ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

High-Throughput Regulatory Part Prototyping and Analysis by Cell-Free Protein Synthesis and Droplet Microfluidics

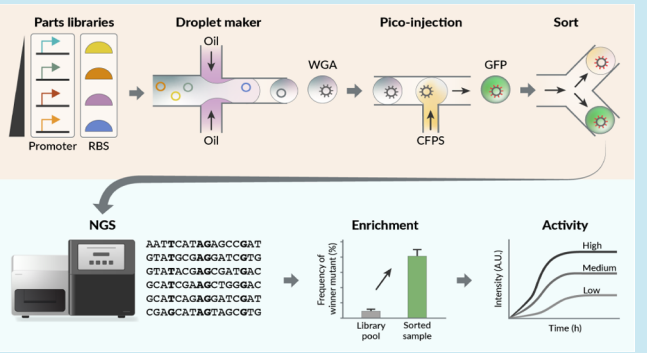
Rui GAN,◆ Maria D. Cabezas,◆ Ming Pan,◆ Huaibin Zhang, Gang Hu, Lauren G. Clark, Michael C. Jewett,* and Robert Nicol*

Cite This: https://doi.org/10.1021/acssynbio.2c00050

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Engineering regulatory parts for improved performance in genetic programs has played a pivotal role in the development of the synthetic biology cell programming toolbox. Here, we report the development of a novel high-throughput platform for regulatory part prototyping and analysis that leverages the advantages of engineered DNA libraries, cell-free protein synthesis (CFPS), high-throughput emulsion droplet microfluidics, standard flow sorting adapted to screen droplet reactions, and next-generation sequencing (NGS). With this integrated platform, we screened the activity of millions of genetic parts within hours, followed by NGS retrieval of the improved designs. This *in vitro* platform is particularly valuable for engineering regulatory parts of nonmodel organisms, where *in vivo* high-throughput screening methods are not readily available. The platform can be extended to multipart screening of complete genetic programs to optimize yield and stability.



KEYWORDS: *regulatory parts, T7 promoter, ribosome binding site, cell-free protein synthesis, microfluidics, DNA library, next-generation sequencing, emulsion droplet, flow cytometry*

INTRODUCTION

In the cell, regulatory parts (e.g., promoters, ribosome binding sites (RBSs), terminators, insulators, etc.) are responsible for the temporal and spatial control of transcription and translation. Advances in synthetic biology have enabled engineering of transcription/translation regulatory parts for tuning gene expression, metabolic pathway optimization, and biological circuit design, among others. Early studies demonstrated that carefully selecting regulatory parts could improve recombinant protein expression by increasing protein folding and preventing cell toxicity.[1,2] Over time, more complex metabolic pathways, which involve the coexpression of several genes, have been reconstituted in heterologous hosts to produce valuable therapeutics, materials, and industrial chemicals. In addition, regulatory parts can also be useful for designing genetic circuits and biosensors that can control metabolic pathways. More recent advances have made it possible to design genetic logic-based circuits that can be programmed to control the expression of large gene networks,[3−6] thus bringing the field closer to achieving rational, predictable, and programmable cellular and multicellular behavior, with important implications for biomedical sciences and biotechnology. However, significant work remains to build tools for the routine and systematic engineering of high yield heterologous pathways while maintaining cell growth and viability. One of these tools is the capability to optimize regulatory parts to balance the flux between metabolites, intermediates, and final products.

However, engineering regulatory parts continue to rely on laborious processes that typically involve: (1) generating a mutagenesis library, (2) cloning the library into a host that expresses a reporter protein (e.g., green fluorescence protein), and (3) performing a high-throughput screen, where the desired regulatory parts can be characterized quantitatively from millions of candidates. Additionally, implementing these platforms in nonmodel hosts has been challenging due to the lack of related genetic tools, difficult to culture conditions, and low transformation rates.

To address some of these challenges, we developed a high-throughput screening methodology for rapid and efficient regulatory part prototyping that combines cell-free protein synthesis (CFPS), droplet microfluidics, standard flow sorting, and next-generation sequencing (NGS). CFPS provides a

Received: January 27, 2022

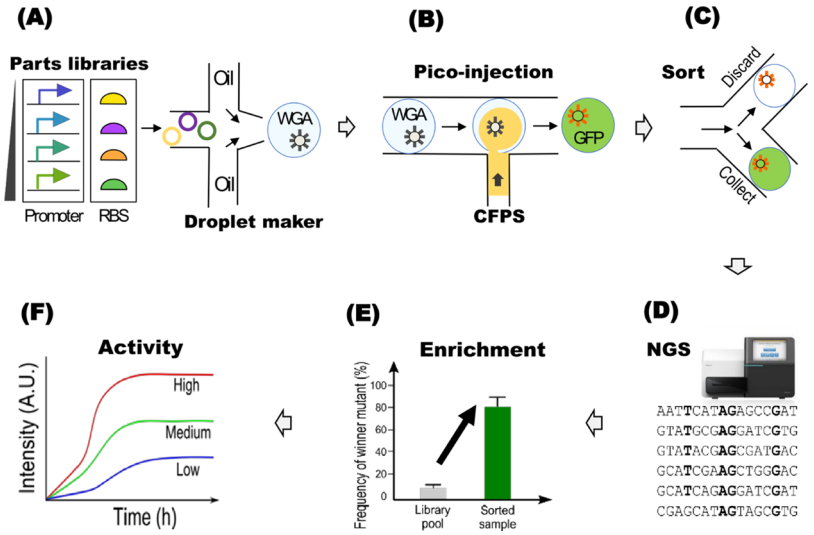


Figure 1. High-throughput screening methodology for rapid regulatory parts prototyping with cell-free protein synthesis (CFPS), microfluidics, standard flow sorting, and next-generation sequencing (NGS). (A) Compartmentalize user-designed parts library into a single droplet using a microfluidic device for whole-genome amplification (WGA). (B) Combine cell-free protein expression with WGA in droplets to express sfGFP reporter protein. (C) Flow sorting droplets based on fluorescence signal with a microfluidic device. (D) Collected parts are sequenced by next-generation sequencing (NGS), and (E) the enrichment of each part is calculated. (F) Activities of selected parts are finally characterized and used for expression regulation.

robust and modular *in vitro* platform that allows precise control over protein expression independent of cell growth. To date, several extract-based CFPS systems derived from multiple host organisms have been successfully developed,[7−17] which enabled applications in metabolic engineering, vaccine development, biomanufacturing, education, and other areas.[18−29] Although previous work has demonstrated that extract-based cell-free systems can be used for prototyping genetic parts in prokaryotes and eukaryotes,[30,31] generating and quantitatively screening a significant population of mutants in an automated, high-throughput fashion remains challenging.

The methodology we developed uses droplet-based microfluidics combined with fluorescence sorting to overcome the throughput limitation, minimize reaction volume, and reduce cost. Similar droplet-based microfluidics strategies have previously been applied for enzyme evolution,[32,33] drug discovery,[34] and genetic circuits prototyping.[35] We expand on these efforts by integrating NGS, CFPS, and microfluidic droplet sorting and demonstrate that this approach can be successfully applied for regulatory part screening. We also demonstrate the integration of standard flow cytometry instruments to screen CFPS droplet microfluidics via double emulsions. By combining these highly scalable technologies, our platform is uniquely positioned to significantly increase the throughput of regulatory parts discovery, which will enable broad applications in biotechnology.

