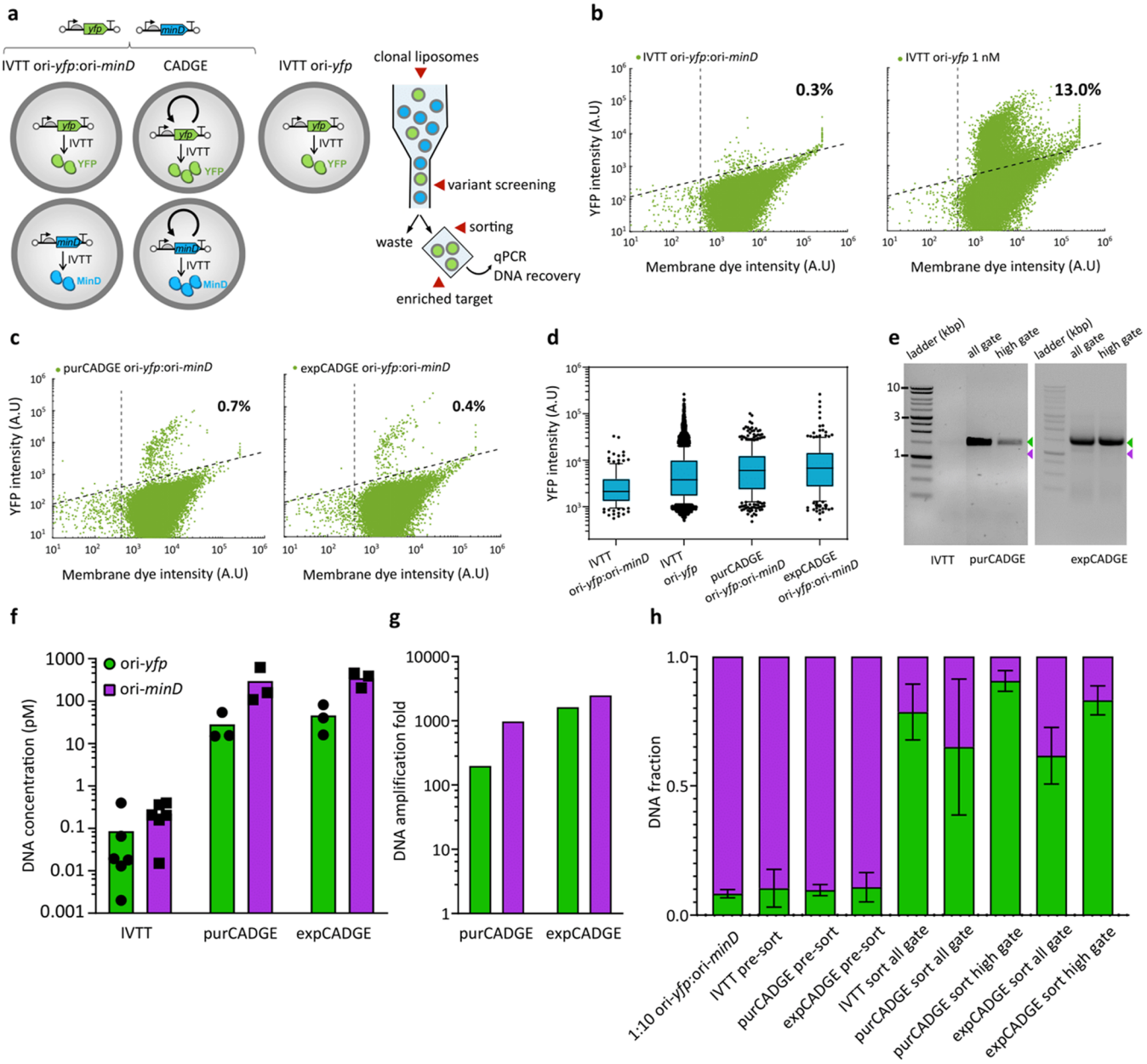


Figure 4. Enrichment of clonally amplified GOI via high-throughput FACS screening. (a) Schematics of clonal gene expression and enrichment experimental design for FACS. (b) Flow cytometry analysis of liposomes prepared from indicated DNA template mixtures in the PURE system: 10 pM total with 1:1 YFP:MinD DNA mixture (left) and 1 nM YFP DNA (right). (c) Flow cytometry analysis of CADGE liposomes prepared from 10 pM total input YFP:MinD DNA template mixture. (d) Box plot analysis of YFP intensity. (e) Agarose gel electrophoresis images of recovered DNA from sorted CADGE liposomes. Green arrowhead indicates the expected DNA size for ori-*yfp* (2 kb), and purple arrowhead indicates the expected size for ori-*minD* (∼1.4 kb). (f) Absolute DNA quantitation by qPCR of pre-sort liposome suspensions. (g) DNA amplification in pre-sort purCADGE and expCADGE liposome samples. Amplification was calculated as DNA concentration of a specific gene in end-point samples with dNTPs compared to end-point samples without dNTPs. Color coding is the same as in (f). (h) Fractions of ori-*minD* and ori-*yfp* DNA mixtures recovered from pre-sort or post-sort liposomes as calculated from absolute DNA quantification by qPCR. Error bars indicate standard deviation from three to seven biological replicates.

unrelated template ori-*minD* based on the fluorescence of expressed YFP by fluorescence-activated cell sorting (FACS). For this, the ori-*yfp* DNA template was mixed with 10-fold excess of ori-*minD* template and the DNA/PURE mixture was encapsulated in liposomes at a total 10 pM DNA concentration (Figure 4a). At such a low template DNA concentration (1 pM ori-*yfp*), expression of YFP is significantly reduced compared to higher DNA concentrations typically used in cell-free reactions, leading to low signal-to-noise ratio (Figure 4b). As expected from previous results, CADGE liposomes exhibited higher dynamic range of YFP fluorescence compared to liposomes that contained the same input DNA mixture concentration but no replication factors (Figure 4b,d). An up to 3-fold increase of the mean intensity of YFP-positive liposomes was measured upon gene amplification (Figure 4d). For sorting, two stringency conditions were tested: the "all-gate", which encompassed the top 1% of all of the liposomes (applied to both nonamplified and CADGE samples), and the "high-gate", which included only the

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

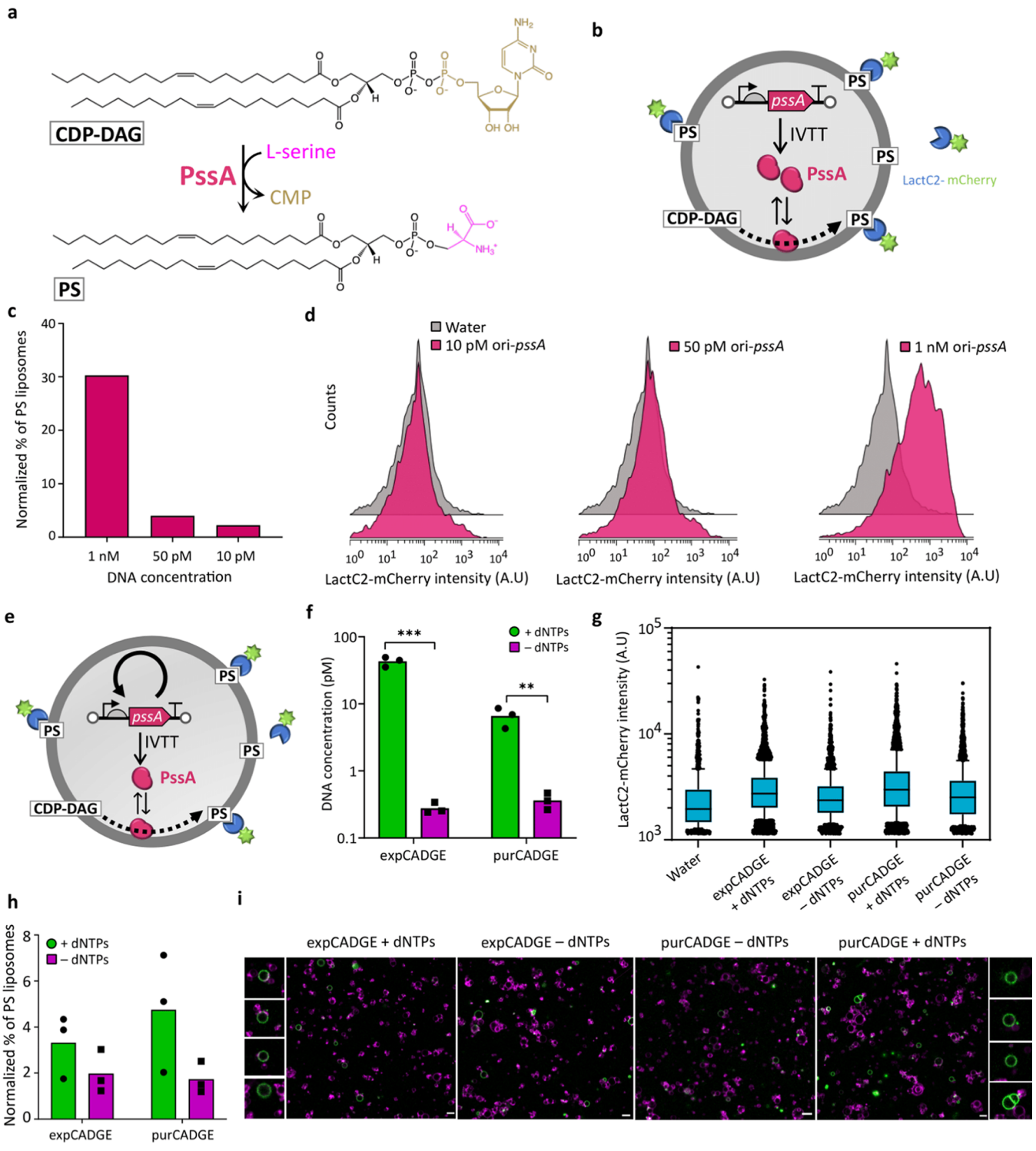


Figure 5. Application of CADGE for improving the enzymatic catalysis of phospholipid headgroup conversion. (a) Schematic of CDP-DAG conversion to PS by PssA. (b) Schematic of in-vesiculo expressed PssA enzyme activity and detection of PS-positive liposomes by LactC2-mCherry binding. Percentage quantitation (c) and histograms (d) of PS-positive liposomes expressing PssA in the PURE system, as assayed by flow cytometry (raw data in Figure S6). (e) Schematic of CADGE liposomes expressing PssA enzyme and detection of PS-positive liposomes by LactC2-mCherry binding. (f) Absolute DNA quantitation by qPCR of lyzed CADGE liposome samples. ** $P \leq 0.01$; *** $P \leq 0.001$. (g) Box plots and (h) quantitation of PS-positive CADGE liposomes expressing PssA, as assayed by flow cytometry (raw data in Figures S8 and S9). (i) Confocal microscopy of CADGE liposomes expressing PssA. Green, LactC2-eGFP; magenta, Texas Red-conjugated lipids. Scale bar: 5 μm.

