ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

Cell-Free Gene Expression Dynamics in Synthetic Cell Populations

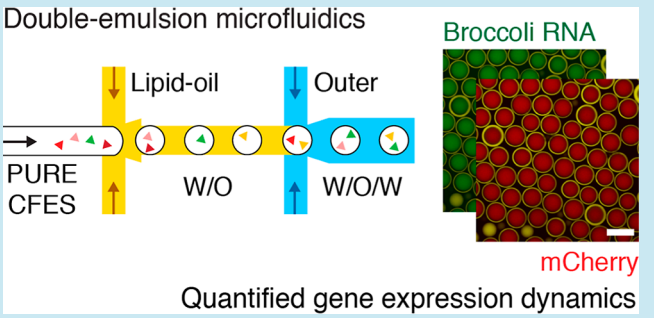
David T. Gonzales, Naresh Yandrapalli, Tom Robinson, Christoph Zechner,* and T-Y. Dora Tang*

Cite This: *ACS Synth. Biol.* 2022, 11, 205−215

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The ability to build synthetic cellular populations from the bottom-up provides the groundwork to realize minimal living tissues comprising single cells which can communicate and bridge scales into multicellular systems. Engineered systems made of synthetic micron-sized compartments and integrated reaction networks coupled with mathematical modeling can facilitate the design and construction of complex and multiscale chemical systems from the bottom-up. Toward this goal, we generated populations of monodisperse liposomes encapsulating cell-free expression systems (CFESs) using double-emulsion microfluidics and quantified transcription and translation dynamics within individual synthetic cells of the population using a fluorescent Broccoli RNA aptamer and mCherry protein reporter. CFE dynamics in bulk reactions were used to test different coarse-grained resource-limited gene expression models using model selection to obtain transcription and translation rate parameters by likelihood-based parameter estimation. The selected model was then applied to quantify cell-free gene expression dynamics in populations of synthetic cells. In combination, our experimental and theoretical approaches provide a statistically robust analysis of CFE dynamics in bulk and monodisperse synthetic cell populations. We demonstrate that compartmentalization of CFESs leads to different transcription and translation rates compared to bulk CFE and show that this is due to the semipermeable lipid membrane that allows the exchange of materials between the synthetic cells and the external environment.



1. INTRODUCTION

The establishment of synthetic multicellular systems that can sustain out-of-equilibrium behavior is a major challenge in bottom-up synthetic biology. This requires the integration of chemical reaction networks for intercellular communication between synthetic cells within populations. Recent examples of bottom-up multicellular systems demonstrate how simple chemical building blocks can be used to make functional systems that scale from molecules to cells to synthetic cell communities[1,2] or tissues.[3,4] This provides a route for characterizing how cell−cell communication could lead to collective behavior in a minimal and multicellular context. Importantly, relative to biological systems, bottom-up synthetic cells and cell populations are highly amenable to perturbations and modifications for experiments and quantitative analysis. However, transitioning from single synthetic cells to multicellular systems is challenging because of our poor ability to control molecular assembly beyond a certain number of modules. Highly engineered systems and controllable methodologies coupled with mathematical modeling can facilitate building more complex and multiscale chemical systems from the bottom-up. It has already been shown how models of cell-free expression systems (CFESs) can help unravel the modular response of highly complex systems responsible for supporting out-of-equilibrium behavior.[5]

CFESs encapsulated in lipid vesicles have been established as one of the most popular and utilized types of synthetic cells as they encompass both cellular compartmentalization and the central biological dogma of transcription and translation in a minimalistic fashion. These minimal synthetic cells have demonstrated simple transcription and translation pathways,[6] gene expression cascades,[7] gene expression bursting,[8] noise from macromolecular crowding,[9] stochastic gene expression,[10] genetic circuits for intercellular communication,[11] and CFESs coupled with other metabolic processes including ATP production[12] and DNA replication,[13] demonstrating their potential as a platform to build communicating populations of synthetic cells. Despite these successes, compartmentalized cell-free gene expression dynamics has not yet been fully quantified and modeled. This can provide a simplified system that focuses on the effect of compartmentalization on gene expression dynamics.

Even though CFESs are dramatically reduced in complexity compared to biological cells, they still contain as many as 37

Received: August 6, 2021
Published: January 4, 2022

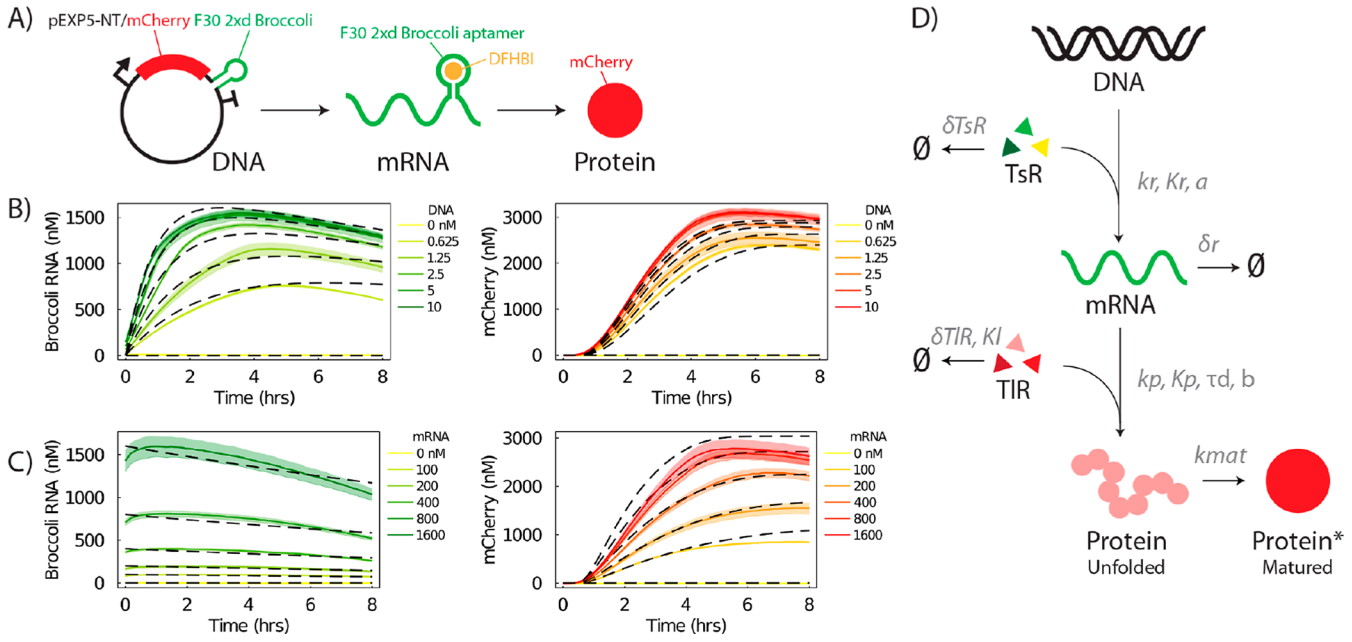


Figure 1. Monitoring transcription and translation dynamics in cell-free expression. (A) Construct of the pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid containing a constitutive T7 RNAP-mediated promoter expressing 6xHis mCherry with an F30-2xdBroccoli RNA aptamer tag. The small-molecule dye DFHBI becomes fluorescent upon binding with the Broccoli RNA aptamer. (B) mRNA and mCherry protein expression levels over time from bulk PURExpress CFES titrated with varying concentrations of pEXP5-NT/6xHis mCherry F30-2xdBroccoli DNA plasmid. (C) mRNA and mCherry protein expression levels over time from bulk PURExpress CFESs titrated with varying concentrations of purified 6xHis mCherry F30-2xdBroccoli RNA transcripts. Solid lines and shaded areas correspond to mean and standard deviation values from triplicate experiments, respectively. Dashed lines are resource-limited CFES model fits. (D) Illustration of the resource-limited gene expression model for CFESs. Parameters are k_t: RNA transcription rate, K_t: dissociation constant between RNAP and DNA, δ_r: RNA degradation rate, k_p: protein translation rate, K_p: dissociation constant between ribosome and RNA, k_{mat}: mCherry maturation rate, δ_{TsR}: TsR degradation rate, δ_{TlR}: TlR degradation rate, K_l: Michaelis–Menten constant for TlR degradation, a: scaling factor for consumption of TsR with transcription, b: scaling factor for consumption of TlR with translation, and τ_d: time delay for protein translation.