RESULTS AND DISCUSSION

Platform Design and Overview. To achieve a robust, rapid, and cost-effective performance, our platform combines droplet-based microfluidics, fluorescence sorting, CFPS, and NGS. As shown in Figure 1A, the platform was designed to compartmentalize user-designed plasmid libraries using a microfluidic device. As shown below, we validated our platform using a randomized ribosome-binding sequence (RBS) as an illustrative regulatory part; however, the platform is flexible and can accommodate other synthetic regulatory parts. Next, each droplet, which contains a single DNA template generated by whole-genome amplification (WGA), a widely used technique for genome sequencing and isothermal amplification of DNA molecules in low abundance, is combined with CFPS reagents to achieve protein expression (Figure 1B). We used a dual-color plasmid system that expresses a reporter super folding green fluorescent protein (sfGFP) plasmid under the control of an RBS mutant with expression being proportional to the strength of the mutant, and a control protein (RFP), which allowed us to account for the background signal. By converting the droplets to a double emulsion system (Figure 1C), we were able to screen our mega-size library of mutants using flow cytometry and quantitatively determine the level of protein expression by measuring the fluorescence ratio of the coexpressed proteins. NGS then provided information about the sequence identity of RBS mutants (Figure 1D), and the enrichment of the mutants was calculated (Figure 1E). Finally, the activity of each mutant was characterized using standard plate-based methods in separate dedicated experiments (Figure 1F). The entire process was designed to yield novel regulatory parts in just one cycle that can be completed in 2 days and require minimal amounts of reagents. Below, we describe detailed validation of each step of our workflow, as well as validation of regulatory parts generated in our proof-of-concept study.

Functional Verification of WGA and CFPS Reactions Carried Out in Droplets Generated by a Microfluidics Device. A droplet maker (DM) chip was designed and fabricated to compartmentalize our library of regulatory parts into single emulsion droplets (Figure 2A). Using this chip, monodispersed water-in-oil droplets can be generated at a rate of ∼2000 to 4000 droplets per second. Droplet size can be precisely controlled at 4 pL by adjusting the flow rates of water and oil (Supplementary Figure S1). In this way, a single plasmid molecule amplified by WGA can be encapsulated into each droplet following a Poisson distribution. We first confirmed that WGA could efficiently amplify as low as 0.4 pg of plasmid (pJL1-sfGFP) to a level that could be easily

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

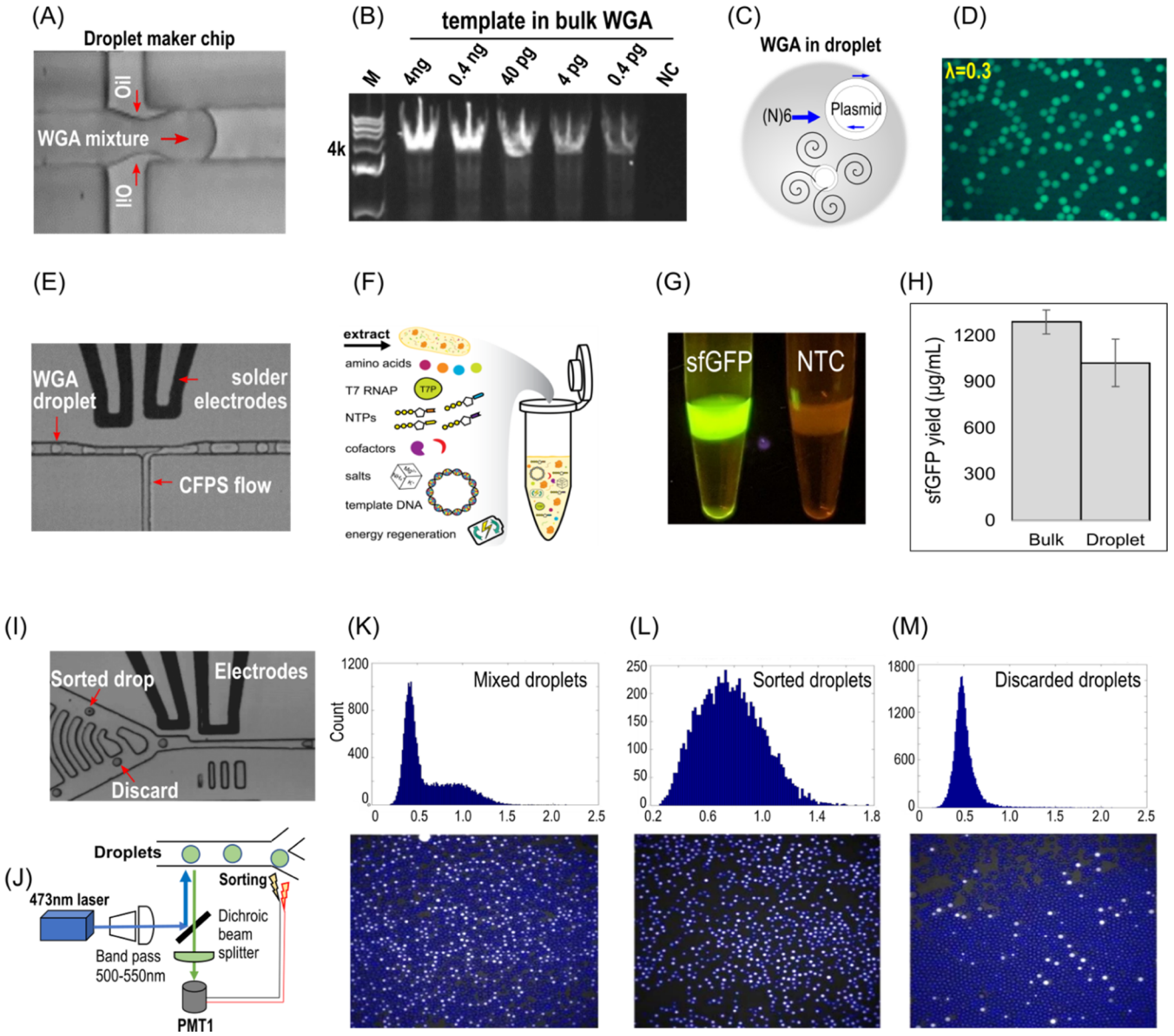


Figure 2. Microfluidic chip design and function validation. Microfluidic device fabrication for *in vitro* compartmentalization (IVC) of whole-genome amplification (WGA) and cell-free protein synthesis (CFPS) in droplets. (A) Scheme of the droplet maker (DM) chip design. In this device, the WGA reaction flow is sheared into separated droplets by an oil phase at the T-shaped junction. (B) In a bulk WGA reaction (5 μL), as low as 0.4 pg of plasmid template can be amplified and detected by agarose gel. (C) Compartmentalize a single plasmid template in emulsion droplets and perform WGA; as a result (D), the amplified DNA product can be detected in the presence of EvaGreen fluorescent dye (approximately 0.3 template molecule per droplet). (E) Design and mechanism of the pico-injection (PI) chip. The WGA droplet was loaded into the chip and flowed with spacer oil. At the T-shape cross site, each droplet meets a continuous flow of CFPS reaction mixture (F). Upon applying a potential, the WGA droplet coalesces with the CFPS mixture and the reporter proteins are expressed in the droplet. NTC: no-template control of CFPS (G). In droplets, sfGFP yield decreased by less than 30% compared to that of the bulk reaction (H). (I) Diagram of the optical pathway and (J) microfluidic sorting chip design. An assorted population of droplets with fluorescent (positive control) and without fluorescent (negative control) dye was mixed and loaded into the sorting chip. The fluorescence of each droplet was assessed to make a histogram (K) and sorted based on a preset fluorescence threshold. "Sorted" droplets (L) and "discarded" droplets (M) are collected, respectively. At last, all three samples were visualized via fluorescence microscopy to validate the sorting efficiency.

resolved using agarose gel electrophoresis from a 10 μL bulk reaction in a test tube after a 16 h incubation period at 30 °C (Figure 2B). Next, we validated that these results can be replicated in the context of our microfluidics device. We dispersed the pJL1-sfGFP plasmid into the WGA droplet with approximately 0.3 molecules per droplet (Figure 2C). After incubation (~16 h) at 30 °C, the droplets were imaged under a fluorescence microscope. Approximately one third of the droplets showed a strong green fluorescence signal, which indicated that a single plasmid could be successfully amplified by WGA (Figure 2D). Moreover, the size of the droplets remained uniform after the overnight incubation.