top (0.2%) of the high-intensity liposomes (applied to CADGE samples only). It was reproducibly difficult to recover the full-length DNA by PCR from the nonamplified liposome samples, while full-length DNA from liposomes with implemented CADGE was easily recovered (Figure 4e). This finding can be explained by higher DNA titers in the sorted liposomes from CADGE samples. Indeed, as assayed by qPCR, ori-*yfp* and ori-*minD* mixtures in liposomes were considerably (both more than

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

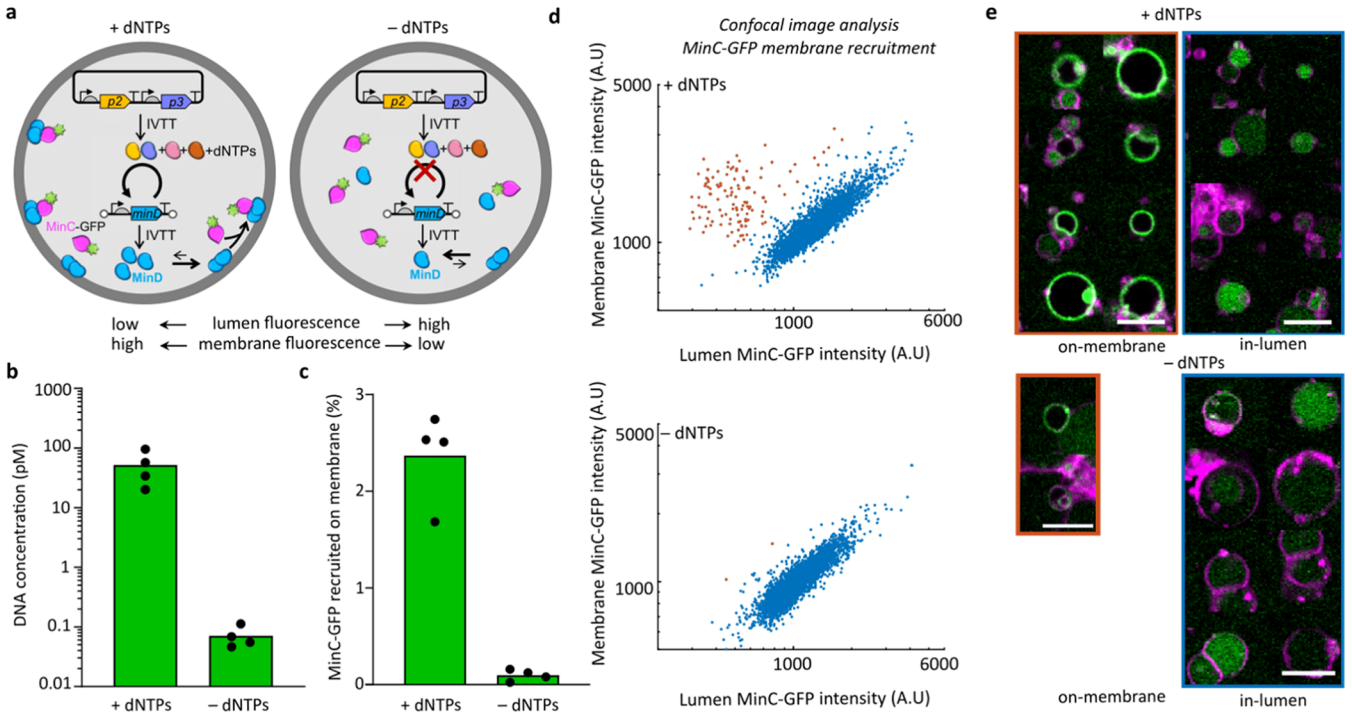


Figure 6. Application of CADGE for facilitating the membrane recruitment of Min proteins. (a) Schematics of MinC−MinD membrane-binding assay in expCADGE liposomes. (b) Absolute quantitation of ori-*minD* DNA by qPCR of lyzed expCADGE liposome samples. Individual symbols are independent biological repeats. (c) Percentage of MinC-GFP recruited to the membrane, as obtained from fluorescence confocal images. (d) Scatter plots of the membrane and lumen fluorescence intensities in individual liposomes. The liposomes marked in orange dots display a marked recruitment of MinC-GFP to the membrane. (e) Montage of randomly selected liposome images taken from the data shown in (d). Scale bars: 5 μm.

a 100-fold) and uniformly (i.e., two genes amplified equally in a single sample) amplified with both purCADGE and expCADGE (Figure 4f−h pre-sorted samples). Furthermore, qPCR analysis of sorted liposome samples gave the enrichment efficiency of ori-*yfp* over ori-*minD* (Table S2). Using the more stringent condition of "high-gate" in CADGE samples results in improved purity of YFP sorting compared to "all-gate" in both CADGE and nonamplified samples (Figure 4h; Figure S5).

Fluorescent proteins expressed from single copies of templates in biomimetic compartments can only be enriched several-fold per sorting round,[36] likely due to low signal-to-noise ratios. Our findings show that CADGE improves enrichment efficiency by enabling selection of liposomes with more stringent fluorescence criteria (enrichment efficiency can reach 89 compared to 31 without amplification, Table S2) and DNA recovery in a single round of mock enrichment experiment, and thus suggest that CADGE may facilitate in vitro protein evolution.

CADGE Improves Phenotypic Output of Synthetic Cell Modules. We previously proposed in vitro evolution as a route to build a synthetic minimal cell.[12] Here, we seek to exploit CADGE for improving the expression of genes that are relevant for the construction of functional cellular modules. One candidate gene is *pssA* from the Kennedy phospholipid biosynthesis pathway.[74,75] The *E. coli pssA* gene encodes an enzyme that conjugates cytidine diphosphate-diacylglycerol (CDP-DAG) with L-serine to produce cytidine monophosphate and phosphatydilserine (PS) (Figure 5a), a precursor of phosphatidylethanolamine. To assay the activity of in-vesiculo synthesized PssA enzyme, we encapsulated PURE and the plasmid encoding the *pssA* gene in phospholipid liposomes containing 5 mol % CDP-DAG and digested the extraliposomal DNA with DNase I (Figure 5b). Since PssA is active as a membrane-associated enzyme, PS would be incorporated into the inner leaflet of the membrane.[76,77] However, as previously suggested,[19] we expected some flipping of phospholipids to the outer membrane such that the enzymatic activity of entrapped PssA could be detected by externally staining the liposomes with a PS-specific probe. To this end, we implemented C2-domain of lactadherin protein (LactC2) fused to a fluorescent protein like mCherry or eGFP (Figure 5b).[19] By flow cytometric analysis of LactC2-mCherry- and Acridine Orange- (membrane marker) stained liposomes, we observed that PS production (and, we assumed, gene expression) reduces considerably at limiting template DNA concentrations (10 and 50 pM DNA) compared to 1 nM (Figure 5c,d; Figure S6). Alternatively, we stained the liposomes with LactC2-eGFP and Texas Red (membrane dye) and imaged them by confocal microscopy (Figure S7). We found that limiting the template DNA concentration visibly reduces LactC2 binding, suggesting that *pssA* gene expression is diminished at low input DNA concentrations.

To test if clonal DNA amplification can improve PssA expression, we co-encapsulated the linear ori-*pssA* DNA fragment (10 or 50 pM) with the required additives for either purCADGE or expCADGE (Figure 5e) and incubated at 30 °C for 4 h. Using qPCR, we confirmed 10- to 100-fold amplification of the *pssA* gene compared to −dNTP controls with input ori-*pssA* concentrations of 10 pM, in both CADGE configurations (Figure 5f). Even though PS synthesis was detectable in −dNTP samples, the number of liposomes exhibiting a PS-positive phenotype and mean intensity of recruited LactC2-eGFP increased with functional CADGE (+dNTPs, Figure 5g,h,i; Figures S8 and S9). Overall, these findings demonstrate that, within a synthetic cell context, clonal amplification of template

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

DNA can improve phospholipid headgroup conversion from in vitro expressed PssA protein.

Besides gene-directed phospholipid production, we decided to explore the benefit of CADGE for implementation of the Min system in clonal liposomes. The Min system is involved in the spatial organization of cytokinesis events in *E. coli*[78] and is therefore a relevant protein system for synthetic cell division. MinD is an ATP-dependent membrane-binding protein that recruits MinC, an FtsZ-polymerization inhibitor.[79] We assembled expCADGE reactions with 10 pM ori-*minD* DNA and purified eGFP-MinC as a reporter of the activity of synthesized MinD (Figure 6a), and encapsulated the mixture in liposomes. Quantitative PCR data showed that ori-*minD* was clonally amplified almost a thousand-fold compared to −dNTPs control samples (Figure 6b). Confocal imaging and analysis of eGFP-MinC fluorescence distribution in the lumen and at the membrane revealed that in the vast majority of the liposomes, the basal amount of expressed MinD is not sufficient to recruit eGFP-MinC to the membrane (−dNTPs, Figure 6c−e). Using expCADGE, a larger fraction of liposomes exhibited an excess fluorescence signal at the membrane (+dNTPs, Figure 6c−e), indicating that clonal amplification led to a relocalization of MinC through improved expression of functional MinD.