enzymes and 32 small-molecule compounds or substrates in purified reconstituted cell-free systems[14,15] and even more for crude extract-based CFES. This makes it challenging to collect sufficient data to test existing models or to develop tractable models that rely on the knowledge of precise chemical species as a function of time. Fortunately, there is an increasing effort toward proteomic and metabolic profiling of CFES reactions that have the potential to provide quantitative molecular analysis to give details on reaction dynamics and limits on resources.[16−19] Alternatively, coarse-grained CFES models circumvent the need to measure all molecular species by focusing on only a few species such as DNA, RNA, proteins, RNA polymerases, and ribosomes. Several models have already been demonstrated to faithfully capture quantified cell-free gene expression dynamics in bulk solutions. For instance, several coarse-grained models have used first-order and Michaelis–Menten kinetics to describe cell-free transcription and translation dynamics[20−22] and extended to explicitly account for RNA polymerase and ribosome species.[23,24] Dynamic models that include the initiation, elongation, and termination steps of translation[25,26] or central carbon metabolism[27] have been used to identify bottlenecks in transcription and translation, which can be experimentally relieved to improve protein productivity in CFESs. These examples demonstrate how quantitative coarse-grained models can provide a better understanding of the CFES building blocks and will be crucial for further engineering of more complex synthetic multicellular systems.

As studies have shown differences in protein production by mammalian cell-free expression within water–oil emulsions,[28] quantifying and modeling compartmentalized cell-free gene expression could provide insights into the effect of encapsulation on gene expression. Methodologies to monitor mRNA and protein dynamics in bulk and semipermeable liposomes using cell-free systems have been demonstrated by utilizing fluorescence resonance energy transfer (FRET) donor–acceptor pairs for mRNA,[22,29] fluorescent proteins such as GFP and YFP,[30,31] and fluorescent Spinach RNA aptamers simultaneously with YFP[32,33] or mCherry.[9] However, expressed mRNA and protein levels in encapsulated CFESs have only been measured as either relative fluorescence units within the synthetic cells or quantified concentrations for mRNA or protein levels separately. To the best of our knowledge, an absolute and simultaneous quantification of both transcription and translation dynamics within liposome synthetic cell populations has not yet been presented. This is crucial for models of compartmentalized cell-free gene expression to provide a better understanding of the effects of physical processes and environments, such as trans-membrane diffusion and surface effects, on gene expression dynamics and provides the basis to model increasingly complex compartmentalized cell-free gene expression circuits. Notably, studies on the effects of liposome compartmentalization to CFESs has revealed that optimization of the outer solution and liposome permeability can increase expression yield and prolong gene expression activity;[6] CFESs encapsulated in small cell-sized volumes can result in stochastic gene expression[10] and rare

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

favorable phenotypes such as high expression,[30] and gene expression resources and macromolecular crowding can affect gene expression bursting.[8,9]

Liposome-encapsulated transcription and translation machinery can be generated by lipid swelling[30,34] or phase transfer methods.[8,10,33,35] These methods result in synthetic cells with large variations in cell size and gene expression profiles.[8,9,30,31] Bulk methods can be advantageous due to their accessibility and opportunity in exploring a large random space in terms of encapsulation and size,[30] and populations of vesicles can be generated without specialized equipment. However, it can also be advantageous to generate uniform populations of synthetic cells for reproducibility and predictability. This can be achieved by leveraging recent developments in microfluidics[36−38] and droplet printing technologies[39] to generate populations of synthetic cells with greater throughput, control, and uniformity compared to standard bulk methods.[40,41] Using microfluidic-generated synthetic cells has already been shown to be effective in generating monodisperse synthetic cells to study the effect of macromolecular crowding on gene expression without the use of synthetic crowding agents[42] and to qualitatively monitor Spinach2 RNA aptamer transcription dynamics.[43] This is especially important for quantitative approaches, as it enables the generation of statistically robust data that is amenable to accurate modeling.

In this study, we monitored and quantified both transcription and translation dynamics in bulk and liposome-encapsulated CFES reactions expressing Broccoli RNA aptamer and mCherry protein reporters. Using double-emulsion microfluidics,[37] we generated monodisperse populations of synthetic cells encapsulating cell-free gene expression systems. Fluorescent readings were converted into absolute concentration units by a standard calibration curve to obtain a simultaneous and quantitative characterization of transcription and translation. Bulk reaction results were used to develop and select from several models of resource-limited cell-free gene expression.[22] These models were compared to each other using the Akaike information criterion (AIC), and profile likelihood analysis was performed to quantify the kinetic parameters and their uncertainties. The best-ranking model was then used to compare rate parameters of gene expression dynamics between bulk experiments and synthetic cell populations. Overall, our work combines bottom-up assembly with robust modeling to provide a quantitative outlook of liposome-compartmentalized gene expression dynamics in synthetic cell populations. This has facilitated direct comparisons between experiments of bulk and compartmentalized CFES reactions and provides the basis for the design and construction of multicellular systems with increased levels of complexity.

2. RESULTS AND DISCUSSION

2.1. Monitoring Transcription and Translation Dynamics in Cell-Free Expression Systems. To monitor both transcription and translation dynamics in CFESs, we constructed the pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid. This plasmid consists of a constitutive T7 RNA polymerase-mediated promoter to express a red fluorescent mCherry protein and two copies of a dimeric Broccoli RNA aptamer stabilized by the F30 stem-loop[44,45] between the stop codon of mCherry and the terminator of the gene construct (Figure 1A). Binding of a small-molecule dye, 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), to the Broccoli RNA aptamer results in a green fluorescence signal. This allows simultaneous fluorescence monitoring of transcribed mRNA and reporter protein levels. PURExpress CFES reactions were titrated with pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid DNA or purified mRNA transcripts from the same plasmid. Reaction mixtures were incubated at 30 °C and monitored for mRNA and protein levels over time in a fluorescence well plate reader (Figure 1B,C). Relative fluorescence units were converted into nM concentration units using calibration curves from serial dilutions of Broccoli RNA aptamer and mCherry protein in the same reaction mix composition and acquisition settings (Section 5 in the Supporting Information). We observed typical profiles of gene expression in CFESs, where a signal is first detected from transcription of mRNA followed by translation of the mCherry protein. The gene expression profiles show a plateau at ∼3 h for mRNA and ∼5 h for protein. Rates of transcription reach a maximum at the initial point, while translation rates peak at 1.5−2 h (Figure S14 in the Supporting Information) and then gradually decrease. Our results also show that the endpoint protein concentrations increase with increasing plasmid DNA or mRNA transcript concentrations until a saturation concentration of 3.75 nM for plasmid DNA and 600 nM for mRNA transcript (Figure 1B,C). This indicates that gene expression rates and yield are dependent on both the consumption and degradation of resources. If gene expression was dependent solely on the consumption of resources, the final protein production would be constant regardless of DNA or mRNA input. This hypothesis is supported by previous work where the addition of ribosomes to PURExpress after exhaustion restores gene expression activity[22] and a delayed addition of a DNA template into PURExpress after incubation results in reduced rates and yield of gene expression.[22,26]

2.2. Resource-Limited Gene Expression Model for Cell-Free Expression. To describe the dynamics of cell-free gene expression, a coarse-grained model based on Stögbauer et al. (2012)[22] was developed to quantitatively compare results across experiments and literature values. This model accounts for both transcription and translation dynamics driven by a limited pool of resources for gene expression. Transcription (TsR) and translation resources (TlR) are assigned unitless quantities initialized at 1 and then gradually decreased to 0 as they are consumed by transcription or translation and degradation. These species serve as a phenomenological proxy to account for the cumulative effect of different limiting factors that fuel transcription and translation processes, such as RNA polymerase, ribosome concentrations, NTP, amino acids, and other energy resources. Based on this model, we generated several candidate CFES models composed of a system of delay and ordinary differential equations. We used our bulk experimental results from both transcription and translation dynamics and mCherry maturation experiments (Section 7 in the Supporting Information), to guide model selection. Candidate models were ranked among each other using the AIC.[46] Profile likelihoods were then used to determine the parameter identifiability and confidence intervals (CIs) for each of the candidate models.[47,48] The best-scoring model resulting from this analysis is shown in eqs 1−6 (Section 9 in the Supporting Information) below.