Our next step was to confirm that we could combine cell-free protein expression with WGA in droplets to express sfGFP reporter protein. For this, we fabricated a pico-injection (PI) chip that allows injection of the CFPS mixture into the WGA reaction droplet that encodes for the reporter protein. In our design (Figure 2E), the WGA droplets were loaded into the PI

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

chip (from the left side), while the CFPS mixture (without template) flowed continuously (from the bottom) and reached each WGA droplet at the T-shaped junction with an average 6 pL of CFPS mixture per WGA droplet (Figure 2F). These components then coalesced upon applying an electric force using a pair of electrodes (Supplementary Figure S1). After incubating at 30 °C for 16 h, the sfGFP yield synthesized in droplets was quantified and compared with a 15 μL CFPS reaction carried out in a tube. The in-droplet CFPS yield of sfGFP decreased by less than 30% (Figure 2G,H). During injection, the CFPS mixture occasionally failed to coalesce into the WGA droplets, thus forming WGA-only and CFPS-only droplets. Since either of these two types of droplets do not contain the reagents necessary to express protein, we reasoned that they would not impact our library screening because they failed to generate any fluorescence signal.

To confirm the sorting capability of our microfluidic device and optical detection system (Supplementary Figure S1) (Figure 2I,J),[36] we prepared a control sample by mixing a 1:1 ratio of droplets with fluorescence (positive drops) and without fluorescence (negative drops) (Figure 2K). A minimum threshold value (merely higher than the detectable fluorescence intensity of the negative drop) was set to perform binary sorting for the mixed droplets. The sorted positive droplets and negative droplets were collected separately, and for each sample, the positive droplets and negative droplets were counted to calculate the error rate. False positive rate was defined as negative droplets that escape into the positive channel, while false negative rate was defined as positive droplets entering the negative channel. We calculated the rate of both false positive and false negative to be under 10% (Figure 2L,M). Droplet coalescence and splitting events were not significantly detected after sorting. Taken together, we verified that our droplet-based microfluidics device can be implemented in combination with WGA and CFPS reactions and fluorescent cell sorting with high level of confidence and reproducibility.

T7 Promoter Library Construction and Screening. To benchmark the performance of our platform as a regulatory parts prototyping methodology, we used a well-validated system, the T7 promoter sequences for *Escherichia coli*.[36−38] We generated a T7 promoter library by diversifying a total of 10 nucleotides that span the transcription start site (not including the start site) (Figure 3A). The library consisted of approximately 10^6 distinct mutants as characterized by NGS (from ∼4.5 million reads) (Figure 3B). Our data revealed that the wild-type T7 promoter (wtT7P) sequence accounts for approximately 5% of the whole library. This high occurrence was probably caused by the remains of a parental plasmid that contained the wtT7P sequence even after we performed a DpnI digestion to remove the parental plasmid. Nevertheless, using our microfluidic platform, the library was encapsulated into droplets by WGA, followed by pico-injection of the CFPS mixture for sfGFP expression. The T7 promoter library was then sorted in the chip according to the fluorescence intensity (Figure 3C). During sorting, we collected the top 1% of droplets displaying strong fluorescence, while the other droplets with lower fluorescence signal and empty droplets were discarded. The DNA templates were then extracted from the sorted droplets, and sequence information was analyzed by NGS. Our data revealed that the occurrence of the wtT7P sequence (an indicator of strong T7 promoter activity) in the sorted pool increased to 83% compared to the 5% prior to

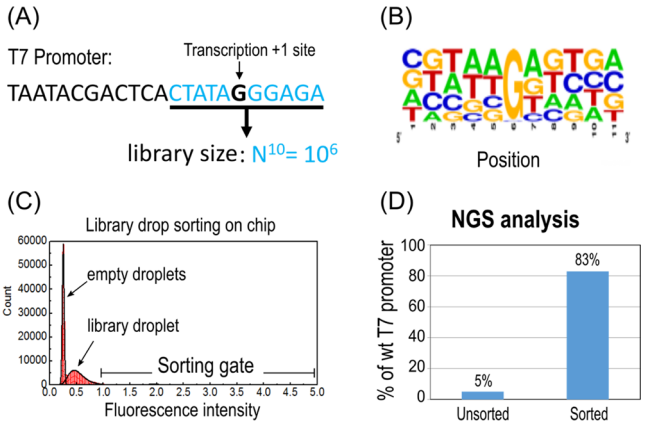


Figure 3. Screening and analysis of T7 promoter library. (A) DNA library was generated using the wild-type T7 promoter by randomizing 10 nucleotides flanking the transcription site "G" (as shown in blue). The size of this library is approximately one million (∼10^6). (B) Mutation rate at each site was confirmed by next-generation sequencing (NGS) and represented with a WebLogo 3 graphic. (C) After WGA and CFPS reactions in droplets, the droplet library was sorted in a microfluidic chip. Droplets with a fluorescence intensity in the range of the "sorting gate" were collected, while the others were discarded. (D) After NGS analysis of the library and sorted sample, the frequency of the wild-type T7 sequence (as an indicator of promoter strength) increased by 16-fold from 5 to 83%.

sorting, thus resulting in a 16-fold enrichment rate per round. (Figure 3D). Overall, these results confirm that the platform achieves enrichment of strong performing regulatory parts based on fluorescence intensity.

Improving Screening Precision and Speed with a Dual-Color CFPS and Double Emulsion. Although we successfully demonstrated the enrichment of a strong wtT7 promoter, a 16-fold enrichment still shows low efficiency compared to the top 1% sorted population. We attributed the low enrichment to increased background noise and reasoned that implementing a dual-color screening strategy should decrease background noise and improve the enrichment rate (Supplementary Figure S2). In addition, we noted that the screening speed in the microfluidic device was too slow, with a top speed of approximately 1500 droplets per second. In contrast, commercial flow cytometers can screen particles at rates of 10,000−100,000 events (cells or particles) per second. Therefore, we set out to convert the primary emulsion to a double emulsion (water-in-oil-in-water droplet) to perform screening with a standard flow cytometer. Next, we constructed a dual-color reporter plasmid, which encodes both sfGFP (GFP Ex/Em: 485/528 nm) and sfmCherry (RFP with Ex/Em: 560/620 nm). We first tested the expression of the dual-color reporter plasmid in a bulk CFPS reaction to demonstrate that a sufficient fluorescence signal was generated from the expression of both reporter proteins without interference between the two fluorescence channels (Figure 4A,B). A new microfluidic chip design allowed us to convert the primary emulsion into a double emulsion (Supplementary Figure S3). To validate the function of this chip, we encapsulated a dual-color CFPS reaction mixture into a primary emulsion and immediately converted it into a double emulsion. After a 16 h incubation at 30 °C, the double emulsion remained stable and uniform in size (Figure 4C), and expression of both sfGFP and sfmCherry was detected from the fluorescence intensity signal in the droplets (Figure 4D,E).

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

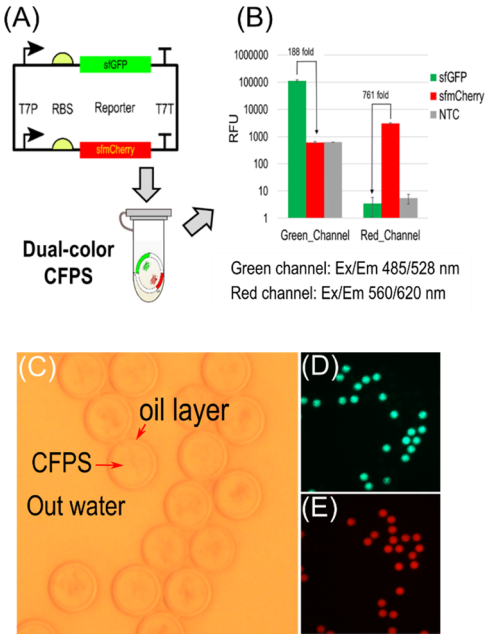


Figure 4. Cell-free protein synthesis of dual-color fluorescence reporters in double emulsion droplets for a normalized sorting. (A) Two fluorescence reporter proteins, sfGFP (green) and sfmCherry (red), were cloned into the same plasmid for coexpression in a bulk CFPS reaction, and protein expression was quantified via fluorescence. The excitation/emission wavelengths for sfGFP and sfmCherry are 485/528 and 560/620 nm, respectively. (B) When reading through the green channel, the signal of RFP was as low as the no-template control (NTC); in contrast, when reading through the red channel, the signal of GFP was as low as the no-template control (NTC). Therefore, there is no spectrum interference between sfGFP and sfmCherry when read with the optimal excitation/emission wavelengths. (C) Primary emulsion droplets containing CFPS reactions were converted into double emulsion droplets of uniform size and stability. The inert aqueous phase was separated steadily from the outer water phase by a thin layer of oil. CFPS works efficiently in the inner phase of the double emulsion, as demonstrated by sfGFP (D) and sfmCherry (E) expression retained in each droplet without detectable crosstalk or leakage.