DISCUSSION

Herein, we established CADGE, a single-step DNA amplification and in situ transcription-translation strategy that can be used for improving clonal gene expression. Our findings suggest that CADGE can be instrumental in facilitating in vitro evolution of a variety of genes, including those that are important for synthetic cell construction. The general applicability of this strategy is enabled by only a few requirements: the possibility of in vitro expression of the GOI, and in vitro activity of the POI.

Advantages of CADGE for clonal amplification, compared to previous strategies[28,51−56,80] include minimum effort (i.e., a single-step amplification and expression), time (around 2 h to set up the reaction), and instrumentation (although microfluidic devices for microcompartmentalization or screening can be implemented, if necessary). The benefit of using CADGE post-enrichment lies in the simplicity of the protocol since only PCR is required to proceed to another round of encapsulation/enrichment, and in the improved DNA recovery yield.

Despite a number of advantages, CADGE is not without some limitations. Thus, DSB was found to exhibit some inhibitory effect on gene expression, which could be mitigated to some extent by optimized DSB concentrations (Figure S2). Moreover, in our fluorescence measurements of CADGE samples, the measured percentage of liposomes with a fluorescence signal above the activity threshold (3 to 5% with the studied ori-*GOI*) is lower than the predicted value of $(1 - e^{-\lambda}) \times 100 = 18\%$, with $\lambda = 0.2$ at 10 pM DNA, if one assumes that all liposomes with at least one DNA copy would give a signal and that all liposomes have a fixed diameter of 4 μm. More accurate percentage values calculated from histograms of liposome sizes in different samples are reported in Table S3, but the conclusions remain the same. The discrepancy between the measured and predicted fraction of "active" liposomes suggests that (i) some DNA molecules are transcriptionally inactive or depleted into the lipid film,[44] thus lowering the apparent λ, (ii) only liposomes with a particularly high concentration of amplified DNA or reporter protein cross the fluorescence detection threshold in our flow cytometer-based activity assay, or (iii) the encapsulation of input DNA does not follow a Poisson distribution. We suspect that under some conditions (Figures 2d and 5h with ori-*yfp* and ori-*pssA*, respectively) competition for resources during *p2* and *p3* expression may limit the yield of synthesized POI in expCADGE. Therefore, optimization of the concentration of the *p2*−*p3* expression plasmid might be necessary for effective channeling of resources toward expression of POI. This drawback may be alleviated to some extent by using purified DNAP and TP (purCADGE, see Figures 2 and 5). Another limitation is the current lack of commercial availability of some of the required components, such as purified TP, SSB, or DSB. However, TP (together with DNAP) can be expressed from a plasmid in situ (expCADGE), while Φ29 SSB might in principle be replaced with a commercially available alternative (such as *E. coli* SSB), provided that it is shown to be compatible. Overall, we recommend that the optimality of expCADGE vs purCADGE, as well as the optimal DSB and *p2*−*p3* expression plasmid concentrations should be determined on a case-by-case basis for each GOI. We also believe that the yield of POI production per DNA template could be further improved through buffer optimization, in particular the concentrations of magnesium and potassium glutamate,[81] tRNAs, and NTPs.[49]

PURE*frex2.0* in vitro transcription-translation system was used here in its standard composition. However, other promoters than T7, such as the bacteriophage SP6,[19,82] T3[83] or native *E. coli*[84] promoters, could also be used in combination with their cognate RNA polymerase supplied in the reaction mixture. One challenge may reside in the management of collision events between the Φ29 DNA polymerase and the RNA polymerases originating from different organisms.[85,86] Moreover, cell lysates, especially from *E. coli*,[87] could be utilized as a cheaper cell-free expression system, in particular for biomanufacturing purposes.[32] The extract could be modified to avoid the degradation of linear PCR fragments by exonucleases, for instance by supplementing inhibitors of RecBCD (ExoV), such as GamS protein[88] or χ-DNA oligonucleotides[89] or using Δ*RecBCD E. coli* strains.[90] Alternatively, purified TP-bound DNA[18] could be used as a template in cell lysates, assuming that the parental TP hinders exonuclease digestion. Application of CADGE in eukaryotic cell extracts—e.g., from insect cells, wheat germ, rabbit reticulocytes, and human cells—might be useful for the production and engineering of proteins with post-translational modifications, such as glycosylation and phosphorylation. While protein yields may remain low even with clonal gene amplification compared to *E. coli*-based cell-free systems, the increased amount of DNA may be sufficient for the recovery of interesting gene variants. In general, codon usage of GOI may be optimized for the chosen cell-free translation system, which should not influence much the DNA replication efficacy given the high template tolerance of the Φ29 DNAP.

In the shown examples, the genotype-to-phenotype coupling was established using phospholipid vesicles. Liposomes are uniquely suited for cell-free evolution of peripheral and transmembrane proteins,[27,39] and for their tunable membrane permeability, which could be relevant to assay the activity of POI through external addition of substrates or cofactors. The lipid-coated bead approach for liposome production[91] was chosen for its simplicity as it does not require specialized equipment, for the easy storage and distribution across laboratories of pre-assembled lipid films deposited on glass microbeads, and for its high biocompatibility due to the absence of organic solvent. One of the major drawbacks of using liposomes prepared by the swelling method is the heterogeneity of liposome sizes (Figure

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

S10) and encapsulation of DNA, PURE, or CADGE components. Other methods for the preparation of more homogeneous liposomes in size and encapsulation efficiency, such as enhanced continuous droplet interface crossing encapsulation[92] and double-emulsion microfluidics,[93] could in principle be utilized as well.

Other types of microcompartments can also potentially be combined with CADGE: water-in-oil emulsion droplets,[52] microfabricated chambers,[40] and peptide-based compartments.[94] Emulsion droplets are particularly appealing for their high monodispersity and because they have already been integrated in microfluidic-based screening platforms for directed evolution of water-soluble enzymes.[28,95]

Application of gene-expressing liposomes empowered with clonal amplification is also relevant to build a synthetic cell from the ground-up. When applied to essential genes, CADGE-assisted directed evolution might accelerate the optimization of individual cellular modules and their integration to achieve higher-level functions.[12] Considering the excellent processivity of Φ29 DNAP,[60,62] the application of CADGE to long synthetic genomes can reasonably be envisaged. Through the example of TP (Figure 3), we showed the implementation of a positive feedback loop, where the GOI can itself assist in its own amplification, thereby circumventing the need for screening. This reaction scheme may in principle be expanded to self-amplification of polymerases and gene circuits based on DNA polymerization, such as in compartmentalized self-replication[96] and compartmentalized partnered replication.[97] Moreover, self-organization and catalytic activity of the peripheral membrane proteins MinD and PssA were detectable by isothermal DNA amplification from clonal amounts. This strategy might be particularly useful for the in vitro evolution of cellular functions starting from a single copy of ori-*GOI* library variants encapsulated in liposomes. The replicating template may contain single or multiple genes encoding entire pathways and multiprotein complexes. For instance, application of CADGE to phospholipid-synthesizing enzymes of the Kennedy pathway located upstream (PlsB, PlsC, and CdsA) and downstream (Psd) of PssA could aid in optimizing synthetic cell growth through directed evolution.

Finally, we anticipate that performing CADGE under mutagenic conditions could extend its utility for in situ library production. For example, a mutator DNA polymerase[98] or mutagenic factors, such as Mn^{2+} and dNTP analogues, could be employed for genetic diversification directly within liposomes, bypassing the step of external gene library preparation. Such an error-prone CADGE strategy might be particularly interesting for introducing random mutations across long (>10 kbp) DNA templates, for instance, large synthetic genomes for the evolutionary construction of a minimal cell.[12]

METHODS

Buffers and Solutions. All buffers and solutions were made using Milli-Q grade water with 18.2 MΩ resistivity (Millipore, USA). Chemicals were purchased from Sigma-Aldrich unless otherwise indicated.

Construct design. G365 (pUC-ori-YFP) was constructed by subcloning of the YFP gene (amplified by primers 1106 ChD/1107 ChD from plasmid G79) into Φ29 origins-containing vector G96[18] (amplified by primers 1104 ChD/1105 ChD) via the Gibson Assembly method.[99] G368 (pUC-ori-pssA) was cloned by subcloning of the *pssA* gene (amplified by primers 1115 ChD/1116 ChD from plasmid G149) into Φ29 origins-containing vector G96 (amplified by primers 1104 ChD/1105 ChD) via the Gibson Assembly method. Plasmid G338 (pUC-ori-TP) was obtained as a result of subcloning the fragment ori-p2p3, which was PCR-amplified from plasmid G95 (plasmid encoding for ori-*p2*−*p3*)[18] using the primers 961 ChD/962 ChD (with overhangs containing KpnI and HindIII restriction sites) into the KpnI-HindIII-linearized pUC19 vector, during which a spontaneous recombination event flipped out the entire (t7)promoter-p2-(vsv)terminator fragment, only leaving the shorter oriL-(t7)promoter-p3-(t7)terminator-oriR insert. G437 (pUC-ori-minD) was obtained by subcloning the *minD* gene (amplified by primers 91 ChD/397 ChD from plasmid pUC57-MinD)[20] into the Φ29 origins-containing vector G365 (amplified by primers 535 ChD/562 ChD) via the Gibson Assembly method. The cloning of G85 (pUC57-DNAP) was previously reported.[18] All of the plasmids were cloned by heat-shock transformation of *E. coli* Top10 strain, and plasmids were extracted from individual cultures outgrown in LB/ampicillin (50 μg/mL) using the PURE Yield Plasmid Miniprep kit (Promega). Individual clones were screened and confirmed by Sanger sequencing at Macrogen-Europe B.V. Primer sequences and plasmid descriptions can be found in Tables S4 and S5.