$$\frac{\mathrm{d}[\mathrm{DNA}]}{\mathrm{d}t} = 0 \tag{1}$$

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

Table 1. Parameter Estimates and Likelihood-Based 95% CI from the Resource-Limited Gene Expression Model Fitting on Bulk DNA and RNA Titration Experiments ($\hat{\theta}_{Bulk}$) and Synthetic Cell Population DNA Titration Experiments ($\hat{\theta}_{Cell}$)[a]

parameter	description	$\hat{\theta}_{bulk}$	95% CI	$\hat{\theta}$cell	95% CI	units
k_r	RNA transcription rate	2894	2728 - 3674	1899	1631 - 3537	nM/h
K_r	Dissociation constant between RNAP and DNA	3.67	2.89 - 5.68	8.86	6.97 - 18.66	nM
δ_r	RNA degradation rate	0.0392	0.0361 - 0.0422	0.0081	0.00239 - 0.0143	1/h
k_p	Protein translation rate	2568	2211 - 3108	1954	1617 - 2696	nM/h
K_p	Dissociation constant between ribosome and RNA	703	530 - 1347	1319	819 - 2038	nM
k_{mat}	mCherry maturation rate	2.15	(±0.12)	2.15	(±0.12)	1/h
δ_{TsR}	TsR degradation rate	0.231	0.171 - 0.298	0.154	0.136 - 0.175	1/h
δ_{TlR}	TlR degradation rate	0.0884	0.0441 - 0.1187	0.244	0.184 - 0.684	1/h
K_l	Michaelis−Menten constant for TlR degradation	1.21×10^{-6}	−∞ - +∞	0.232	−∞ - 0.713	
A	scaling factor for consumption of TsR with transcription	4.45×10^{-4}	4.18×10^{-4} - 4.57×10^{-4}	6.60×10^{-4}	6.21×10^{-4} - 6.74×10^{-4}	
b	scaling factor for consumption of TlR with translation	1.78×10^{-4}	1.18×10^{-4} - 2.42×10^{-4}	4.46×10^{-13}	−∞ - +∞	
τ_d	time delay for protein translation	0.433	0.254 - 0.560	0.0576	−∞ - 0.279	h
τ_l	time lag between the reaction start and data collection	2.81×10^{-9}	−∞ - +∞	0.457	0.342 - 0.535	h

[a]Parameters with CIs at −∞ and/or +∞ are non/weakly-identifiable within one order of magnitude from $\hat{\theta}$.

$$\frac{d[RNA]}{dt} = \frac{k_r TsR[DNA]}{K_r + [DNA]} - \delta_r[RNA] \quad (2)$$

$$\frac{d[protein]}{dt} = \frac{k_p TlR[RNA](t - \tau_d)}{K_p + [RNA](t - \tau_d)} - k_{mat}[protein] \quad (3)$$

$$\frac{d[protein^*]}{dt} = k_{mat}[protein^*] \quad (4)$$

$$\frac{dTsR}{dt} = -\frac{ak_r TsR[DNA]}{K_r + [DNA]} - \delta_{TsR} TsR \quad (5)$$

$$\frac{dTlR}{dt} = -\frac{bk_p TlR[RNA](t - \tau_d)}{K_p + [RNA](t - \tau_d)} - \frac{\delta_{TlR} TlR}{K_l + TlR} \quad (6)$$

This model uses Michaelis−Menten-type kinetics for transcription and translation. Translation is additionally modeled by a delay differential equation with a time delay (τ_d) to account for the time delay of protein expression observed in our mRNA titration experiments. Transcription and translation resources (TsR and TlR, respectively) are consumed by transcription and translation processes and degraded independently. These are consumed during transcription and translation with a scaling factor, a and b, respectively. Both resources, also, spontaneously degrade with first-order and Michaelis−Menten kinetics for TsR and TlR, respectively. RNA degradation and mCherry protein maturation are assigned first-order reactions (see model derivation in Section 8 of the Supporting Information). In contrast to the previously published model,[22] TsR degradation was included to account for the independent exhaustion of transcription resources. Lastly, we included a time lag (τ_l) in the fitting procedure of the model to account for the time between starting the CFES reaction and acquiring the first data point (Section 9 in the Supporting Information). This was negligible for our bulk experiments that took less than 10 min from adding the DNA or RNA template into the CFE bulk reactions to acquiring the first data points in the plate reader. However, it was important for the encapsulated experiments which had a lag period between sample preparation and measurements of ∼30 min. The model fit and optimized rate parameters and CIs are shown in Figure 1B,C (dashed lines) and Table 1. All parameters except for K_l and τ_l are well-identifiable. The estimates of K_l and τ_l turn out to be very low, and varying these parameters within one order of magnitude does not significantly affect the model fit. Overall, the model captures the general behavior of gene expression dynamics across different initial DNA and RNA conditions. The remaining quantitative mismatch is likely due to additional chemical complexity not captured by our coarse-grained model. The mCherry maturation rate parameter k_{mat} was determined independently using a protein maturation assay[49] at 2.15 ± 0.12 h^{-1} (Section 7 in the Supporting Information). This corresponds to a maturation half-time ($t_{0.5}$) of 19.31 ± 2.24 min, which is also comparable to previous reports of mCherry maturation at 15 min in *Escherichia coli*.[50] Assuming that a standard PURExpress reaction contains 100 nM T7 RNA polymerase[14] and 2.4 μM ribosomes (NEB), T7 RNA polymerase transcription and ribosome translation rates are approximately 8.2−11.1 NTP/s and 0.20−0.28 amino acid/s, respectively (calculated from k_r = 2728−3674 nM/h for a 1087 bp transcript and k_p = 2211−3108 nM/h for a 777 aa protein in Table 1). These values are lower than the reported *in vivo* rates in *E. coli* bacterial cells (230 ± 20 NTP/s[51] and 8-18 amino acid/s[52]). However, the polymerase transcription rates and ribosome translation rates are similar to previous work in PURExpress expressing GFP at 37 °C (2.2 NTP/s and 0.03 amino acid/s, respectively).[22] Using a FRET sensor to measure RNA transcription in PURExpress, initial transcription rates from 10 nM DNA plasmid template were previously measured at ∼7 nM/min.[29] This is also comparable to our initial transcription rate measurements at 15.9 nM/min for 10 nM DNA plasmid (Figure S14 in the Supporting Information). The differences could be attributed to different reaction conditions, T7 RNA polymerase concentrations, the encoding gene, and/or batch-to-batch variability of the expression system.

2.3. Production of Synthetic Cell Populations with Low Variability. Having established a quantitative model for cell-free gene expression in bulk reactions, we next wanted to test its applicability on populations of compartmentalized reactions. To this end, we encapsulated PURExpress CFESs in lipid-based synthetic cell populations using either an inverse