This result confirmed that CFPS from a dual-expression plasmid took place within the double emulsion and without leakage or cross contamination.

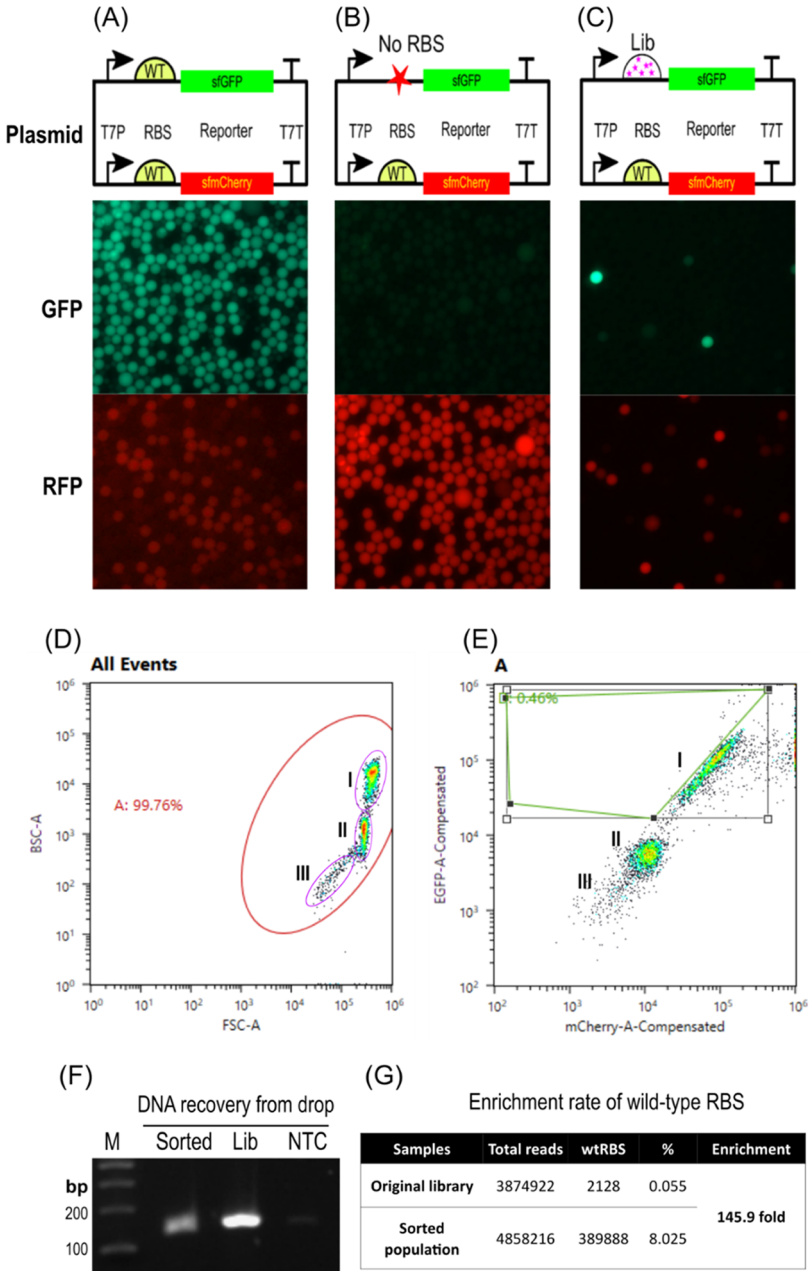
Prototyping an RBS Library Generated in Double Emulsion Droplets Using Flow Cytometry. Using this dual-color library sorting approach, we constructed an RBS library in the pJL1-NBGCheFP plasmid (Supplementary Table 2) (Supplementary Figure S4A). To generate a dual-color RBS library, 10 nucleotides in the RBS region upstream of the start site of sfGFP were fully randomized to A, C, G, or T (library size $\sim 10^4$), while the RBS of the sfmCherry plasmid corresponded to the WT sequence (Figure 5A). Mutagenesis was confirmed by NGS, with the overall distribution of the four nucleotides close to the expected 25%, ranging from 17 to 30% (Supplementary Figure S4B). We note that at each mutation site, all four nucleotides were present with a slightly biased distribution caused by the DNA synthesis.

The library was dispersed into the droplets, with ~0.3 molecules per droplet (λ = 0.3) for the WGA reaction. As a positive control plasmid, we used sfGFP and sfmCherry with the wild-type RBS (wtRBS), which exhibits a strong expression of both reporter proteins (Figure 5A). For the negative control plasmid, we removed the RBS of sfGFP while keeping the wtRBS for sfmCherry; thus, the negative control plasmid only expresses sfmCherry (Figure 5B). Next, the control plasmids were dispersed into droplets (with thousands of copies per droplet) for WGA and subsequent CFPS via pico-injection. In the droplets containing the positive control plasmid, strong fluorescence was detected from both sfGFP and sfmCherry (Figure 5A), while the fluorescence of sfmCherry was only detected in the droplets containing the negative control plasmid (Figure 5B). For the droplets loaded with constructs from the RBS library, we observed that a portion of the droplets showed both green and red fluorescence as expected for expression of dual-color plasmid that contain an active RBS mutant (Figure 5C). We noticed that a portion of the droplet population loaded with the dual-color plasmid RBS library failed to display red fluorescence, indicating that the plasmid template may be missing or that the CFPS reaction failed.

The CFPS droplets were converted into double emulsion for sorting using a flow cytometer. Following sorting, we identified three distinct droplet populations (Figure 5D). Population I exhibited a strong gross side fluorescence, which indicates that the WGA and CFPS reactions expressing sfmCherry and sfGFP were successful. Although population II showed a similar forward scatter pattern compared to population I, it was a magnitude lower. We hypothesize that these droplets contained no library template (empty droplets due to Poisson distribution) and therefore could not express any fluorescence protein. The droplets in population III were less uniform, smaller in size, and showed only a basal level fluorescence compared to population I and II. We reasoned that this population contained failed droplets that include WGA-only droplets or CFPS-only droplets, which were caused by a failure in pico-injection, or ruptured CFPS droplets or oil-in-water emulsion droplets, which were generated during double emulsion conversion. From this analysis, we identified population I as having our desired library sample and set a sorting gate to collect the top 0.46% droplets with a high ratio of green-to-red signal from a total ~3 million droplets (Figure 5E). With one round of screening toward highly active parts (high ratio of green: red), we have a higher probability of recovering a distribution from a randomized library consisting of strong/moderate parts while maintaining a population of weak parts for assessment. A DNA fragment spanning the RBS region was amplified by PCR from the DNA extracted from collected droplets (Figure 5F), and the mutation type was analyzed by NGS. We found that the wtRBS (an indicator of strong-performance mutant) was enriched by 146-fold (Figure 5G), a significant improvement of efficiency compared to a single-color setup.