Linear DNA templates were prepared by PCR using 5′-phosphorylated primers (491 ChD/492 ChD). Reactions were set up in 100 μL volume, 500 nM each primer, 200 μM dNTPs, 10 pg/μL DNA template, and 2 units of Phusion High-Fidelity DNA Polymerase (NEB) in HF Phusion buffer, and thermal cycling was performed as follows: 98 °C 30 s initial denaturation, and thermal cycling at (98 °C for 5 s, 72 °C for 90 s) × 20, and final extension at 72 °C for 5 min. Extra care was taken to not over-amplify the DNA by too many thermal cycles, as it was found to adversely affect the quality of purified DNA. The amplified PCR fragments were purified using QIAquick PCR purification buffers (Qiagen) and RNeasy MinElute Cleanup columns (Qiagen) using the manufacturer's guidelines for QIAquick PCR purification, except for longer pre-elution column drying step (4 min at 10,000*g* with open columns), and elution with 14 μL ultrapure water (Merck Milli-Q) in the final step. The purified DNA was quantified by Nanodrop 2000c spectrophotometer (Isogen Life Science) and further analyzed for size and purity by gel electrophoresis.

Purification of DNAP, TP, SSB, DSB, LactC2-eGFP, and LactC2-mCherry. Purified Φ29 DNA replication proteins were produced as described in ref. 18. Stock concentrations and storage buffers are: DNAP (320 ng/μL in 50 mM Tris, pH 7.5, 0.5 M NaCl, 1 mM EDTA, 7 mM β-mercaptoethanol (BME), 50% glycerol), TP (400 ng/μL in 25 mM Tris, pH 7.5, 0.5M NaCl, 1 mM EDTA, 7 mM BME, 0.025% Tween 20, 50% glycerol), SSB (10 mg/mL in 50 mM Tris, pH 7.5, 60 mM ammonium sulfate, 1 mM EDTA, 7 mM BME, 50% glycerol), DSB (10 mg/mL in 50 mM Tris, pH 7.5, 0.1 M ammonium sulfate, 1 mM EDTA, 7 mM BME, 50% glycerol). The proteins were aliquoted and stored at −80 °C. The DNAP and TP proteins were diluted before immediate use into PURE*frex2.0* solution I (GeneFrontier). Both LactC2-eGFP and LactC2-mCherry genes were cloned into pET11 vector, under control of the T7-LacO promotor and in frame with an N-terminal His-tag. LactC2-mCherry was expressed in *E. coli* BL21(DE3) (NEB) and LactC2-eGFP protein was expressed in *E. coli* strain ER2566 (NEB). Overnight pre-cultures were prepared in Luria Broth (LB) medium containing 50 μg/mL ampicillin. The overnight cultures were diluted 1:100 in fresh LB medium with 50 μg/mL

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

ampicillin and incubated at 37 °C while shaking, until an OD_{600} of 0.4−0.6 was reached. Protein expression was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside. The cells were incubated at 26 °C for 4 h or overnight at 16 °C while shaking, and harvested at 4000*g* for 15 min. Pellet of 1 L cells was resuspended in 10 mL of lysis buffer (50 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 10% glycerol). The cells were disrupted by sonication on ice, using 7 pulses of 30 s and 1 min intervals, with an amplitude of 40%. The cell suspension was centrifuged for 30 min at 30,000*g* at 4 °C to remove the cell debris. To the cell-free extract, 10 mM imidazole and SetIII protease inhibitor-EDTA-free (1:1000 dilution, Calbiochem) were added. The proteins were purified with HisPure Ni-NTA resin (Thermo Scientific). The Ni-NTA (∼3 mL) was equilibrated with buffer (50 mM HEPES-KOH, 500 mM NaCl, 10% glycerol, 10 mM imidazole, pH 7.5). The cell-free extract was mixed with the equilibrated resin and incubated for 1 to 16 h while tumbling in the cold room. After incubation the resin with bound protein was transferred into a gravity column, the unbound fraction was removed by gravity and subsequently the resin was washed with 20 equiv volume wash buffer (50 mM HEPES-KOH, 500 mM NaCl, 10% glycerol, 40 mM imidazole, pH 7.5). The protein was eluted with 5 mL elution buffer (50 mM HEPES-KOH, 500 mM NaCl, 500 mM imidazole, 10% glycerol, pH 7.5) and fractions of ∼1 mL were collected. The fluorescent fractions were pooled together and buffer-exchanged with storage buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 10% glycerol) using a 10-MWCO Amicon Ultra-15 centrifugal filter unit (Merck). The concentration of the protein was determined with a Bradford assay.

CADGE in Bulk Reactions. Bulk reactions were set up in PURE*frex2.0* (GeneFrontier). A 20 μL reaction consisted of 10 μL solution I, 1 μL solution II, 2 μL solution III, 20 mM ammonium sulfate, 300 μM dNTPs, 375 μg/mL purified Φ29 SSB protein, 105 μg/mL purified Φ29 DSB protein, 0.6 units/μL of Superase·In RNase inhibitor (Ambion), 10 pM target DNA and either plasmid DNA (2 nM plasmid G85 encoding for the *p2* gene in ori-*p3* clonal amplification experiments or 1 nM G95 encoding for *p2* and *p3* genes in ori-*yfp*, ori-*minD*, and ori-*pssA* experiments) or 3 ng/μL each purified Φ29 DNAP and TP. Reactions were incubated in a nuclease-free PCR tube (VWR) in a ThermalCycler (C1000 Touch, Bio-Rad) at a default temperature of 30 °C. Incubation time was indicated when appropriate.

To analyze the reactions by gel electrophoresis, 10 μL reaction was treated with 0.2 mg/mL RNase A (Promega), 0.25 units RNase One (Promega) at 30 °C for 2 h, followed by 1 mg/mL Proteinase K (Thermo Scientific) at 37 °C for 1 h, and column-purified using the QIAquick PCR purification buffers (Qiagen) and RNeasy MinElute Cleanup columns (Qiagen) using the manufacturer's guidelines for QIAquick PCR purification, except for longer pre-elution column drying step (4 min at 10,000*g* with open columns), and elution with 14 μL ultrapure water (Merck Milli-Q) in the final step. A fraction (6 μL) of the eluate was mixed with an equal volume of 6× purple gel loading dye (NEB) and loaded in 1% agarose gel with ethidium bromide, following which DNA was separated using an electrophoresis system (Bio-Rad). The BenchTop 1-kb DNA Ladder (Promega) was used to estimate the size of DNA.

Mass Spectrometry. LC-MS/MS analysis with QconCATs was employed for the absolute quantification of de novo synthesized proteins in bulk PURE reactions. Pre-ran PURE reaction solutions were mixed with one-third volume of heavily labeled QconCAT(^{15}N)[72] in a 50 mM Tris (pH 8.0) buffer containing 1 mM $CaCl_2$. The samples were then incubated at 90 °C for 10 min and cooled down to 4 °C. Trypsin was then added at a 250 μg/mL final concentration and digestion incubation was carried out overnight at 37 °C. The trypsin-digested samples were treated with trifluoroacetic acid (TFA) 10% and centrifuged for 10 min. The supernatant was then transferred to a glass vial with a small insert for LC-MS/MS analysis. Measurements were performed on a 6460 Triple Quad LC-MS system (Agilent Technologies, USA) using Skyline software.[100] Samples of 5.5 μL were injected per run into an ACQUITY UPLC Peptide CSH C18 Column (Waters Corporation, USA). The peptides were separated in a gradient of buffer A (25 mM formic acid in Milli-Q water) and buffer B (50 mM formic acid in acetonitrile) at a flow rate of 500 μL per minute and at a column temperature of 40 °C. The column was initially equilibrated with 98% buffer A. After sample injection, buffer A gradient was changed to 70% (over the first 20 min), 60% (over the next 4 min), and 20% (over the next 30 s). This final ratio was conserved for another 30 s and the column was finally flushed with 98% buffer A to equilibrate it for the next run. The selected peptides and their transitions for both synthesized proteins and heavily labeled QconCATs were measured by multiple reaction monitoring (MRM). The recorded LC-MS/MS data were analyzed with Skyline for fraction calculation between unlabeled and labeled peptides (^{14}N/^{15}N ratio) on both cell-free core/produced proteins and the initially added QconCATs. With these fraction values, and considering the regular concentration of core ribosomal peptides within PURE system (2 μM), we could estimate the concentration of the cell-free expressed proteins using the following equation

$$\frac{\left(\frac{^{14}N}{^{15}N}\right)_{PP} \times 4}{\left(\frac{^{15}N}{^{15}N}\right)_{CPP1} + \left(\frac{^{15}N}{^{15}N}\right)_{CPP2}} \quad (1)$$

where PP refers to the detected peptide of produced protein, CPP1 refers to the detected peptide 1 of the core ribosomal protein (GVVVAIDK), and CPP2 refers to the detected peptide 2 of the core ribosomal protein (VVGQLGQVLGPR). MS/MS measurement details for each of the analyzed proteins can be found in Table S6.