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

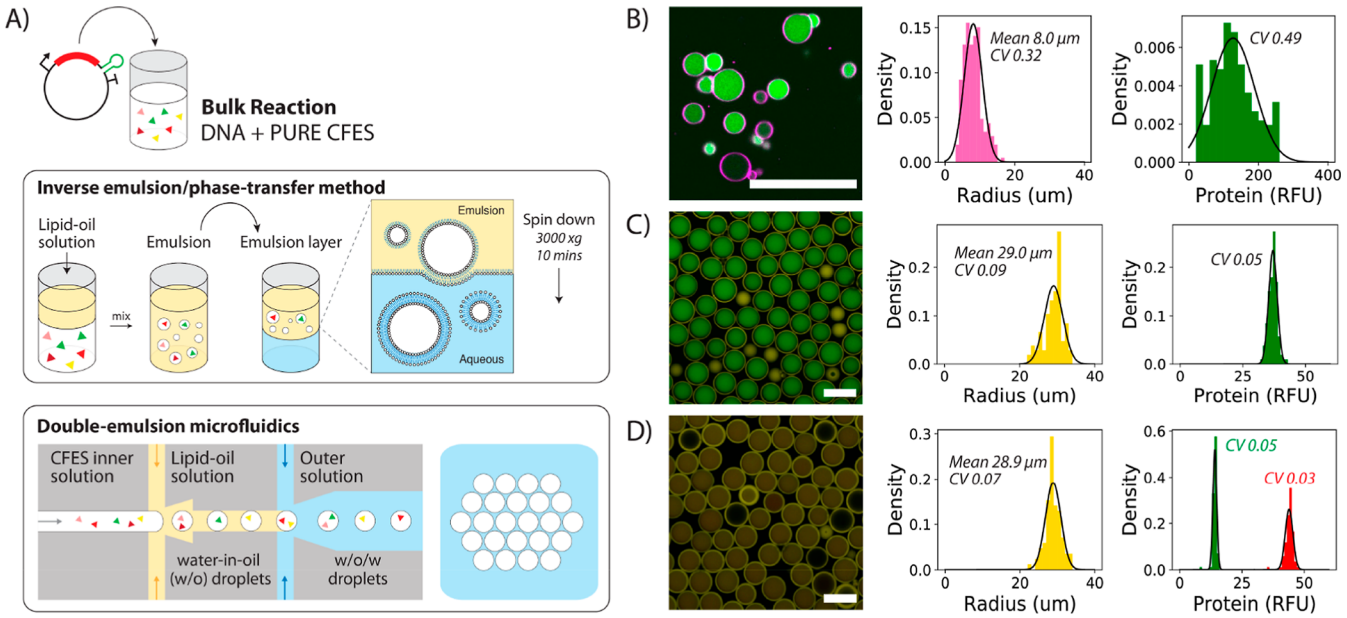


Figure 2. Variability in synthetic cell populations. (A) Schematic of the bulk inverse emulsion phase transfer method and double-emulsion microfluidics to generate liposomes or synthetic cells. (B) Synthetic cell population expressing eGFP protein from 1.17 nM pEXP5-NT/6xHis eGFP plasmid DNA generated using the bulk inverse emulsion phase transfer method. (C) Microfluidic-generated synthetic cells expressing eGFP protein from 4.5 nM pEXP5-NT/6xHis eGFP F30-2xdBroccoli plasmid DNA. (D) Merged image of the synthetic cell population expressing both eGFP and mCherry protein from two plasmids (4.5 nM pEXP5-NT/6xHis eGFP and 4.5 nM pEXP5-NT/6xHis mCherry plasmid DNA). Endpoint histograms of radius and protein RFU are plotted alongside each of the synthetic cell populations (B–D). The number of cells analyzed is 206, 106, and 85 for (B–D), respectively. Black lines are Gaussian distributions obtained by fitting mean and variance of the data. These experiments show the relative levels of expressed protein and do not refer to absolute concentrations. RFU values between the microfluidic-generated synthetic cells in (C,D), but not the inverse emulsion-made synthetic cells, are comparable, as these images were acquired using the same microscopy settings. However, CV values can be compared across all populations. All images are taken at the endpoint after 12 h of incubation at 30 °C using confocal microscopy with a 40× objective for (B) and 10× objective for (C,D). Scale bars are all 100 μm. Calibrated units of (C,D) are shown in Section 15 of the Supporting Information.

emulsion/phase transfer method[53] or a double-emulsion microfluidic methodology[36,37] (Figure 2A) (Materials and Methods and Sections 10–12 in the Supporting Information). The inner solution was composed of the PURExpress CFES and a plasmid DNA for constitutive T7 RNAP-mediated expression of a fluorescent protein gene (eGFP or mCherry). Confocal microscopy images for the inverted emulsion and microfluidic-generated synthetic cells were segmented to obtain relative fluorescence units (RFU) of the expressed protein in single cells in each population. These were then used to calculate the coefficient of variation (CV) of the distribution of the expressed protein in each cell population. The CV allows comparison of the variability of distributions with different scales of measurement. The inverted emulsion method generated liposomes with a mean radius of 8.0 μm and a CV of 0.32 (Figure 2B). In comparison, the microfluidic-generated synthetic cells were larger with a mean radius of 29.0 μm and exhibited lower size variation with a CV of 0.09 (Figure 2C) as expected. Protein expression in the inverted emulsion-generated synthetic cells also showed a greater variation (CV 0.49, Figure 2B) compared to synthetic cells produced in microfluidics (CV 0.05) (Figure 2C). These results are in agreement with previous studies of inverse emulsion-generated cells with expressed protein concentration CVs ranging from 0.20 to 0.80.[10] We further show that simultaneous encapsulation of two plasmids in a microfluidic-generated synthetic cell population results in expression of both eGFP and mCherry proteins in each cell at a consistent ratio (3.11 ± 0.133 eGFP/mCherry RFU) (Figure 2D). This demonstrates the robustness of our synthetic cell production where the inner CFES solution is well-mixed and the microfluidic method maintains the homogeneity throughout encapsulation. The increased variance in phase transfer-generated cells is most likely due to a combined result of fluctuations in cell size and encapsulation. In contrast, the synthetic cell populations generated using double-emulsion microfluidics resulted in larger and more uniform cell populations, making them highly suitable for our quantitative analysis. In addition, it was also observed that the fluorescence from expression of the pEXP5-NT/6xHis eGFP plasmid is decreased in the two-plasmid synthetic cells (mean RFU 14.1, Figure 2D) compared to the single-plasmid synthetic cells (mean RFU 37.0, Figure 2C) (see also Section 15 in the Supporting Information). This is a result of gene expression resources being split between the expression of both eGFP and mCherry proteins in the two-plasmid synthetic cells.

2.4. Quantitative Transcription and Translation Dynamics in Synthetic Cell Populations. Using our microfluidic platform, we generated synthetic cell populations comprising large populations of monodisperse liposome-encapsulated CFESs and quantified RNA and protein levels over time to study transcription and translation dynamics using fluorescence microscopy methods. To alleviate non-identifiabilities during model fitting, we prepared three populations of synthetic cells with different DNA concentrations (1.75, 3.5, and 7.0 nM pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid DNA) during one microfluidic session from one batch of CFES master mix and outer feeding buffer solution.