Characterization of Beneficial RBS Mutants. After NGS analysis of the sorted RBS population and original library, the enrichment of sorted RBS mutants was calculated, and 48 mutants with an enrichment rate that ranges from 100- to 2000-fold were randomly chosen (Supplementary Table S3). The 48 RBS mutants were then cloned into the pJL1-sfGFP plasmid, and the activity of each mutant was measured by quantifying the sfGFP yield. Extract-based CFPS was performed at 30 °C for 16 h, and the end-point fluorescence was recorded. As shown in Figure 6A, after one round of activity enrichment, eight out of the 48 mutants exhibited activities similar to or slightly higher than the wtRBS (−90 to 116%). The remaining 40 mutants in the set contributed to a

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX



Samples	Total reads	wtRBS	%	Enrichment
Original library	3874922	2128	0.055	
Sorted population	4858216	389888	8.025	145.9 fold

Figure 5. RBS library construction and screening. (A) Positive control plasmid with the wtRBS expresses both sfGFP and sfmCherry in droplets. (B) Negative control plasmid, which lacks the RBS for sfGFP expression, only expresses sfmCherry. (C) Library plasmid was constructed by randomizing the RBS sequence for sfGFP. Protein expression was visualized via fluorescence. (D) Scatter plot of the double emulsion droplets was analyzed by flow cytometry. Droplets in population I correspond to intact droplets that contain the RBS library, while population II and III either failed to form a double emulsion or ruptured. (E) Droplets that showed a high sfGFP to sfmCherry ratio were sorted out (in green gate) as putative beneficial mutants. The DNA fragment covering the RBS region was amplified from the sorted population ("Sorted") and original RBS library ("Lib"). (F) NTC denotes no-template control of PCR. (G) Frequency of wild-type RBS (wtRBS), as a representation of highly active mutants, was identified by NGS, and the enrichment rate was calculated accordingly.

uniform distribution in activity, thus fulfilling our goal to prototype a set of regulatory parts for translation.

The kinetics of combined transcription and translation of nine strong mutants were characterized using the PURExpress system. To determine the optimal template concentration, we first generated kinetic traces of the pJL1-sfGFP plasmid having the wtRBS sequence in a 5 μL PURExpress reaction at a range of concentrations (2.5, 1.25, 0.625, 0.313, and 0.156 nM) (Supplementary Figure S5) and determined that optimal sfGFP production was obtained using 0.313 nM of DNA template. Using this concentration, we tested nine RBS mutants with activity close to the wtRBS in PURExpress reaction (Figure 6B). While the wtRBS outperformed all of the evolved mutants in the initial translation rate (within the first hour), several mutants (Rm838, Rm858, and Rm835) reached wtRBS end point after 3 h.

We further investigated the sequence similarity between the nine beneficial mutants and the wtRBS. Interestingly, these beneficial mutants and wtRBS shared a relatively low sequence similarity (Figure 6C). For instance, mutants Rm708, Rm861,

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

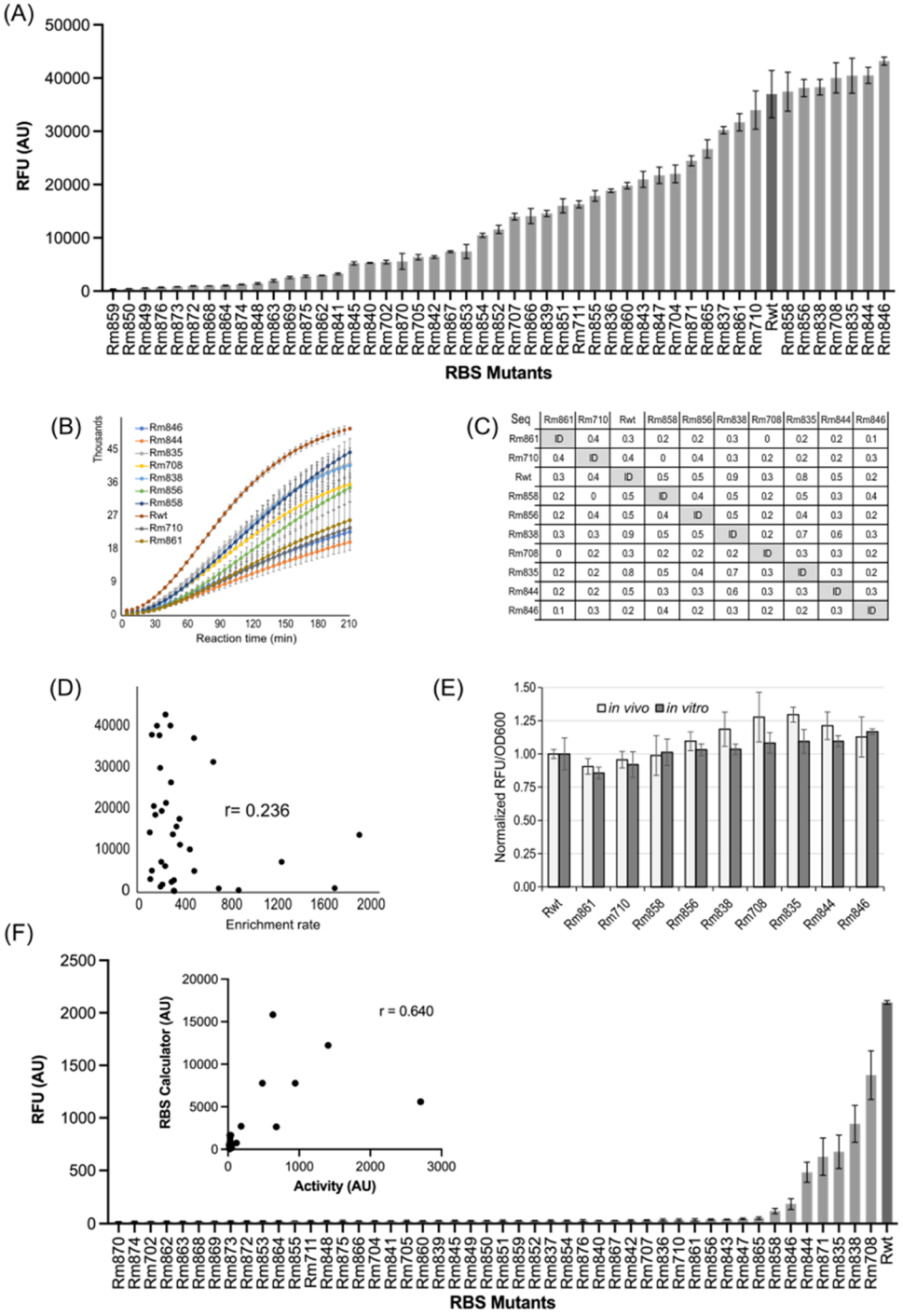


Figure 6. Kinetics characterization and sequence analysis of selected RBS mutants. (A) Forty-eight RBS mutants were sampled randomly from the screened set after one round. The activity was characterized quantitatively by measuring the end-point fluorescence yield of sfGFP in extract-based *E. coli* CFPS reaction. The activity of the wtRBS was used as a reference (solid black bar). The error bar represents a deviation of three independent reactions from each mutant. (B) Nine RBS mutants that showed high end-point protein yield was further characterized for their coupled transcription/translation kinetics using a reconstituted cell-free protein synthesis system (PURExpress) for 210 min. The error bar represents the standard deviation of two independent reactions from each mutant. (C) Sequence similarity analysis of nine RBS mutants show maximum end-point sfGFP yield. The analysis was performed using ClustalW software. "ID" denotes identical. (D) Correlation coefficient between RBS activity and enrichment. (E) Nine RBS mutants that showed high end-point protein yield were further characterized in *E. coli* cells. The ratio of RFU (sfGFP) to OD600 that indicates the activity of RBS mutant was measured at the late exponential phase (after 4 h induction). The ratio values of all mutants were normalized to the wild-type RBS (Rwt), respectively, to perform a comparison between *in vitro* and *in vivo*. The standard deviation was estimated from three individual colonies of each sample. (F) RBS mutant library screening in *Vibrio natriegens* CFPS at 5 μL scale using the Echo acoustic liquid handling system. The error bar represents a deviation of four replicate reactions.

and Rm846 exhibited a similarity to wtRBS of 30, 30, and 20%, respectively. Yet, they were capable of expressing sfGFP with the same-level end point reading in extract-based CFPS. Moreover, the sequence similarity between these mutants is

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

also low, with 0% sequence similarity between Rm708 and Rm861, 20% between Rm708 and Rm846, and 10% between Rm861 and Rm846. This sequence diversity implies that pairing between the RBS and the ribosome may not be strictly determined by the canonical Shine−Dalgarno sequence but rather by other interactions outside the RBS such as *trans*/*cis* factors in *E. coli*, as well as RBS resistance against nuclease degradation.