In-Vesiculo Protein Expression. The procedure was adapted from ref 18 with minor modifications. To prepare lipid-coated beads as precursors of liposomes with the standard lipid composition, in a 5 mL round-bottom glass flask, a primary lipid mixture was prepared consisting of DOPC (50.8 mol %), DOPE (35.6 mol %), DOPG (11.5 mol %), and 18:1 cardiolipin (2.1 mol %). The resulting mixture was additionally spiked with DSPE-PEG(2000)-biotin (1 mass %) and DHPE-Texas Red (0.5 mass %) for a total mass of 2 mg. Finally, the lipid mixture was complemented with 25.4 μmol of rhamnose (Sigma-Aldrich) dissolved in methanol. To prepare liposomes containing CDP-DAG, the primary lipid mixture composition was modified as follows: DOPC (47.5 mol %), DOPE (34.2 mol %), DOPG (11.4 mol %), 18:1 cardiolipin (1.9 mol %), and CDP-DAG (5 mol %), with additional DSPE-PEG(2000)-biotin (1 mass %) and, if indicated, DHPE-TexasRed (0.5 mass %) for a total mass of 2 mg. All lipids were purchased at Avanti Polar Lipids and dissolved in chloroform, except the DHPE-Texas Red (Invitrogen). Finally, 600 mg of 212−300 μm glass beads (Sigma-Aldrich) were added to the lipid/rhamnose

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

solution, and the organic solvent was removed by rotary evaporation at 200 mbar for 2 h at room temperature (rotary evaporator, Heidolph), followed by overnight desiccation. The dried lipid-coated beads were stored under argon at −20 °C until use. A 20 μL PURE*frex2.0* reaction solution was assembled from 10 μL of buffer solution, 1 μL of enzyme solution, 2 μL of ribosome solution, indicated amount of input DNA template in RNase-free Milli-Q water. To the well-mixed reaction, 10 mg lipid-coated beads, additionally pre-dessicated for at least 30 min before use, were added. The 1.5 mL Eppendorf tube containing the bead-PURE*frex2.0* mixture was gently rotated on an automatic tube rotator (VWR) at 4 °C along its axis for 30 min for uniform liposome swelling. The mixtures were then subjected to four freeze/thaw cycles (5 s in liquid nitrogen followed by 10 min on ice). From this step onward, the liposome suspension was handled gently and only with cut pipette tips to prevent liposome breakage. Finally, 10 μL of the supernatant liposome suspension (the beads sediment to the bottom of the tube) was transferred to a PCR tube, where it was mixed with 0.5 units of DNase I (NEB). The reactions were incubated at 30 °C in a thermocycler for the indicated time periods.

In-Vesiculo Clonal Amplification and Expression of Genes. The liposome suspensions were assembled as above, except that the necessary CADGE components were pre-mixed with the PURE*frex2.0* solution prior to the addition of the lipid-coated beads and swelling. The following compounds were supplemented (all final concentrations): 20 mM ammonium sulfate, 0.75 U/μL SUPERase (Ambion), 10−50 pM template DNA (as indicated), 375 μg/μL purified SSB, 21−52.5 μg/μL of purified DSB, 1−3 ng/μL each of purified Φ29 DNAP and TP proteins, 250 pM *p2*−*p3* encoding G95 plasmid, and 300 μM PCR Nucleotide mix (Promega). The liposome suspensions were incubated at 30 °C in a thermocycler for the indicated time periods.

Quantitative PCR. Upon completion of the bulk or in-liposome CADGE reactions at 30 °C, 2 μL samples were harvested, heated at 75 °C in the thermocycler for 15 min to inactivate the DNase I, and diluted 100-fold in Milli-Q water prior to addition to the qPCR mixtures. Ten microliter reactions consisted of PowerUP SYBR Green Master Mix (Applied Biosystems), 400 nM each primer (1121 ChD/1122 ChD for *yfp*, 980 ChD/981 ChD for *p3*, 1125 ChD/1126 ChD for *pssA*, 1208 ChD/1209 ChD for *minD*), and 1 μL of diluted sample. The thermal cycling and data collection were performed on Quantstudio 5 Real-Time PCR instrument (Thermo Fisher), using the thermal cycling protocol 2 min at 50 °C, 5 min at 94 °C, (15 s at 94 °C, 15 s at 56 °C, 30 s at 68 °C) × 45, 5 min at 68 °C, followed by melting curve from 65 to 95 °C. The concentration of nucleic acids was calibrated using 10-fold serial dilutions of corresponding standard DNA templates ranging from 1 fM to 1 nM. Data were analyzed using the Quantstudio Design and Analysis software v1.4.3 Software (Thermo Fisher).

Flow Cytometry. The liposome suspension (1−3 μL) was diluted in 300 μL (final volume) of PB buffer consisting of 20 mM HEPES-KOH, pH 7.6, 180 mM potassium glutamate, and 14 mM magnesium acetate. To remove any remaining beads or large debris, the diluted liposome suspension was gently filtered through the 35 μm nylon mesh of the cell-strainer cap from the 5 mL round-bottom polystyrene test tubes (Falcon). When indicated, DsGreen (Lumiprobe) dye was added at a 1:100,000 stock concentration to stain dsDNA, or Acridine Orange (6 μM) and LactC2-mCherry protein (300 nM) were added to stain the liposome membrane and phosphatydilserine, respectively. The mixture was incubated for 1 h at room temperature to equilibrate binding. Liposomes were screened with the FACSCelesta flow cytometer (BD Biosciences) using the 488 nm laser and 530/30 filter for detection of DsGreen, GFP, YFP, or Acridine Orange, and the 561-nm laser and 610/20 filter for detection of PE-Texas Red or mCherry. The following acquisition parameters were used: photon multiplier tube voltages set at 375 V for forward scatter, 260 V for side scatter (SSC), DsGreen detection at 500 V, PE-Texas Red detection at 300−370 V, YFP detection at 550 V, GFP detection at 700 V, mCherry detection at 370 V, Acridine Orange detection at 400 V, threshold for SSC at 200 V, sample flow 1 (∼1000 events/s), injection volume 50−200 μL, recording of 10−100,000 total events. Data were analyzed using Cytobank (https://community.cytobank.org/). Raw data were pre-processed as described in Figure S11 to filter out possible aggregates and debris.

Confocal Microscopy and Image Analysis. A custom-made glass imaging chamber was functionalized with BSA-biotin:BSA and Neutravidin as previously described.[17] The liposome suspension (3−7 μL) was supplemented with PB buffer to a maximum volume of 7 μL and transferred into the functionalized chamber. The LactC2-GFP probe was used at a final concentration of ∼260 nM. After 30 to 60 min incubation at room temperature to let the liposomes sediment, the sample was imaged with a Nikon A1R Laser scanning confocal microscope equipped with a ×100 objective and operated via the NIS Elements software (Nikon). The laser lines 488 nm (for MinC-eGFP), 514 nm (for YFP), and 561 nm (for DHPE-Texas Red and LactC2-mCherry) were used in combination with appropriate emission filters. The position of the focal plane was manually adjusted to image as many liposomes as possible across their equatorial plane. Image analysis was performed using ImageJ (https://imagej.nih.gov/ij/) and an in-house developed code, called SMELDit, which enables the identification of individual liposomes, as well as the quantification of fluorescence signals at the membrane and in the lumen.

Mock Enrichment of *p3* Gene. The linear DNA constructs ori-*p3* and ori-*p6* were mixed at a 1:1 molar ratio for a total DNA concentration of either 10 or 50 pM in PURE*frex2.0* solutions containing 20 mM ammonium sulfate, 300 μM dNTPs, 375 μg/mL purified SSB, 105 μg/mL purified DSB, and 0.6 units/μL of Superase·In RNase inhibitor. The reactions were also supplemented with 2 nM plasmid DNA encoding for Φ29 DNAP (G85 plasmid). The well-mixed solution was encapsulated in liposomes as described above. Then, 5 μL of bead-free liposome suspension was transferred to a PCR tube, where it was mixed with 0.25 units of DNase 1 (Thermo Scientific), and incubated at 30 °C for 16 h. Upon completion, 2 μL samples were harvested from both + and − dNTPs reactions for quantitative PCR as described above. The enrichment efficiency of ori-*p3* over ori-*p6* was calculated as

$$\frac{\text{fraction } p3\ (+\text{dNTPs})/\text{fraction } p6\ (+\text{dNTPs})}{\text{fraction } p3\ (-\text{dNTPs})/\text{fraction } p6\ (-\text{dNTPs})} \quad (2)$$

Mock Enrichment of *yfp* Gene. The linear DNA constructs ori-*yfp* and ori-*minD* were mixed at a 1:10 molar ratio (1 pM ori-*yfp* and 9 pM ori-*minD* final concentrations) in either gene expression solution (PURE*frex2.0*: 50% v/v solution I, 5% v/v solution II, and 10% v/v solution III supplemented with 0.6 units/μL of Superase·In RNase inhibitor) or gene

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

expression-coupled replication solution (PURE*frex2.0* with an addition of 20 mM ammonium sulfate, 300 μM dNTPs, 375 μg/mL purified SSB, 52.5 μg/mL purified DSB, 3 ng/μL purified Φ29 DNAP, 3 ng/μL purified TP, and 0.6 units/μL of Superase·In RNase inhibitor). The well-mixed solution was encapsulated in liposomes as described above. Then, 10 μL of bead-free liposome suspension was transferred to a PCR tube, where it was mixed with 0.5 units of Proteinase K (Thermo Scientific), and incubated at 30 °C for 16 h. The liposome suspension (3 μL) was mixed with 497 μL of PB buffer and filtered through the 35 μm nylon mesh of the cell-strainer cap from the 5 mL round-bottom polystyrene test tubes (Falcon).

Fluorescence-activated cell sorting was conducted on FACSMelody (BD Biosciences). Lasers PE-CF594(YG) and FITC-BB515, 100 μm nozzle, 23.14 PSI pressure, and 34.2 kHz drop frequency were used. Photon multiplier tube voltages applied were 320 V for forward scatter, 455 for side scatter, 337 V for Texas Red, and 673 V for GFP, and a threshold of 359 V at the side scatter was applied. Liposomes with 1% highest YFP signal were sorted out from liposomes prepared in gene expression solution ("all-gate"), and the same gate was applied to the liposomes prepared in gene expression-coupled replication solution or an adjusted gate including only 0.2% highest YFP signal ("high-gate"). Around 50,000 (low-gate) or 10,000 (high-gate) liposomes were sorted into a 1.5 mL Eppendorf tube. Liposomes from the "all-gate" were further concentrated by centrifugation at 12,000*g* for 3 min, and removing three fourth of the supernatant volume. The proteinase K was heat-inactivated at 95 °C for 5 min.