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205–215

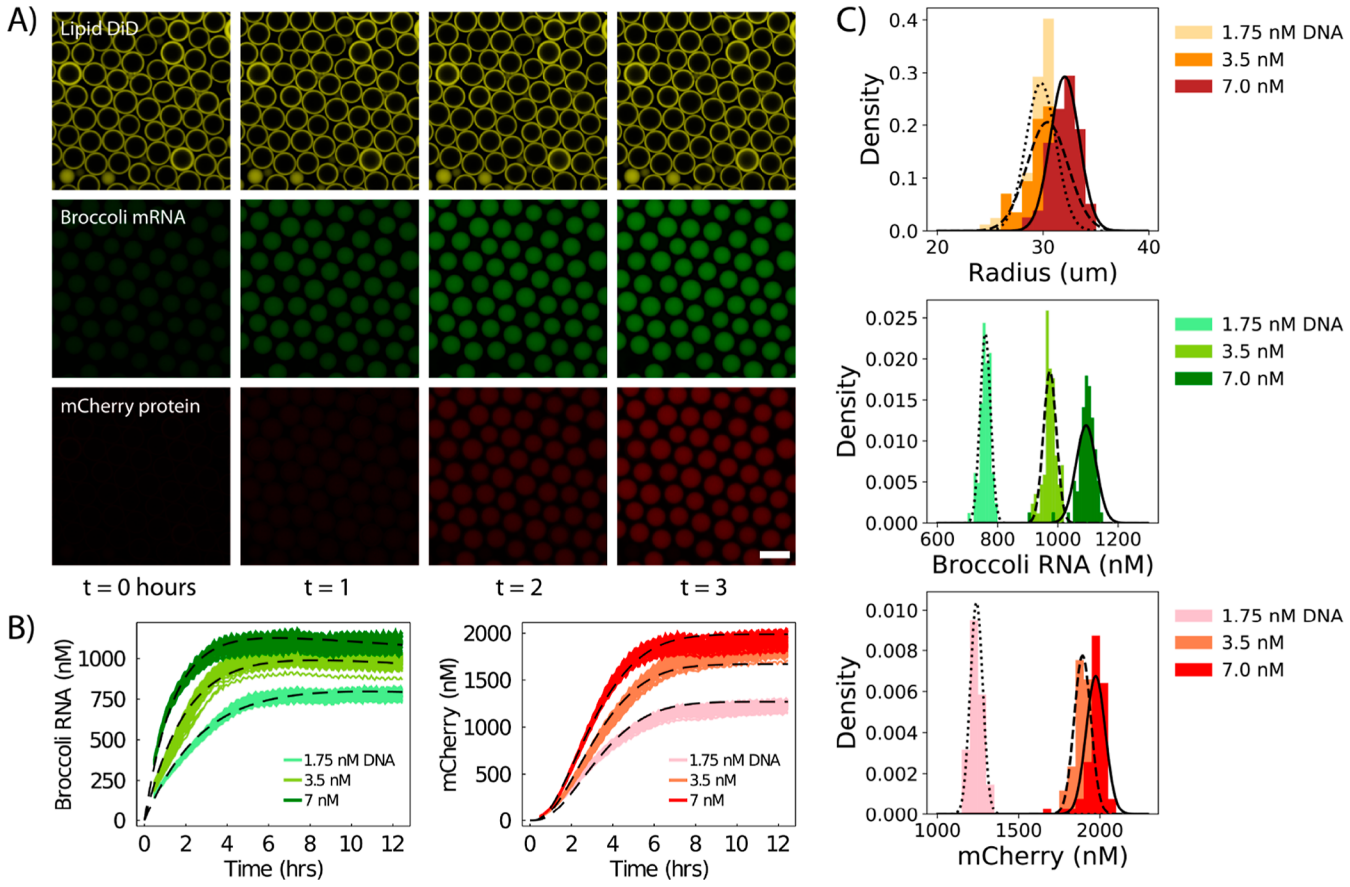


Figure 3. Quantifying transcription and translation dynamics in synthetic cell populations. (A) Timelapse confocal images of a synthetic cell population containing PURE CFES and 3.5 nM pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid DNA. Images are divided into three channels: DiD dye-tagged lipid membrane (top row), Broccoli mRNA (middle row), and mCherry protein (bottom row). Timelapse images were taken every 5 min for a total of 12 h incubation at 30 °C using confocal microscopy. Scale bars are all 100 μm. (B) Single-cell traces of mRNA and protein expression in three synthetic cell populations with 1.75, 3.5, and 7.0 nM pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid DNA with 82, 85, and 78 cells, respectively. RFU values are converted into nM concentrations units. (C) Endpoint distributions of radius, mRNA, and protein concentrations of the synthetic cell populations.

RNA and protein levels in the synthetic cell populations were then monitored with confocal microscopy at 30 °C for 12 h (Figure 3A). Similar to the bulk CFES experiments, RFU was converted into absolute concentrations using a standard calibration curve, and our image analysis protocols were used (Sections 5 and 13 in the Supporting Information) to quantify RNA and protein dynamics in the synthetic cell populations (Figure 3B). Cell sizes from the three populations containing different plasmid DNA concentrations were monodisperse at ~30 μm radius with a CV ranging from 0.04 to 0.065. The variability of gene expression from mRNA to protein remained constant with CV values ranging from 0.02 to 0.03 (Figure 3C and Table 2). This indicates a low degree of variability in translation across the synthetic cells, as CV values were not altered between mRNA and protein levels. Based on cell size and concentration measurements, copy numbers of DNA, mRNA, and protein molecules in a single synthetic cell are estimated to be in the order of 10^5, 10^7, and 10^8, respectively. Other components required for gene expression in the PURExpress CFES are also present in similar or higher concentrations,[10,54] such that stochastic effects associated with low copy numbers should be virtually absent. Time scales of active gene expression were comparable between bulk (Figure 1B) and encapsulated reactions (Figure 3B) (approx. 8 h). Maximum gene expression rates and endpoint mRNA and protein concentrations differ between the bulk expression and compartmentalized expression (Figures S14–S15 and S45–S46 in the Supporting Information). For example, protein expression levels in the liposomes are consistently lower than in bulk reactions. To quantify gene expression dynamics, mean RNA and protein dynamics from all three synthetic cell populations were globally fit to the resource-limited CFES model in eqs 1–6. Sample preparation of the synthetic cells typically took 0.5–1 h due to the encapsulation of different plasmid concentrations. As a result, the initial points of gene expression were not fully captured in the time series data. To account for sample preparation time, a time lag parameter (τ_l)

Table 2. Mean and Standard Deviations of the Microfluidic-Generated Synthetic Cell Populations[a]

Population	DNA (nM)	Radius (μm)	Broccoli RNA (nM)	mCherry (nM)
1	1.75	29.8 ± 1.4 (0.048)	759.0 ± 17.3 (0.022)	1240.5 ± 38.4 (0.031)
2	3.5	30.4 ± 1.9 (0.064)	973.8 ± 21.6 (0.022)	1892.7 ± 51.5 (0.028)
3	7.0	32.0 ± 1.4 (0.043)	1093.7 ± 33.5 (0.03)	1973.9 ± 59.1 (0.03)

[a]The total number of cells analyzed is 82, 85, and 78 for populations 1, 2, and 3, respectively.

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205–215

was included into the fitting procedure of the model for the time between starting the CFES reaction and acquisition of the first data point.

The fitted rate parameters and 95% CIs obtained by fitting the experimental data to the model are shown in Table 1. It is important to note that the parameter estimation was performed on only one batch of experiments. This was done to avoid known batch-to-batch variability in CFESs that is not accounted for in our parameter estimation methodology. However, we also observed that different batches of synthetic cell populations prepared on different days can result in different endpoint protein concentrations and maximum translation rates with a batch-wise CV of 0.10 and 0.16, respectively (Figure S48 in the Supporting Information). This is comparable to previously reported batch CV values of expressed eGFP or RFP in bulk PURE systems at 0.05−0.2.[15,55] The identifiable parameters are comparable to the bulk reaction parameters within one order of magnitude. However, more parameters were weakly identifiable as only three DNA concentrations (1.75, 3.5, and 7 nM) were considered for model fitting in the synthetic cell population experiment (Section 14 in the Supporting Information). The parameters K_l and τ_d were weakly identifiable, and b is non-identifiable. Additional populations were not prepared, as the time required to generate the populations of vesicles would have led to less data being obtained in the initial rates of gene expression, which are important for the purpose of parameter inference. We hypothesized that deviations of rate parameters between bulk and encapsulated formats are due to the different chemical conditions. Specifically, the composition of the outer feeding buffer solution can affect the inner CFES reaction by diffusion of materials across the semipermeable membrane.[6] In preparing inner and outer solutions, we ensured that inner CFESs and outer solutions were osmotically balanced by matching freezing-point osmometer measurements. However, the outer feeding buffer and inner PURExpress CFES had slightly different compositions. To determine whether the composition of the outer solution would significantly affect gene expression in the microfluidic-generated synthetic cells, synthetic cell populations with identical inner and outer solutions (except for the DNA plasmid template in the outer solution) were compared with progressively diluted outer solutions. Diluted outer solutions resulted in a lower expression of RNA and protein in the synthetic cell populations, which shows that the composition of the outer solution influences the dynamics of the encapsulated CFES reaction (Figures S53 and S54 in the Supporting Information). The higher expression in the undiluted PURExpress outer solution agrees with previous experiments, showing that an outer solution that has been chemically tuned can improve gene expression in liposome-encapsulated CFES.[6,56]

3. CONCLUSIONS

In summary, our study tested different variations of a coarse-grained model of CFES reactions using simultaneously quantified RNA and protein dynamics with likelihood-based methods for model selection and parameter identification. Using a coarse-grained model, gene expression parameters were estimated without the knowledge of the full composition of the CFES. This is particularly useful for crude extract systems or proprietary CFESs such as the NEB PURExpress CFES that we used in this study. Several models have been developed to include more details of CFE reactions such as initiation and elongation factors[26] or multiple translating ribosomes on an mRNA template.[24] These models provide a more detailed interpretation of the data but require either additional information on the time-varying states of these gene expression factors or additional unknown parameters that can result in overparameterized models and non-identifiability. In the present study, a coarse-grained model of transcription and translation was able to recapitulate the full gene expression dynamics across DNA and RNA titration experiments. While we focused on a simple constitutively expressed gene in this model, it can be readily extended to more complex gene circuitry, CFES characteristics, and protein maturation properties. We then showed that large populations of highly monodisperse synthetic cells can be reproducibly generated using double-emulsion microfluidics. Gene expression in these synthetic cells is uniform and deterministic. Using our methodologies, we demonstrated that bulk and encapsulated CFES reactions result in different gene expression dynamics. These differences are attributed to the semipermeable lipid membrane, which allows the exchange of ions and water that alters the internal composition of the synthetic cells. This emphasizes the importance of the physical environment to compartmentalized biochemical reactions. Our results demonstrate a high degree of control over synthetic cell production and relative ease of analysis compared to synthetic cells with high variability generated by bulk encapsulation methods which will be critical for bottom-up synthetic biology to build synthetic multicellular systems.