We also investigated the correlation between RBS mutant activity and enrichment rate (Supplementary Table S3). If sequence enrichment rate could work as a reliable indicator of RBS activity, it would dramatically accelerate directed evolution for various applications. Unfortunately, the correlation coefficient in our study is low (R = 0.236) (Figure 6D). This may occur because numerous factors contributed to a heavy bias amid the multistep screening protocol, as discussed in the previous section (Supplementary Figure S2). Moreover, NGS library preparation can lead to biased read counts. Additionally, other minor factors that should be considered are mutant stability during the CFPS reaction, as well as droplet split/coalescence during double emulsion conversion and screening. These parameters will be further examined and optimized in our follow-up studies.

Finally, we characterized nine evolved RBS mutants that showed high activity *in vitro* in living *E. coli* cells (Rm846, Rm844, Rm835, Rm708, Rm838, Rm856, Rm858, Rm710, and Rm861) as well as the wild-type (Rwt) as a control. *E. coli* cells harboring the RBS construct were grown for a total of 8 h, and the expression of sfGFP was induced from the third hour. The ratio of fluorescence of sfGFP to OD600 was measured per hour (Supplementary Figure S6). The ratio value after a 4 h induction (late exponential phase) was utilized as an indicator to the activity of RBS mutant. To make the activity values comparable between *in vitro* and *in vivo*, we normalized the values of all of the mutants to Rwt, which was counted as 1. As shown in Figure 6E, all of the nine mutants exhibited strong activity in *E. coli* cells and the trends of these mutants are quite similar between *in vitro* and *in vivo*. This verified that our *in vitro*-evolved active RBS mutants can also work well in living *E. coli* cells.

Apply Evolved *E. coli* RBS Parts to *V. natriegens* CFPS. Using the results from our RBS library, we attempted to extend the applicability to the nonmodel organism *V. natriegens* CFPS because only a handful of genetic parts for *V. natriegens* have been identified.[39] While the results reveal that the wtRBS outperforms all of the RBS mutants, five mutants showed significant activities ranging from 25 to 65% of the observed wtRBS activity (Figure 6E), likely due to the conserved nature of 3′ end region of the 16 sRNA between *E. coli* and *V. natriegens*. We further compared our evolution strategy to the performance of these mutants using the Salis RBS calculator, a computational tool that allowed us to predict translation initiation rates (TIR).[40] While this comparison did not yield a strong correlation coefficient (R = 0.650), we did observe that the top-five performing mutants also corresponded to those with high TIRs (Figure 6E inset). We hypothesize that the factors related to mRNA stability and degradation of the 5′UTR inherent in *V. natriegens* contributed to deviations from theoretical calculations, and further experiments will provide more insight. Overall, combining high-throughput experimentation with a guided computational approach can expand the capability of prototyping regulatory parts for nonmodel organisms.

CONCLUSIONS

In this study, we established a quantitative high-throughput screening platform for regulatory part prototyping by leveraging the advantages of microfluidics, NGS, and CFPS. Using this platform, we successfully prototyped a large randomized RBS mutant library (10^6 mutants) in *E. coli* CFPS. We also demonstrate a general double emulsion-based method for CFPS reaction droplet sorting based on fluorescence intensity and standard flow cytometry quantification according to the performance of each RBS mutant. Importantly, NGS enabled sequence analysis of each RBS mutant in the library. Based on the NGS data, we further characterized the 48 beneficial RBS mutants that were significantly enriched. We identified 8−9 mutants that exhibited equal activities to the wtRBS after only one round of screening despite low similarity to the wtRBS in sequence. This demonstrates the ability of this platform to prototype unique regulatory parts from nonmodel strains, such as *V. natriegens*. Overall, although further optimization will be needed to improve the performance of the platform, current proof-of-concept results highlight the utility of integrating droplet-based microfluidics, NGS, and CFPS, especially when complemented with flow cytometry and computational methods to facilitate sorting and design, respectively. We expect that the high-throughput regulatory part prototyping we describe here will pave the way toward the construction and analysis of sequence-function maps for regulatory parts, including for nonmodel organisms, as well as facilitate future *in vivo* cell engineering, biosensing, and biomanufacturing efforts.

MATERIAL AND METHODS

General Reagents for Molecular Biology Research. All DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc (IDT) (Supplementary Table S1). Phi29 DNA polymerase from *Bacillus subtilis* (M0269L), Inorganic pyrophosphatase (Yeast) (M2403S), NdeI (R0111S), SalI-HF (R3138S), T4 DNA Ligase (M0202S), Gibson assembly Master Mix (E2611S), NEBNext Ultra II End Repair/dA-Tailing Module (E7546S), NEBNext Ultra II Ligation Module (E7595S), 10-β competent *E. coli* cells (C3019H), and PURExpress *In Vitro* Protein Synthesis Kit (E6800S) were purchased from New England Biolabs (NEB, Ipswich, MA). E-Gel EX Agarose Gels 1% (G402021), 2% (G402002), and 4% (G401004) were purchased from Thermo Fisher Scientific (Waltham, MA). 1*H*,1*H*,2*H*,2*H*-Perfluoro-1-octanol (PFO) (370533−5G), Trichloro(1*H*,1*H*,2*H*,2*H*-perfluorooctyl)silane (448931−10G), and Poly(sodium 4-styrenesulfonate) (243051) were purchased from Sigma-Aldrich (St. Louis, MO). QIAquick PCR Purification Kit (28104) and QIAprep Spin Miniprep Kit (27104) were purchased from Qiagen (Germantown, MD). SsoAdvanced Universal SYBR Green Supermix (172−5270) and CFX96 Touch Real-Time PCR Detection System (1855196) were purchased from Bio-Rad (Billerica, MA). Miseq kit v3 (150 cycle) and Miseq DNA sequencer for next-generation sequencing (NGS) were obtained from Illumina (Illumina, San Diego, CA). Novec HFE-7500 was purchased from 3M (Saint Paul, MN).

Microfluidic Device Fabrication. All microfluidic devices were fabricated following a standard lithographic process. Briefly, the mold was fabricated by patterning SU-8 on a Si wafer using photolithography. A poly(dimethyl siloxane)

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(PDMS) replica was then fabricated from the mold by soft lithography.[32,41] The cured PDMS replica was peeled off from the mold and punctured with 1 mm diameter inlet and outlet holes. The PDMS replica was then bonded to a glass slide after being treated with oxygen plasma for 45 s.

For the fabrication of the pico-injector and sorter, the microelectrode was made by filling the electrode channel with solder (51In/32.5Bi/16.5Sn, Indium Corporation) at 85 °C to make electric connections with the metal pins that connect to a power supplier (Figure 2E). After cooling to room temperature, conductivity between the metal pin and the solidified solder electrode was tested using a multimeter.

The single emulsion droplet generator, pico-injector, and sorter were treated with 1.5% (v/v) fluorosilane (1*H*,1*H*,2*H*,2*H*-perfluorooctyl trichlorosilane) in HFE-7500 to render the channel surface fluorophilic. Residual silane solution was removed by flushing with neat HFE-7500 and dried by blowing air. The double emulsion droplet generator was first treated with PDADMAC (Poly-(diallyldimethylammonium chloride), 2 mg/mL in 0.5 M NaCl), followed by PSS (Poly(sodium 4-styrenesulfonate), 2 mg/mL in 0.5 M NaCl) to render the channel surface hydrophilic. Residual polyelectrolyte solution was removed by flushing milli-Q water and dried by blowing nitrogen gas.

Droplet Generation and Manipulation. Water-in-fluorocarbon (w/o) single emulsions were generated from a single emulsion droplet generator, which was a two-inlets flow-focusing device with a channel height of 20 μm and a nozzle size of 20 μm. The continuous phase contained fluorosurfactant (2% PFPE-PEG surfactant in HFE-7500) with a flow rate of 0.45 mL/h, and the disperse phase contained a library of template constructs with a flow rate of 0.15 mL/h. The average droplet volume was ∼4 pL. The droplets were collected from the outlet tubing in a PCR tube.