The DNA contained in sorted liposomes was used as a template for PCR amplification using phosphorylated primers (ChD 491/ChD 492). Reactions were set up in 100 μL volume, 300 nM each primer, 400 μM dNTPs, 10 μL sorted, heat-inactivated liposome solution, and 2 units of KOD Xtreme Hotstart DNA polymerase in Xtreme buffer, and thermal cycling was performed as follows: 2 min at 94 °C for polymerase activation, and thermal cycling at (98 °C for 10 s, 65 °C for 20 s, 68 °C for 1.5 min) × 30. The amplified PCR fragments were purified using QIAquick PCR purification buffers (Qiagen) and RNeasy MinElute Cleanup columns (Qiagen) using the manufacturer's guidelines for QIAquick PCR purification, except for longer pre-elution column drying step (4 min at 10,000*g* with open columns), and elution with 14 μL ultrapure water (Merck Milli-Q) in the final step. The purified DNA was quantified by the Nanodrop 2000c spectrophotometer (Isogen Life Science).

The enrichment efficiency of ori-*yfp* over ori-*minD* was calculated as

$$\frac{\text{fraction } yfp \text{ (post sort)}/\text{fraction } minD \text{ (post sort)}}{\text{fraction } yfp \text{ (pre sort)}/\text{fraction } minD \text{ (pre sort)}} \quad (3)$$

Statistical Analysis of DNA Occupancy. The probability that a liposome contains k molecules of DNA ($k = 0, 1, 2, 3, \cdots$) according to a Poisson distribution is

$$P(k) = \frac{\lambda^k}{k!}e^{-\lambda} \quad (4)$$

where λ is the expected average number of input DNA molecules per liposome. It can be calculated as a function of the diameter d of the liposomes and the bulk concentration C of input DNA templates, as

$$\lambda = \frac{\pi N_A C d^3}{6} \quad (5)$$

where N_A is the Avogadro constant. A CADGE reaction may occur in a liposome if one or more copies of linear DNA template are encapsulated, whose corresponding probability is given by

$$P(k \geq 1) = 1 - e^{-\lambda} \quad (6)$$

With expCADGE, the concentration of *p2*−*p3*-plasmid largely exceeds that of ori-*GOI* such that only the concentration of ori-*GOI* template limits the percentage of liposomes exhibiting CADGE: $P(k_{p2-p3\text{-plasmid}} \geq 1) \times P(k_{ori\text{-}GOI} \geq 1) \approx P(k_{ori\text{-}GOI} \geq 1)$.

Statistics. Box and whiskers plots in Figures 2, 4, and 5 have the following characteristics: middle line is the median, the whiskers of the plot are drawn from the 10th percentile up to the 90th percentile, any data point outside the whiskers is drawn as an individual point.

■ ASSOCIATED CONTENT

Data Availability Statement

Flow cytometry data were analyzed using Cytobank (https://community.cytobank.org/). MATLAB scripts and a user manual for SMELDit are made available upon request.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.2c00668.

Additional data for LC-MS, FACS, qPCR, and confocal image analysis, as well as a list of primer sequences and plasmids used in this study (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Christophe Danelon − *Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, 2629HZ Delft, The Netherlands; Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRAE, INSA, 31077 Toulouse, France;* orcid.org/0000-0002-0961-6640; Email: c.j.a.danelon@tudelft.nl, danelon@insa-toulouse.fr

Authors

Zhanar Abil − *Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, 2629HZ Delft, The Netherlands*

Ana Maria Restrepo Sierra − *Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, 2629HZ Delft, The Netherlands*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.2c00668

Author Contributions

[§]Z.A. and A.M.R.S. contributed equally. Z.A. and C.D. conceptualized the research and acquired funding. Z.A. and A.M.R.S. performed the experiments. Z.A., A.M.R.S., and C.D. designed the experiments, analyzed the data, and wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): Z.A. and C.D. have filed patent 22315279.4 titled "Micro-compartmentalized ultra-high throughput screening from single copy gene libraries".

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

■ ACKNOWLEDGMENTS

The authors are grateful to Margarita Salas and Mario Mencía Caballero (Universidad Autónoma de Madrid), and Alicia del Prado and Miguel de Vega (Centro de Biología Molecular Severo Ochoa, Madrid) for kindly providing them with the purified DNAP, TP, SSB, and DSB proteins. They thank Yannick Rondelez and Thibault Di Meo (ESPCI, Paris) for fruitful discussions. They also thank Ilja Westerlaken for the purification of LactC2-mCherry and LactC2-eGFP proteins, and Elisa Godino for helping with the preparation and imaging of MinD-containing liposome samples. The authors are grateful to Duco Blanken, Flora Yang, and Anne Doerr for their assistance with the mass spectrometry experimental setup. Finally, they are thankful to Tess Bevers for assistance with cloning of the ori-*p3* plasmid and Mats van Tongeren for developing SMELDit. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 707404, and from the Netherlands Organization for Scientific Research (NWO/OCW) via the "BaSyC—Building a Synthetic Cell" Gravitation grant (024.003.019).

■ REFERENCES

(1) Romero, P. A.; Arnold, F. H. Exploring Protein Fitness Landscapes by Directed Evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866−876.
(2) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev. Genet.* **2015**, *16*, 379−394.
(3) Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121*, 12384−12444.
(4) Bornscheuer, U. T.; Hauer, B.; Jaeger, K. E.; Schwaneberg, U. Directed Evolution Empowered Redesign of Natural Proteins for the Sustainable Production of Chemicals and Pharmaceuticals. *Angew. Chem., Int. Ed.* **2019**, *58*, 36−40.
(5) Turner, N. J. Directed Evolution Drives the next Generation of Biocatalysts. *Nat. Chem. Biol.* **2009**, *5*, 567−573.
(6) Engqvist, M. K. M.; Rabe, K. S. Applications of Protein Engineering and Directed Evolution in Plant Research. *Plant Physiol.* **2019**, *179*, 907−917.
(7) Hida, K.; Hanes, J.; Ostermeier, M. Directed Evolution for Drug and Nucleic Acid Delivery. *Adv. Drug Delivery Rev.* **2007**, *59*, 1562−1578.
(8) Xiong, A. S.; Peng, R. H.; Zhuang, J.; Davies, J.; Zhang, J.; Yao, Q. H. Advances in Directed Molecular Evolution of Reporter Genes. *Crit. Rev. Biotechnol.* **2012**, *32*, 133−142.
(9) Johannes, T. W.; Zhao, H. Directed Evolution of Enzymes and Biosynthetic Pathways. *Curr. Opin. Microbiol.* **2006**, *9*, 261−267.
(10) Chatterjee, R.; Yuan, L. Directed Evolution of Metabolic Pathways. *Trends Biotechnol.* **2006**, *24*, 28−38.
(11) Zhou, S.; Wu, Y.; Xie, Z. X.; Jia, B.; Yuan, Y. J. Directed Genome Evolution Driven by Structural Rearrangement Techniques. *Chem. Soc. Rev.* **2021**, *50*, 12788−12807.
(12) Abil, Z.; Danelon, C. Roadmap to Building a Cell: An Evolutionary Approach. *Front. Bioeng. Biotechnol.* **2020**, *8*, 927.
(13) Yu, W.; Sato, K.; Wakabayashi, M.; Nakaishi, T.; Ko-Mitamura, E. P.; Shima, Y.; Itaru Urabe, T.; Yomo, E. Synthesis of Functional Protein in Liposome. *J. Biosci. Bioeng.* **2001**, *92*, 590−593.
(14) Nomura, S. I. M.; Tsumoto, K.; Hamada, T.; Akiyoshi, K.; Nakatani, Y.; Yoshikawa, K. Gene Expression within Cell-Sized Lipid Vesicles. *ChemBioChem* **2003**, *4*, 1172−1175.
(15) Noireaux, V.; Libchaber, A. A Vesicle Bioreactor as a Step toward an Artificial Cell Assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17669−17674.
(16) Fujii, S.; Matsuura, T.; Sunami, T.; Nishikawa, T.; Kazuta, Y.; Yomo, T. Liposome Display for in Vitro Selection and Evolution of Membrane Proteins. *Nat. Protoc.* **2014**, *9*, 1578−1591.
(17) Blanken, D.; Van Nies, P.; Danelon, C. Quantitative Imaging of Gene-Expressing Liposomes Reveals Rare Favorable Phenotypes. *Phys. Biol.* **2019**, *16*, No. 045002.
(18) Van Nies, P.; Westerlaken, I.; Blanken, D.; Salas, M.; Mencía, M.; Danelon, C. Self-Replication of DNA by Its Encoded Proteins in Liposome-Based Synthetic Cells. *Nat. Commun.* **2018**, *9*, No. 1583.
(19) Blanken, D.; Foschepoth, D.; Serrão, A. C.; Danelon, C. Genetically Controlled Membrane Synthesis in Liposomes. *Nat. Commun.* **2020**, *11*, No. 4317.
(20) Godino, E.; López, J. N.; Foschepoth, D.; Cleij, C.; Doerr, A.; Castellà, C. F.; Danelon, C. De Novo Synthesized Min Proteins Drive Oscillatory Liposome Deformation and Regulate FtsA-FtsZ Cytoskeletal Patterns. *Nat. Commun.* **2019**, *10*, No. 4969.
(21) 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*, No. 539.
(22) Kattan, J.; Doerr, A.; Dogterom, M.; Danelon, C. Shaping Liposomes by Cell-Free Expressed Bacterial Microtubules. *ACS Synth. Biol.* **2021**, *10*, 2447−2455.
(23) Berhanu, S.; Ueda, T.; Kuruma, Y. Artificial Photosynthetic Cell Producing Energy for Protein Synthesis. *Nat. Commun.* **2019**, *10*, No. 1325.
(24) Pourmir, A.; Johannes, T. W. Directed Evolution: Selection of the Host Organism. *Comput. Struct. Biotechnol. J.* **2012**, *2*, No. e201209012.
(25) Shimizu, Y.; Kuruma, Y.; Ying, B. W.; Umekage, S.; Ueda, T. Cell-Free Translation Systems for Protein Engineering. *FEBS J.* **2006**, *273*, 4133−4140.
(26) Perez, J. G.; Stark, J. C.; Jewett, M. C. Cell-Free Synthetic Biology: Engineering Beyond the Cell. *Cold Spring Harbor Perspect. Biol.* **2016**, *8*, No. a023853.
(27) Fujii, S.; Matsuura, T.; Sunami, T.; Kazuta, Y.; Yomo, T. In Vitro Evolution of α-Hemolysin Using a Liposome Display. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 16796−16801.
(28) Holstein, J. M.; Gylstorff, C.; Hollfelder, F. Cell-Free Directed Evolution of a Protease in Microdroplets at Ultrahigh Throughput. *ACS Synth. Biol.* **2021**, *10*, 252−257.
(29) Newton, M. S.; Cabezas-Perusse, Y.; Tong, C. L.; Seelig, B. In Vitro Selection of Peptides and Proteins - Advantages of MRNA Display. *ACS Synth. Biol.* **2020**, *9*, 181−190.
(30) Watts, R. E.; Forster, A. C. Update on Pure Translation Display with Unnatural Amino Acid Incorporation. In *Ribosome Display and Related Technologies*, Methods in Molecular Biology; Springer: New York, 2012; Vol. *805*, pp 349−365.
(31) Uyeda, A.; Watanabe, T.; Kato, Y.; Watanabe, H.; Yomo, T.; Hohsaka, T.; Matsuura, T. Liposome-Based in Vitro Evolution of Aminoacyl-TRNA Synthetase for Enhanced Pyrrolysine Derivative Incorporation. *ChemBioChem* **2015**, *16*, 1797−1802.
(32) Silverman, A. D.; Karim, A. S.; Jewett, M. C. Cell-Free Gene Expression: An Expanded Repertoire of Applications. *Nat. Rev. Genet.* **2020**, *21*, 151−170.
(33) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. Cell-Free Translation Reconstituted with Purified Components. *Nat. Biotechnol.* **2001**, *19*, 751−755.
(34) Dodevski, I.; Markou, G. C.; Sarkar, C. A. Conceptual and Methodological Advances in Cell-Free Directed Evolution. *Curr. Opin. Struct. Biol.* **2015**, *33*, 1−7.
(35) Tawfik, D. S.; Griffiths, A. D. Man-Made Cell-like Compartments for Molecular Evolution. *Nat. Biotechnol.* **1998**, *16*, 652−656.
(36) Sunami, T.; Sato, K.; Matsuura, T.; Tsukada, K.; Urabe, I.; Yomo, T. Femtoliter Compartment in Liposomes for in Vitro Selection of Proteins. *Anal. Biochem.* **2006**, *357*, 128−136.
(37) Cho, E.; Lu, Y. Compartmentalizing Cell-Free Systems: Toward Creating Life-Like Artificial Cells and Beyond. *ACS Synth. Biol.* **2020**, *9*, 2881−2901.
(38) Nishikawa, T.; Sunami, T.; Matsuura, T.; Ichihashi, N.; Yomo, T. Construction of a Gene Screening System Using Giant Unilamellar Liposomes and a Fluorescence-Activated Cell Sorter. *Anal. Chem.* **2012**, *84*, 5017−5024.