4. MATERIALS AND METHODS

4.1. Plasmid Design. The plasmids pEXP5-NT/6xHis eGFP[57] and pEXP5-NT/6xHis mCherry[58] were kindly provided by J. L. Ross Anderson, University of Bristol. These plasmids consist of a constitutive T7 RNA polymerase-mediated promoter with a strong ribosomal binding site to express 6xHis-tagged eGFP and mCherry fluorescent proteins, respectively. The pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid was made by inserting the F30-2xdBroccoli fragment downstream the mCherry stop codon and upstream the terminator of the mCherry gene. This results in transcribed mRNA that includes the F30-2xBroccoli sequence but a translated protein without the F30-2xBroccoli sequence. The F30 structure acts as a stable RNA scaffold for the two dimeric Broccoli units (2xdBroccoli).[59,60] Broccoli binds and activates the fluorescence of the small molecule (*Z*)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1*H*-imidazol-5(4*H*)-one (DFHBI) (Sigma, USA). Plasmid construction protocols and sequences are further described in Section 2 of the Supporting Information. The pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid was sequenced, confirmed by Sanger sequencing, and is available in Addgene (www.addgene.org, plasmid ID 169233).

4.2. Bulk CFES Experiments. Bulk CFES expression experiments were run using a standard half-volume (12.5 μL) reaction mix of the PURExpress *in vitro* protein synthesis kit (NEB, USA). All CFES experiments were supplemented with sucrose at a final concentration of 80.4 mM. The additional sucrose was included to balance the osmolarity between inner and outer buffer solutions in the encapsulated experiments and was also included in the bulk experiments to maintain the same reaction conditions. To detect levels of the Broccoli RNA aptamer, 10 μM DFHBI was added in CFES reactions using

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

the pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid or its purified transcripts. All plasmid DNA templates were prepared and purified by ethanol precipitation using the QIAGEN plasmid maxi kit (QIAGEN, Germany) and then dissolved in nuclease-free water. 6xHis mCherry F30-2xdBroccoli mRNA transcripts were prepared by in vitro transcription of the pEXP5-NT/6xHis mCherry F30-2xdBroccoli plasmid using the HiScribe T7 high-yield RNA synthesis kit (NEB, USA), treated with DNAse I (NEB, USA), purified using the QIAGEN RNeasy mini kit, and dissolved in nuclease-free water. Triplicate CFES reactions with the required DNA or mRNA template concentrations were prepared in 384-well plates, sealed with a clear film, and incubated in a TECAN Spark 20M plate reader at 30 °C. Fluorescence measurements were undertaken for each sample at 10 min intervals for 8 h. Excitation and emission wavelengths used were 485/535, 570/620, and 450/510 nm with a bandwidth of ±20 nm each, for eGFP, mCherry, and Broccoli RNA, respectively. Fluorescence values were then converted into concentration units using a linear calibration curve from serial dilutions of purified eGFP protein, mCherry protein, and Broccoli RNA in the same CFES reaction mix and plate reader acquisition settings. Further details for the calibration and bulk CFES experiments are available in Sections 5−7 of the Supporting Information.

4.3. Encapsulated CFES Experiments. CFES reactions were encapsulated into liposomes using a double-emulsion microfluidic device and methodology as presented in ref 37. Inner CFES solutions were prepared with a plasmid DNA template similar to the bulk CFES experiments. The lipid oil phase was composed of 1-octanol (Sigma, USA) with 6.5 mM L-α-phosphatidylcholine (Egg PC) phospholipids (Avanti, USA) and 53.3 μM 1,1′-dioctadecyl-3,3,3′,3′-tetramethylindodicarbocyanine, 4-chlorobenzenesulfonate salt (DiD) fluorescent dye (Invitrogen, USA). The outer aqueous solution was composed of a CFES feeding buffer solution modified using previous work,[8] which contains NTPs (6 mM ATP (Sigma, USA), 4 mM CTP (Sigma, USA), 4 mM UTP (Sigma, USA), 6 mM GTP (Roche, Switzerland)), amino acids (0.5 mM each) (Sigma, USA), 1.5 mM spermidine(Sigma, USA), 1.5 mM dithiothreitol (Thermo, USA), 0.02 mM folinic acid (Sigma, USA), 280 mM potassium glutamate (Sigma, USA), 20 mM magnesium glutamate (Sigma, USA), 100 mM HEPES (Roth, Germany), 480 mM glucose (Sigma, USA), and 2% (w/v) Pluronic F-68 (Gibco, USA) at pH 7.6. We can generate hundreds to thousands of synthetic cells per microfluidic session but typically prepare up to a hundred synthetic cells in a glass slide for imaging. Prepared synthetic cell populations were imaged by confocal laser scanning microscopy using an inverted Zeiss LSM 880 with Airyscan and a 10X/0.45 Plan-Apochromat M27 objective. The samples were maintained at 30 °C. Laser excitation wavelengths were 488, 488, 561, and 633 nm for Broccoli RNA, eGFP protein, mCherry protein, and DiD dye, respectively. Emission wavelengths were 499−561, 499−561, 579−641, and 640−720 nm for Broccoli RNA, eGFP protein, mCherry protein, and DiD dye detection, respectively. Images were focused at the equator of the synthetic cells and then acquired every 5 min for a total of 12 h. z-stack images of the samples were taken at the 12 h endpoint. Timelapse and z-stack images were processed using Fiji (v1.53c)[61] and Python (v3.6) with Scikit-image.[62] Synthetic cells were segmented, and fluorescence values for each cell were taken and converted to concentration units using a linear calibration curve from serial dilutions of purified eGFP protein, mCherry protein, and Broccoli RNA in bulk CFES reaction solutions with the same confocal microscopy acquisition settings and corrected for changes in laser power (Section 5 in the Supporting Information). Further details for microfluidic chip fabrication and pretreatment, bulk phase transfer and microfluidic CFES encapsulation, image analysis, and synthetic cell population experiments are available in Sections 10−16 in the Supporting Information.

4.4. CFES Model Selection and Parameter Estimation. A cell-free gene expression model was developed based on a previously published resource-limited gene expression model.[22] We tested seven variations of the model using mass action or Michaelis−Menten kinetics for transcription and translation and the degradation and consumption of transcription resources (TsR) and translation resources (TlR). These models were fit on the Broccoli RNA aptamer and mCherry protein time series data from our bulk experiments. The agreement between the experimental data and model was measured by the negative natural logarithm of the likelihood of the model parameters given the experimental data

$$-\mathrm{LL} = -\ln p(D|\theta) \tag{7}$$

where $\theta = \{\theta_1, \cdots, \theta_k\}$ is the set of parameters for the model and D is the experimental data. The term on RHS is the log-likelihood of observing data D given model parameters θ. Rate parameters of a model were estimated by minimizing the log-likelihood, that is

$$\hat{\theta} = \mathrm{argmin}_\theta[-\ln p(D|\theta)] \tag{8}$$

where $\hat{\theta}$ is the maximum likelihood estimator (MLE) of the model parameters. The different models were ranked according to the AIC

$$\mathrm{AIC} = 2k - 2\mathrm{LL} \tag{9}$$

where k is the number of parameters and LL is the log-likelihood evaluated at the MLE $\hat{\theta}$. The models with the lowest AIC values were selected and used for profile likelihood analysis. Profile likelihoods and likelihood-based CIs from the parameter estimates were calculated to assess the parameter identifiability.[47,48] The profile likelihoods of each parameter were calculated by

$$\mathrm{PL}(\theta_i) = \min_{\theta_{j\neq i}}[-\ln p(D|\theta)] \tag{10}$$

which is the minimum of the negative log-likelihood with respect to all parameters $\theta_{j\neq i}$ while holding the parameter θ_i fixed. Likelihood-based CIs of each parameter were estimated by the regions in

$$\{\theta_i|\mathrm{PL}(\theta_i) + \ln p(D|\hat{\theta}) < \chi^2(\alpha, \mathrm{df})\} \tag{11}$$

where $\chi^2(\alpha,\mathrm{df})$ is the chi-squared distribution with $\alpha = 0.95$ confidence level and df is the degree of freedom, which is the number of parameters of the model.[48] The final model used (eqs 1−6) was chosen based on the AIC. Further details and results of the model selection and parameter estimation are available in Section 9 of the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00376. Data, code, and plasmid sequences used in this study available in https://dx.doi.org/10.17617/3.77. The pEXP5-NT/6xHis