Water-in-fluorocarbon-in-water (w/o/w) double emulsion droplets were generated from a double emulsion droplet generator (Supplementary Figure S3), which was a three-inlets flow-focusing device with a channel height of 20 μm and a nozzle size of 20 μm. The previously generated water-in-fluorocarbon (w/o) single emulsion was used directly as the dispersed phase, which was injected into inlet 1 at a flow rate of 0.04 mL/h. An additional 2% fluorosurfactant in HFE-7500 was injected from inlet 2, with a flow rate of 0.06 mL/h, which served as the "spacer oil" to separate individual aqueous droplets, thus preventing multiple encapsulations. The continuous phase, which contained 0.2% Triton 100-X in 1X PBS, was injected from inlet 3, with a flow rate of 0.2 mL/h. The droplets were collected from the outlet tubing into a PCR tube for further experiments.

Pico-injection was performed by flowing droplets at 0.2 mL/h from inlet 1. Two percent surfactant was injected at 0.4 mL/h to separate the droplets apart. CFPS was injected at 0.3 mL/h. Pico-injection of CFPS into droplets occurred when the electric field was turned on. The estimated pico-injection volume was 6 pL per WGA droplet (Figure 2E). The pico-injected droplets were then collected from the outlet tubing in a PCR tube. On-chip sorting was performed according to previous publications with some modifications.[32,41]

Whole-Genome Amplification (WGA) and Single-Molecule Encapsulation in Droplets. WGA reaction was prepared with Phi29 DNA polymerase (NEB) with minor modifications. Specifically, the WGA reaction mix contained 1 pg/μL plasmid, 1x phi29 reaction buffer, 2 μM exo-nuclease-resistant hexanucleotide primers (5′-NNNN*N*N-3′), 0.2 mM of each deoxynucleoside triphosphate (dNTP), 0.8 U/μL phi29 DNA polymerase, and 2 U/mL pyrophosphatase. The reaction components were mixed in 1.5 mL tubes by adding DNA template, nuclease-free water, and hexamers and then heated to 95 °C for 3 min to allow template denaturing. Next, the mixture was quickly transferred to ice for primer annealing, after which the remaining components were added, and the reaction was incubated for 16−18 h at 30 °C. For single-molecule primed WGA in the droplet, the reaction was prepared following the same protocol as the regular reaction except that the concentration of template plasmid was adjusted to ≤1 molecule per droplet. The whole reaction was then encapsulated into droplets using the microfluidic device. Fluids were loaded into the microfluidics device at 100 μL/h for the aqueous phase and 400 μL/h for the carrier oil. The droplets were approximately 3−4 pL in volume and were collected off-chip in the form of an emulsion, which was incubated for 16−18 h at 30 °C. To confirm the final product, WGA droplets were ruptured with 1*H*,1*H*,2*H*,2*H*-Perfluoro-1-octanol (PFO) and the aqueous phase containing DNA product was separated from the organic phase by a centrifugation (16, 000*g* for 10 min). The collected product was digested by restriction endonucleases (NdeI or SalI) and analyzed with a 1% agarose gel.

***E. coli* Cell-Free Protein Synthesis and Pico-Injection into WGA Droplets with Microfluidic Device.** *E. coli* extract-based CFPS was performed according to a previously published report with minor modifications.[42,43] *E. coli* strain C321.ΔA.759 (*endA*$^-$ *gor*$^-$ *rne*$^-$ *mazF*$^-$) was used for preparing the whole-protein extract used in this study.[44] The CFPS reaction was prepared in a 1.5 mL tube and then centrifuged (16, 000*g* for 10 min) to remove insoluble debris that could jam the microfluidic channel. During pico-injection, the WGA droplet was reloaded into the chip and flowed through the channel with an interval distance of approximately 100 μm. CFPS mixture was then pumped into the chip as a continuous flow. The flow rate of WGA droplets was 0.2 mL/h versus 0.3 mL/h for CFPS, resulting in a 1:1.5 volume ratio of WGA to CFPS (Supplementary Figure S1). CFPS droplets were then collected off-chip and incubated for ∼16 h at 30 °C.

Plasmid Construction. Sequences of key plasmids used in this study are listed in Supplementary Table S2. Superfolder GFP (sfGFP) reporter plasmid pJL1-sfGFP was used for numerous previous studies.[42−44] An *E. coli* codon-optimized superfolder mCherry (sfmCherry) sequence was obtained from IDT, and cloned into the pJL1-sfGFP backbone to replace the sfGFP-coding sequence between the NdeI and SalI restriction sites. The resulting sfmCherry reporter plasmid was named pJL1-sfmCherry.

The dual-color reporter plasmid containing both sfGFP and sfmCherry was constructed by inserting the entire sfmCherry expression cassette into the backbone of pJL1-sfGFP. First, a fragment containing the T7 promoter, RBS, sfmCherry-coding sequence and T7 terminator was amplified from the pJL1-sfCherry plasmid by primers G524-f and G524-r. The pJL1-sfGFP plasmid was amplified with primers G525-f and G525-r as a linear vector. The expression cassette of sfmCherry was then ligated into the vector following a standard Gibson assembly protocol (NEB). The resulting dual-color reporter plasmid was named pJL1-GCheFP. The pJL1-NBGChFP plasmid was constructed by deleting the ribosome-binding sequence (RBS) of sfGFP from pJL1-GCheFP.

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Construction of the T7 Promoter Library and RBS Library. The pJL1-sfGFP plasmid containing a T7 promoter expression cassette was amplified by PCR using G484-f and G484-r primers (Supplementary Table S1). During PCR, the primer G484-r introduced a total of 10 randomized nucleotides (N) flanking the transcription initiation site "G" (Figure 3A). The PCR products were purified by QIAquick PCR Purification Kit (Qiagen), digested by DpnI (NEB) to remove parental plasmid, and followed by self-ligation using T4 DNA ligase (NEB). The circular DNA constructs consisted of a library of T7 promoter variants that express sfGFP. Randomization of the 10 nucleotides was confirmed by NGS (150 cycles) covering the entire mutation site. Prior to encapsulating a single molecule into droplets, the library was quantified by a real-time PCR with the primers G499-f and G499-r and SsoAdvanced Universal SYBR Green Supermix (Bio-Rad).

The construction of the RBS library is shown in Supplementary Figure S4. DNA oligos G589-r carried 10 randomized nucleotides that correspond to the position of the wtRBS. G589-f and G589-r were annealed and extended via PCR to create a duplex library fragment for insertion. Meanwhile, a linear vector was prepared by amplifying the pJL1-NGCheFP plasmid with the primers G590-f and G590-r. The insert and vector were ligated following the standard Gibson assembly protocol from NEB (E2611S). The randomization of the 10 nucleotides was confirmed by NGS. Prior to single molecule encapsulation, the library was quantified by real-time PCR with the primers G499-f and G499-r and SsoAdvanced Universal SYBR Green Supermix (Bio-Rad).

T7 promoter Library Screening with Fluorescence-Activated Droplet Sorting (FADS). The optical setup consisted of an inverted microscope (Carl Zeiss SAS) mounted on a vibration-dampening platform (Thorlabs GmbH). A 20 mW, 488 nm solid-state laser (LAS; Newport-Spectra-Physics) was mounted on the platform via a heat sink (Newport-Spectra-Physics). The laser beam was shaped into a 10 × 150 μm line by combining a 25 mm diameter cylindrical lens (effective focal length: −50 mm; Thorlabs GmbH) and a 25 mm diameter plano-convex lens (effective focal length: 25 mm; Thorlabs GmbH) having a 5 cm distance between them (LL). The shaped beam was guided to the side camera port of the microscope via a series of periscope assemblies (Thorlabs GmbH). Inside the microscope, the laser light was reflected up into an LD Plan Neofluar 40/0.6 microscope objective (OBJ; Carl Zeiss SAS) and focused across a channel within the microfluidic device (CHIP). A high-speed digital camera (CAM; Vision Research) was mounted on the top camera port of the microscope to capture digital images during droplet production and sorting. A 488 nm notch filter (F1; Semrock Inc.) positioned in front of the camera protected the camera's sensor from the reflected laser light. Light emitted from fluorescing droplets was captured by the objective and channeled back along the path of the laser into the system of periscope assemblies. The emitted light was separated from the laser beam by a 488/532 nm-wavelength transmitting dichroic beam splitter (DBS; Semrock Inc.), passed through a 510 nm bandpass filter (F2; 20 nm bandwidth; Semrock Inc.), and collected in a H5784-20 photomultiplier tube (PMT; Hamamatsu Photonics K.K.). Data acquisition (DAQ) and control were performed by a PCI-7831R Multifunction Intelligent DAQ card (National Instruments Corporation) executing a program written in LabView 8.2 (National Instruments Corporation). The data acquisition rate for the system was 100 kHz. To sort a particular droplet, the DAQ card provided a signal to a model 623B high voltage amplifier (Trek Inc.), connected to the electrodes of the microfluidic device.[41] The emulsion containing T7 promoter mutants was sorted with a rate of 1000 droplets per second. By adjusting the threshold value during sorting, approximately 1% of the droplets were sorted out for sequence analysis.