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203

(39) Uyeda, A.; Nakayama, S.; Kato, Y.; Watanabe, H.; Matsuura, T. Construction of an in Vitro Gene Screening System of the *E. Coli* EmrE Transporter Using Liposome Display. *Anal. Chem.* **2016**, *88*, 12028−12035.
(40) Zhang, Y.; Minagawa, Y.; Kizoe, H.; Miyazaki, K.; Iino, R.; Ueno, H.; Tabata, K. V.; Shimane, Y.; Noji, H. Accurate High-Throughput Screening Based on Digital Protein Synthesis in a Massively Parallel Femtoliter Droplet Array. *Sci. Adv.* **2019**, *5*, No. eaav8185.
(41) Stögbauer, T.; Windhager, L.; Zimmer, R.; Rädler, J. O. Experiment and Mathematical Modeling of Gene Expression Dynamics in a Cell-Free System. *Integr. Biol.* **2012**, *4*, 494−501.
(42) Chory, J.; Kaplan, S. The in Vitro Transcription-Translation of DNA and RNA Templates by Extracts of Rhodopseudomonas Sphaeroides. Optimization and Comparison of Template Specificity with Escherichia Coli Extracts and in Vivo Synthesis. *J. Biol. Chem.* **1982**, *257*, 15110−15121.
(43) Niederholtmeyer, H.; Xu, L.; Maerkl, S. J. Real-Time MRNA Measurement during an in Vitro Transcription and Translation Reaction Using Binary Probes. *ACS Synth. Biol.* **2013**, *2*, 411−417.
(44) Nourian, Z.; Danelon, C. Linking Genotype and Phenotype in Protein Synthesizing Liposomes with External Supply of Resources. *ACS Synth. Biol.* **2013**, *2*, 186−193.
(45) Doerr, A.; De Reus, E.; Van Nies, P.; Van Der Haar, M.; Wei, K.; Kattan, J.; Wahl, A.; Danelon, C. Modelling Cell-Free RNA and Protein Synthesis with Minimal Systems. *Phys. Biol.* **2019**, *16*, No. 025001.
(46) Ueno, H.; Sawada, H.; Soga, N.; Sano, M.; Nara, S.; Tabata, K. V.; Su'etsugu, M.; Noji, H. Amplification of over 100 Kbp DNA from Single Template Molecules in Femtoliter Droplets. *ACS Synth. Biol.* **2021**, *10*, 2179−2186.
(47) Libicher, K.; Hornberger, R.; Heymann, M.; Mutschler, H. In Vitro Self-Replication and Multicistronic Expression of Large Synthetic Genomes. *Nat. Commun.* **2020**, *11*, No. 904.
(48) 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*, 3507−3517.
(49) Sakatani, Y.; Ichihashi, N.; Kazuta, Y.; Yomo, T. A Transcription and Translation-Coupled DNA Replication System Using Rolling-Circle Replication. *Sci. Rep.* **2015**, *5*, No. 10404.
(50) Sakatani, Y.; Yomo, T.; Ichihashi, N. Self-Replication of Circular DNA by a Self-Encoded DNA Polymerase through Rolling-Circle Replication and Recombination. *Sci. Rep.* **2018**, *8*, No. 13089.
(51) Mazutis, L.; Araghi, A. F.; Miller, O. J.; Baret, J.-C.; Frenz, L.; Janoshazi, A.; Valé, R. T.; Miller, B. J.; Hutchison, J. B.; Link, D.; Griffiths, A. D.; Ryckelynck, M. Droplet-Based Microfluidic Systems for High-Throughput Single DNA Molecule Isothermal Amplification and Analysis. *Anal. Chem.* **2009**, *81*, 4813−4821.
(52) Fallah-Araghi, A.; Baret, J. C.; Ryckelynck, M.; Griffiths, A. D. A Completely in Vitro Ultrahigh-Throughput Droplet-Based Microfluidic Screening System for Protein Engineering and Directed Evolution. *Lab Chip.* **2012**, *12*, 882−891.
(53) Galinis, R.; Stonyte, G.; aidotas Kiseliovas, V.; apolas Zilionis, R.; Studer, S.; Hilvert, D.; Janulaitis, A.; Mazutis, L. DNA Nanoparticles for Improved Protein Synthesis In Vitro. *Angew. Chem., Int. Ed.* **2016**, *55*, 3120−3123.
(54) Paul, S.; Stang, A.; Lennartz, K.; Tenbusch, M.; Überla, K. Selection of a T7 Promoter Mutant with Enhanced in Vitro Activity by a Novel Multi-Copy Bead Display Approach for in Vitro Evolution. *Nucleic Acids Res.* **2013**, *41*, No. e29.
(55) Diamante, L.; Gatti-Lafranconi, P.; Schaerli, Y.; Hollfelder, F. In Vitro Affinity Screening of Protein and Peptide Binders by Megavalent Bead Surface Display. *Protein Eng. Des. Sel.* **2013**, *26*, 713−724.
(56) Lindenburg, L.; Huovinen, T.; Van De Wiel, K.; Herger, M.; Snaith, M. R.; Hollfelder, F. Split & Mix Assembly of DNA Libraries for Ultrahigh Throughput On-Bead Screening of Functional Proteins. *Nucleic Acids Res.* **2020**, *48*, No. e63.
(57) Plesa, C.; Sidore, A. M.; Lubock, N. B.; Zhang, D.; Kosuri, S. Multiplexed Gene Synthesis in Emulsions for Exploring Protein Functional Landscapes. *Science* **2018**, *359*, 343−347.
(58) Restrepo Sierra, A. M.; Arold, S. T.; Grünberg, R. Efficient Multi-Gene Expression in Cell-Free Droplet Microreactors. *PLoS One* **2022**, *17*, No. e0260420.
(59) Mencía, M.; Gella, P.; Camacho, A.; De Vega, M.; Salas, M. Terminal Protein-Primed Amplification of Heterologous DNA with a Minimal Replication System Based on Phage Φ29. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 18655−18660.
(60) Rodríguez, I.; Lázaro, J. M.; Blanco, L.; Kamtekar, S.; Berman, A. J.; Wang, J.; Steitz, T. A.; Salas, M.; De Vega, M. A Specific Subdomain in Φ29 DNA Polymerase Confers Both Processivity and Strand-Displacement Capacity. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6407−6412.
(61) Kamtekar, S.; Berman, A. J.; Wang, J.; Lázaro, J. M.; De Vega, M.; Blanco, L.; Salas, M.; Steitz, T. A. Insights into Strand Displacement and Processivity from the Crystal Structure of the Protein-Primed DNA Polymerase of Bacteriophage Φ29. *Mol. Cell.* **2004**, *16*, 609−618.
(62) Blanco, L.; Bernads, A.; Lharo, J. M.; Martins, G.; Garmendia, C.; Salasn, M. Highly Efficient DNA Synthesis by the Phage Phi29 DNA Polymerase. *J. Biol. Chem.* **1989**, *264*, 8935−8940.
(63) Kamtekar, S.; Berman, A. J.; Wang, J.; Lá Zaro, J. M.; De Vega, M.; Blanco, L.; Salas, M.; Steitz, T. A. The Phi29 DNA Polymerase:Protein-Primer Structure Suggests a Model for the Initiation to Elongation Transition. *EMBO J.* **2006**, *25*, 1335−1343.
(64) Blanco, L.; Gutierrez, J.; Lazaro, J. M.; Bernad, A.; Salas, M. Replication of Phage φ29 DNA in Vitro: Role of the Viral Protein P6 in Initiation and Elongation. *Nucleic Acids Res.* **1986**, *14*, 4923−4937.
(65) Freire, R.; Serrano, M.; Salas, M.; Hermoso, J. M. Activation of Replication Origins in Phi29-Related Phages Requires the Recognition of Initiation Proteins to Specific Nucleoprotein Complexes. *J. Biol. Chem.* **1996**, *271*, 31000−31007.
(66) Peñalva, M. A.; Salas, M. Initiation of Phage Phi29 DNA Replication in Vitro: Formation of a Covalent Complex between the Terminal Protein, P3, and 5′-DAMP. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5522−5526.
(67) Méndez, J.; Blanco, L.; Salas, M. Protein-Primed DNA Replication: A Transition between Two Modes of Priming by a Unique DNA Polymerase. *EMBO J.* **1997**, *16*, 2519−2527.
(68) Martìin, G.; Làzaro, J.; Mèndez, E.; Salas, M. Characterization of the Phage π29 Protein P5 as a Single-Stranded DNA Binding Protein. Function in π29 DNA-Protein P3 Replication. *Nucleic Acids Res.* **1989**, *17*, 3663−3672.
(69) Serrano, M.; Gutierrez, J.; Prieto, I.; Hermoso, J. M.; Salas, M. Signals at the Bacteriophage Phi 29 DNA Replication Origins Required for Protein P6 Binding and Activity. *EMBO J.* **1989**, *8*, 1879−1885.
(70) Blanco, L.; Prieto, I.; Gutierrez, J.; Bernad, A.; Lazaro, J. M.; Hermoso, J. M.; Salas, M. Effect of NH4+ Ions on Phi29 DNA-Protein P3 Replication: Formation of a Complex between the Terminal Protein and the DNA Polymerase. *J. Virol.* **1987**, *61*, 3983−3991.
(71) Carlson, E. D.; Gan, R.; Hodgman, C. E.; Jewett, M. C. Cell-Free Protein Synthesis: Applications Come of Age. *Biotechnol. Adv.* **2012**, *30*, 1185−1194.
(72) 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*, No. 1898.
(73) Meijer, W. J. J.; Horcajadas, J. A.; Salas, M. Φ29 Family of Phages. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 261−287.
(74) Raetz, C. R. H. Phosphatidylserine Synthetase Mutants of Escherichia Coli. Genetic Mapping and Membrane Phospholipid Composition. *J. Biol. Chem.* **1976**, *251*, 3242−3249.
(75) DeChavigny, A.; Heacock, P. N.; Dowhan, W. Sequence and Inactivation of the Pss Gene of Escherichia Coli. Phosphatidylethanolamine May Not Be Essential for Cell Viability. *J. Biol. Chem.* **1991**, *266*, 5323−5332.
(76) Larson, T. J.; Dowhan, W. Ribosomal-Associated Phosphatidylserine Synthetase from Escherichia Coli: Purification by Substrate-Specific Elution from Phosphocellulose Using Cytidine 5′-Diphospho-1,2-Diacyl-Sn-Glycerol. *Biochemistry* **1976**, *15*, 5212−5218.
(77) Hirabayashi, T.; Larson, T. J.; Dowhan, W. Membrane-Associated Phosphatidylglycerophosphate Synthetase from Escherichia