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

mCherry F30-2xdBroccoli plasmid generated in this study is available in Addgene (addgene.org) with plasmid ID 169233.

Chemicals, materials, and equipment; plasmid construction; mRNA expression and purification; protein expression and purification; protein and mRNA calibration; bulk cell-free expression experiments; mCherry maturation; resource-limited gene expression models; parameter estimation, model selection, and profile likelihoods; microfluidic chip fabrication and pretreatment; encapsulation of CFES in liposomes using microfluidics; inverse emulsion phase transfer method; image acquisition and analysis; DNA titration series in synthetic cell populations; two plasmids in a synthetic cell population; dilution of CFES outer solution in synthetic cell populations (PDF)

AUTHOR INFORMATION

Corresponding Authors

Christoph Zechner − *Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; Center for Systems Biology Dresden, 01307 Dresden, Germany; Physics of Life, Cluster of Excellence, TU Dresden, 01603 Dresden, Germany*; Email: zechner@mpi-cbg.de

T-Y. Dora Tang − *Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; Center for Systems Biology Dresden, 01307 Dresden, Germany; Physics of Life, Cluster of Excellence, TU Dresden, 01603 Dresden, Germany*; orcid.org/0000-0001-9030-6991; Email: tang@mpi-cbg.de

Authors

David T. Gonzales − *Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; Center for Systems Biology Dresden, 01307 Dresden, Germany*

Naresh Yandrapalli − *Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany*

Tom Robinson − *Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00376

Author Contributions

D.T.G., T-Y.D.T., and C.Z. designed the study. D.T.G. performed the experiments. N.Y. designed and fabricated the microfluidic devices. D.T.G. analyzed the data. D.T.G., N.Y., T.R., C.Z., and T-Y.D.T. wrote the article.

Funding

Open access funded by Max Planck Society. We acknowledge financial support from the MaxSynBio Consortium, which is jointly funded by the Federal Ministry of Education and Research (Germany) and the Max Planck Society. The MPI-CBG, the Cluster of Excellence Physics of Life of TU Dresden (Grant Number EXC-1056, T-Y.D.T).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Light Microscopy Facility (LMF) and the Scientific Computing Facility (SCF) of the Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG) for their technical support and useful discussions.

REFERENCES

(1) Joesaar, A.; Yang, S.; Bögels, B.; van der Linden, A.; Pieters, P.; Kumar, B. V. V. S. P.; Dalchau, N.; Phillips, A.; Mann, S.; de Greef, T. F. A. DNA-based communication in populations of synthetic protocells. *Nat. Nanotechnol.* **2019**, *14*, 369−378.

(2) Yang, S.; Pieters, P. A.; Joesaar, A.; Bögels, B. W. A.; Brouwers, R.; Myrgorodska, I.; Mann, S.; De Greef, T. F. A. Light-Activated Signaling in DNA-Encoded Sender-Receiver Architectures. *ACS Nano* **2020**, *14*, 15992.

(3) Booth, M. J.; Schild, V. R.; Graham, A. D.; Olof, S. N.; Bayley, H. Light-activated communication in synthetic tissues. *Sci. Adv.* **2016**, *2*, No. e1600056.

(4) Dupin, A.; Simmel, F. C. Signalling and differentiation in emulsion-based multi-compartmentalized in vitro gene circuits. *Nat. Chem.* **2019**, *11*, 32−39.

(5) Niederholtmeyer, H.; Sun, Z. Z.; Hori, Y.; Yeung, E.; Verpoorte, A.; Murray, R. M.; Maerkl, S. J. Rapid cell-free forward engineering of novel genetic ring oscillators. *Elife* **2015**, *4*, No. e09771.

(6) 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.

(7) Ishikawa, K.; Sato, K.; Shima, Y.; Urabe, I.; Yomo, T. Expression of a cascading genetic network within liposomes. *FEBS Lett.* **2004**, *576*, 387−390.

(8) Caveney, P. M.; Norred, S. E.; Chin, C. W.; Boreyko, J. B.; Razooky, B. S.; Retterer, S. T.; Collier, C. P.; Simpson, M. L. Resource Sharing Controls Gene Expression Bursting. *ACS Synth. Biol.* **2017**, *6*, 334−343.

(9) Norred, S. E.; Caveney, P. M.; Chauhan, G.; Collier, L. K.; Collier, C. P.; Abel, S. M.; Simpson, M. L. Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns. *ACS Synth. Biol.* **2018**, *7*, 1251−1258.

(10) Nishimura, K.; Tsuru, S.; Suzuki, H.; Yomo, T. Stochasticity in Gene Expression in a Cell-Sized Compartment. *ACS Synth. Biol.* **2015**, *4*, 566−576.

(11) Adamala, K. P.; Martin-Alarcon, D. A.; Guthrie-Honea, K. R.; Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem.* **2017**, *9*, 431−439.

(12) Berhanu, S.; Ueda, T.; Kuruma, Y. Artificial photosynthetic cell producing energy for protein synthesis. *Nat. Commun.* **2019**, *10*, 1325.

(13) 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*, 1583.

(14) Shimizu, Y.; Kanamori, T.; Ueda, T. Protein synthesis by pure translation systems. *Methods* **2005**, *36*, 299−304.

(15) Lavickova, B.; Maerkl, S. J. A Simple, Robust, and Low-Cost Method To Produce the PURE Cell-Free System. *ACS Synth. Biol.* **2019**, *8*, 455−462.

(16) Jewett, M. C.; Swartz, J. R. Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnol. Bioeng.* **2004**, *87*, 465−471.

(17) Garenne, D.; Beisel, C. L.; Noireaux, V. Characterization of the all-E. coli transcription-translation system myTXTL by mass spectrometry. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 1036−1048.

(18) Bujara, M.; Schümperli, M.; Pellaux, R.; Heinemann, M.; Panke, S. Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. *Nat. Chem. Biol.* **2011**, *7*, 271−277.

(19) Miguez, A. M.; McNerney, M. P.; Styczynski, M. P. Metabolic Profiling of Escherichia coli-Based Cell-Free Expression Systems for Process Optimization. *Ind. Eng. Chem. Res.* **2019**, *58*, 22472−22482.

(20) Karzbrun, E.; Shin, J.; Bar-Ziv, R. H.; Noireaux, V. Coarse-grained dynamics of protein synthesis in a cell-free system. *Phys. Rev. Lett.* **2011**, *106*, 1−4.

(21) Siegal-Gaskins, D.; Tuza, Z. A.; Kim, J.; Noireaux, V.; Murray, R. M. Gene circuit performance characterization and resource usage in a cell-free "breadboard. ". *ACS Synth. Biol.* **2014**, *3*, 416−425.

(22) 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.