Coli: Purification by Substrate Affinity Chromatography on Cytidine 5′ Diphospho-1,2-Diacyl-Sn-Glycerol Sepharose. *Biochemistry* **1976**, *15*, 5205−5211.
(78) Rowlett, V. W.; Margolin, W. The Min System and Other Nucleoid-Independent Regulators of Z Ring Positioning. *Front. Microbiol.* **2015**, *6*, 478.
(79) Rowlett, V. W.; Margolin, W. The Bacterial Min System. *Curr. Biol.* **2013**, *23*, R553−R556.
(80) Lu, W. C.; Levy, M.; Kincaid, R.; Ellington, A. D. Directed Evolution of the Substrate Specificity of Biotin Ligase. *Biotechnol. Bioeng.* **2014**, *111*, 1071−1081.
(81) Borkowski, O.; Koch, M.; Zettor, A.; Pandi, A.; Batista, A. C.; Soudier, P.; Faulon, J. L. Large Scale Active-Learning-Guided Exploration for in Vitro Protein Production Optimization. *Nat. Commun.* **2020**, *11*, No. 1872.
(82) Ishikawa, K.; Sato, K.; Shima, Y.; Urabe, I.; Yomo, T. Expression of a Cascading Genetic Network within Liposomes. *FEBS Lett.* **2004**, *576*, 387−390.
(83) Kobori, S.; Ichihashi, N.; Kazuta, Y.; Yomo, T. A Controllable Gene Expression System in Liposomes That Includes a Positive Feedback Loop. *Mol. Biosyst.* **2013**, *9*, 1282−1285.
(84) Maddalena, L. L. D.; Niederholtmeyer, H.; Turtola, M.; Swank, Z. N.; Belogurov, G. A.; Maerkl, S. J. GreA and GreB Enhance Expression of Escherichia Coli RNA Polymerase Promoters in a Reconstituted Transcription-Translation System. *ACS Synth. Biol.* **2016**, *5*, 929−935.
(85) Elías-Arnanz, M.; Salas, M. Bacteriophage Phi29 DNA Replication Arrest Caused by Codirectional Collisions with the Transcription Machinery. *EMBO J.* **1997**, *16*, 5775−5783.
(86) Elías-Arnanz, M.; Salas, M. Resolution of Head-on Collisions between the Transcription Machinery and Bacteriophage Phi29 DNA Polymerase Is Dependent on RNA Polymerase Translocation. *EMBO J.* **1999**, *18*, 5675−5682.
(87) Marshall, R.; Noireaux, V. Synthetic Biology with an All E. Coli TXTL System: Quantitative Characterization of Regulatory Elements and Gene Circuits. In *Synthetic Biology*, Methods in Molecular Biology; Springer: New York, 2018; Vol. *1772*, pp 61−93.
(88) Sitaraman, K.; Esposito, D.; Klarmann, G.; Le Grice, S. F.; Hartley, J. L.; Chatterjee, D. K. A Novel Cell-Free Protein Synthesis System. *J. Biotechnol.* **2004**, *110*, 257−263.
(89) Marshall, R.; Maxwell, C. S.; Collins, S. P.; Beisel, C. L.; Noireaux, V. Short DNA Containing χ Sites Enhances DNA Stability and Gene Expression in E. Coli Cell-Free Transcription-Translation Systems. *Biotechnol. Bioeng.* **2017**, *114*, 2137−2141.
(90) Batista, A. C.; Levrier, A.; Soudier, P.; Voyvodic, P. L.; Achmedov, T.; Reif-Trauttmansdorff, T.; DeVisch, A.; Cohen-Gonsaud, M.; Faulon, J. L.; Beisel, C. L.; Bonnet, J.; Kushwaha, M. Differentially Optimized Cell-Free Buffer Enables Robust Expression from Unprotected Linear DNA in Exonuclease-Deficient Extracts. *ACS Synth. Biol.* **2022**, *11*, 732−746.
(91) Nourian, Z.; Roelofsen, W.; Danelon, C. Triggered Gene Expression in Fed-Vesicle Microreactors with a Multifunctional Membrane. *Angew. Chem., Int. Ed.* **2012**, *51*, 3114−3118.
(92) Van de Cauter, L.; Fanalista, F.; Van Buren, L.; De Franceschi, N.; Godino, E.; Bouw, S.; Danelon, C.; Dekker, C.; Koenderink, G. H.; Ganzinger, K. A. Optimized CDICE for Efficient Reconstitution of Biological Systems in Giant Unilamellar Vesicles. *ACS Synth. Biol.* **2021**, *10*, 1690−1702.
(93) Gonzales, D. T.; Yandrapalli, N.; Robinson, T.; Zechner, C.; Tang, T. Y. D. Cell-Free Gene Expression Dynamics in Synthetic Cell Populations. *ACS Synth. Biol.* **2022**, *11*, 205−215.
(94) Vogele, K.; Frank, T.; Gasser, L.; Goetzfried, M. A.; Hackl, M. W.; Sieber, S. A.; Simmel, F. C.; Pirzer, T. Towards Synthetic Cells Using Peptide-Based Reaction Compartments. *Nat. Commun.* **2018**, *9*, No. 3862.
(95) Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D. Emergence of a Catalytic Tetrad during Evolution of a Highly Active Artificial Aldolase. *Nat. Chem.* **2017**, *9*, 50−56.
(96) Ghadessy, F. J.; Ong, J. L.; Holliger, P. Directed Evolution of Polymerase Function by Compartmentalized Self-Replication. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4552−4557.
(97) Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D. Directed Evolution of Genetic Parts and Circuits by Compartmentalized Partnered Replication. *Nat. Biotechnol.* **2014**, *32*, 97−101.
(98) de Vega, M.; Lazaro, J. M.; Salas, M.; Blanco, L. Primer-Terminus Stabilization at the 3′−5′ Exonuclease Active Site of Phi29 DNA Polymerase. Involvement of Two Amino Acid Residues Highly Conserved in Proofreading DNA Polymerases. *EMBO J.* **1996**, *15*, 1182−1192.
(99) Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343−345.
(100) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26*, 966−968.

https://doi.org/10.1021/acssynbio.2c00668
ACS Synth. Biol. 2023, 12, 1187−1203