(23) Gyorgy, A.; Murray, R. M. Quantifying resource competition and its effects in the TX-TL system. *2016 2016 IEEE 55th Conference on Decision and Control CDC*, 2016; pp 3363−3368.
(24) Marshall, R.; Noireaux, V. Quantitative modeling of transcription and translation of an all-E. coli cell-free system. *Sci. Rep.* **2019**, *9*, 11980.
(25) Nieß, A.; Failmezger, J.; Kuschel, M.; Siemann-Herzberg, M.; Takors, R. Experimentally Validated Model Enables Debottlenecking of in Vitro Protein Synthesis and Identifies a Control Shift under in Vivo Conditions. *ACS Synth. Biol.* **2017**, *6*, 1913−1921.
(26) 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*, 025001.
(27) Horvath, N.; Vilkhovoy, M.; Wayman, J. A.; Calhoun, K.; Swartz, J.; Varner, J. D. Toward a genome scale sequence specific dynamic model of cell-free protein synthesis in Escherichia coli. *Metab. Eng. Commu* **2020**, *10*, No. e00113.
(28) Wang, S.; Majumder, S.; Emery, N. J.; Liu, A. P. Simultaneous monitoring of transcription and translation in mammalian cell-free expression in bulk and in cell-sized droplets. *Synth. Biol.* **2018**, *3*, ysy005.
(29) 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.
(30) Blanken, D.; van Nies, P.; Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **2019**, *16*, 045002.
(31) Saito, H.; Kato, Y.; Le Berre, M.; Yamada, A.; Inoue, T.; Yosikawa, K.; Baigl, D. Time-resolved tracking of a minimum gene expression system reconstituted in giant liposomes. *ChemBioChem* **2009**, *10*, 1640−1643.
(32) van Nies, P.; Nourian, Z.; Kok, M.; van Wijk, R.; Moeskops, J.; Westerlaken, I.; Poolman, J. M.; Eelkema, R.; van Esch, J. H.; Kuruma, Y.; et al. Unbiased tracking of the progression of mRNA and protein synthesis in bulk and in liposome-confined reactions. *ChemBioChem* **2013**, *14*, 1963−1966.
(33) Van Nies, P.; Canton, A. S.; Nourian, Z.; Danelon, C. *Monitoring mRNA and Protein Levels in Bulk and in Model Vesicle-Based Artificial Cells*; Elsevier Inc., 2015.
(34) 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*, 4969.
(35) Norred, S. E.; Caveney, P. M.; Chauhan, G.; Collier, L. K.; Collier, C. P.; Abel, S. M.; Simpson, M. L. Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns. *ACS Synth. Biol.* **2018**, *7*, 1251−1258.
(36) Petit, J.; Polenz, I.; Baret, J.-C.; Herminghaus, S.; Bäumchen, O. Vesicles-on-a-chip: A universal microfluidic platform for the assembly of liposomes and polymersomes. *Eur. Phys. J. E* **2016**, *39*, 59.
(37) Yandrapalli, N.; Petit, J.; Bäumchen, O.; Robinson, T. Surfactant-free production of biomimetic giant unilamellar vesicles using PDMS-based microfluidics. *Chem. Commun.* **20212021**, *41*, 100.
(38) Deshpande, S.; Caspi, Y.; Meijering, A. E.; Dekker, C. Octanol-assisted liposome assembly on chip. *Nat. Commun.* **2016**, *7*, 10447.
(39) Villar, G.; Graham, A. D.; Bayley, H. A Tissue-Like Printed Material. *Science* **2013**, *340*, 48−52.
(40) Ugrinic, M.; Zambrano, A.; Berger, S.; Mann, S.; Tang, T.-Y. D.; Demello, A. Microfluidic formation of proteinosomes. *Chem. Commun.* **2018**, *54*, 287−290.
(41) Love, C.; Steinkühler, J.; Gonzales, D. T.; Yandrapalli, N.; Robinson, T.; Dimova, R.; Tang, T. Y. D. Reversible pH-Responsive Coacervate Formation in Lipid Vesicles Activates Dormant Enzymatic Reactions. *Angew. Chem., Int. Ed.* **2020**, *59*, 5950−5957.
(42) Vibhute, M. A.; Schaap, M. H.; Maas, R. J. M.; Nelissen, F. H. T.; Spruijt, E.; Heus, H. A.; Hansen, M. M. K.; Huck, W. T. S. Transcription and Translation in Cytomimetic Protocells Perform Most Efficiently at Distinct Macromolecular Crowding Conditions. *ACS Synth. Biol.* **2020**, *9*, 2797−2807.
(43) Deng, N.-N.; Yelleswarapu, M.; Zheng, L.; Huck, W. T. S. Microfluidic Assembly of Monodisperse Vesosomes as Artificial Cell Models. *J. Am. Chem. Soc.* **2017**, *139*, 587−590.
(44) Filonov, G. S.; Moon, J. D.; Svensen, N.; Jaffrey, S. R. Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *J. Am. Chem. Soc.* **2014**, *136*, 16299−16308.
(45) Filonov, G. S.; Jaffrey, S. R. RNA Imaging with Dimeric Broccoli in Live Bacterial and Mammalian Cells. *Curr. Protoc. Chem. Biol.* **2016**, *8*, 1−28.
(46) Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Control* **1974**, *19*, 716−723.
(47) Maiwald, T.; Hass, H.; Steiert, B.; Vanlier, J.; Engesser, R.; Raue, A.; Kipkeew, F.; Bock, H. H.; Kaschek, D.; Kreutz, C.; et al. Driving the model to its limit: Profile likelihood based model reduction. *PLoS One* **2016**, *11*, No. e0162366.
(48) Raue, A.; Kreutz, C.; Maiwald, T.; Bachmann, J.; Schilling, M.; Klingmüller, U.; Timmer, J. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **2009**, *25*, 1923−1929.
(49) Garamella, J.; Marshall, R.; Rustad, M.; Noireaux, V. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synth. Biol.* **2016**, *5*, 344−355.
(50) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **2004**, *22*, 1567−1572.
(51) Golomb, M.; Chamberlin, M. Characterization of T7 specific ribonucleic acid polymerase. IV. Resolution of the major in vitro transcripts by gel electrophoresis. *J. Biol. Chem.* **1974**, *249*, 2858−2863.
(52) Dai, X.; Zhu, M.; Warren, M.; Balakrishnan, R.; Patsalo, V.; Okano, H.; Williamson, J. R.; Fredrick, K.; Wang, Y. P.; Hwa, T. Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. *Nat. Microbiol.* **2017**, *2*, 16231.
(53) Moga, A.; Yandrapalli, N.; Dimova, R.; Robinson, T. Optimization of the inverted emulsion method for high-yield production of biomimetic giant unilamellar vesicles. *ChemBioChem* **2019**, *20*, 2674.
(54) 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.
(55) Chizzolini, F.; Forlin, M.; Yeh Martín, N.; Berloffa, G.; Cecchi, D.; Mansy, S. S. Cell-Free Translation Is More Variable than Transcription. *ACS Synth. Biol.* **2017**, *6*, 638−647.
(56) Caschera, F.; Lee, J. W.; Ho, K. K. Y.; Liu, A. P.; Jewett, M. C. Cell-free compartmentalized protein synthesis inside double emulsion templated liposomes with in vitro synthesized and assembled ribosomes. *Chem. Commun.* **2016**, *52*, 5467−5469.
(57) Li, M.; Green, D. C.; Anderson, J. L. R.; Binks, B. P.; Mann, S. In vitro gene expression and enzyme catalysis in bio-inorganic protocells. *Chem. Sci.* **2011**, *2*, 1739.
(58) Tang, T.-Y. D.; van Swaay, D.; DeMello, A.; Ross Anderson, J. L.; Mann, S. In vitro gene expression within membrane-free coacervate protocells. *Chem. Commun.* **2015**, *51*, 11429−11432.
(59) Filonov, G. S.; Moon, J. D.; Svensen, N.; Jaffrey, S. R. Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **2014**, *136*, 16299−16308.
(60) Filonov, G. S.; Kam, C. W.; Song, W.; Jaffrey, S. R. In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. *Chem. Biol.* **2015**, *22*, 649−660.
(61) Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **2012**, *9*, 676−682.

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215

(62) van der Walt, S.; Schönberger, J. L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J. D.; Yager, N.; Gouillart, E.; Yu, T. Scikit-image: Image processing in python. *PeerJ* **2014**, No. e453.

https://doi.org/10.1021/acssynbio.1c00376
ACS Synth. Biol. 2022, 11, 205−215