Screening of Dual-Color RBS Library with Fluorescence-Activated Cell Sorting (FACS). The double emulsion bearing the RBS library was analyzed or sorted using a flow cytometer (SONY SH800). To detect sfGFP, a laser tuned to 488 nm was used for excitation, and the emission was detected through a 525 nm filter. To detect sfmCherry, a laser tuned to 560 nm was used for excitation, and the emission was detected through a 600 nm filter. The droplets were analyzed and sorted at rates of 10,000 events per second. Data were analyzed using FlowJo (Tree Star, Ashland, OR). The sorting gate was defined as the ratio of sfGFP to sfmCherry. The threshold value was slightly adjusted during sorting, and approximately 0.46% of the droplets were sorted from the original library for sequence analysis.

NGS and Data Analysis of the RBS Library. The DNA segment having 10 nucleotide mutations was amplified by PCR from library templates extracted out of the droplets collected. Primers G526-f and G649-r were used to generate a 142 base pair DNA fragment. The number of cycles was determined empirically to prevent over-propagation. The products were further refined and indexed with NEBNext Ultra II End Repair/dA-Tailing Module (E7546S, NEB) and NEBNext Ultra II Ligation Module (E7595S, NEB). Library quality was assessed using the Agilent DNA 1000 Bioanalyzer kit (5067−1504, Agilent), according to the manufacturer's instructions. Libraries were sequenced on a MiSeq instrument using Miseq Reagent Kit v3 150 cycles (Illumina).

Characterization of RBS Mutants with Cell-Free Proteins Synthesis. Forty-eight RBS mutants were selected after screening and were assessed by their high enrichment rates. Based on these sequences, 48 DNA oligos were synthesized bearing these RBS mutant sequences and cloned into the pJL1-sfGFP vector to substitute the wtRBS (Supplementary Figure S4). The plasmids containing RBS mutants were extracted from *E. coli* (C3019H, NEB), and the mutated RBS was confirmed by Sanger sequencing. The entire expression cassette spanning from the T7 promoter to the T7 terminator was amplified by PCR and purified using a PCR Clean-Up Kit (28104, Qiagen). To express sfGFP, 200 ng of DNA product of each mutant was added to a 15 μL CFPS reaction.[42,44] After incubating at 30 °C for 16 h, the end-point fluorescence of sfGFP was measured with a plate reader (Synergy H1, BioTek).

To analyze the kinetics of RBS mutants, reconstituted cell-free protein synthesis was set up using PURExpress In Vitro Protein Synthesis Kit (E6800S, NEB) using 0.313 nM of DNA template for each mutant. The kinetic profile was acquired using a real-time PCR machine to collect the fluorescence signal of sfGFP at 5 min intervals for a total of 210 min.

Design and Characterization of RBS Mutants with *V. natriegens* Cell-Free Protein Synthesis. For *V. natriegens* CFPS reactions, the PANOx-SP system was utilized as previously described. Briefly, 5 μL of CFPS reactions were assembled with the Echo 525 in 384-well plates by individually transferring 100 nL of plasmid at 100 μg/nL and 4500 nL of the reaction mix. For the *V. natriegens* CFPS reactions, the

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

reaction mix contained the following components: 1.2 mM ATP; 0.85 mM each of GTP, UTP, and CTP; 34 μg/mL folinic acid; 170 μg/mL of *E. coli* tRNA mixture; 16 μg/mL T7 RNA polymerase; 3 mM for each of the 20 standard amino acids; 0.33 mM nicotinamide adenine dinucleotide (NAD); 0.27 mM coenzyme-A (CoA); 1.5 mM spermidine; 1 mM putrescine; 4 mM sodium oxalate; 290 mM potassium glutamate; 10 mM ammonium glutamate; 6 mM magnesium glutamate; 57 mM HEPES, pH 7.2; 67 mM phosphoenolpyruvate (PEP), and 4 μL (27% v/v) of cell extract. Each reaction was repeated in quadruplicate. Once all reactions were assembled, the 384-well plates were sealed and incubated for 20 h at 30 °C. Fluorescence was measured using a BioTek SynergyTM H1 plate reader with excitation at 485 nm, emission at 528 nm, and cutoff at 510 nm in 384-well plates.

Characterization of the Evolved RBS Mutants by sfGFP Expression in *E. coli* Cell. For *in vivo* characterization of the evolved RBS mutants, the plasmids were transformed into BL21(DE3)pLysS (C606010, Invitrogen). One mL of LB media was inoculated with a single colony in a deep 96-well plate and grown overnight at 37 °C. Three single colonies of each sample were picked and cultivated in parallel to estimate the standard deviation. Next, the saturated culture was diluted 50-fold into 500 μL of 2xYT-rich media in deep 96-well plates with proper antibiotics. Cells were first grown at 37 °C with 200-rpm agitation for 2 h to the early mid-exponential phase when 1 mM IPTG was added to induce sfGFP expression. Then, all samples were cultivated at 30 °C for a total of 6 h. Fluorescence and OD600 readings were measured at an interval of 1 h. The ratio of fluorescence/OD600 was used to indicate the activities of RBS mutants.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.2c00050.

> The entire workflow of regulatory part prototyping implemented by use of cell-free protein synthesis, microfluidics, and NGS (Figure S1); analysis of deviation and improve the precision by use of dual-color strategy (Figure S2); microfluidic chip design to convert single emulsion to double emulsion (Figure S3); construction and verification of an RBS random library in dual-color reporter plasmid (Figure S4); optimize the concentration of DNA template in CFPS (Figure S5); the ratio of RFU (sfGFP) to OD600 of *E. coli* strain containing nine active RBS mutants (Figure S6); DNA oligomers (Table S1); plasmid sequence (Table S2); and sequence and enrichment rate of RBS mutant (Table S3) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Michael C. Jewett − *Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208-3120, United States; Interdisciplinary Biological Sciences Program, Chemistry of Life Processes Institute, and Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208-3120, United States;* orcid.org/0000-0003-2948-6211; Email: m-jewett@northwestern.edu

Robert Nicol − *Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States;* orcid.org/0000-0003-1932-1392; Email: nicol@broadinstitute.org

Authors

Rui GAN − *Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States*

Maria D. Cabezas − *Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208-3120, United States*

Ming Pan − *Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States;* Present Address: Xellar Biosystems, 767 Concord Avenue, Suite 2−2, Cambridge, Massachusetts 02138, United States

Huaibin Zhang − *Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States;* Present Address: Arbor Biotechnologies, 20 Acorn Park Dr Tower 500, Cambridge, Massachusetts 02140, United States

Gang Hu − *Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States;* Present Address: Transition Bio, Inc. One Mifflin Place, STE. 400, Cambridge, Massachusetts 02138, United States.

Lauren G. Clark − *Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208-3120, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.2c00050

Author Contributions

◆R.G., M.D.C. and M.P. contributed equally to this work. R.G., M.D.C., and L.G.C., conducted the experiments; M.P., H.Z., G.H., and R.N. developed the microfluidic platform; and R.G., M.D.C., M.C.J., and R.N. designed the work and wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the U.S. Defense Advanced Research Projects Agency's (DARPA) Living Foundries program Award HR0011-15-C-0084. M.C.J. also gratefully acknowledges the Packard Foundation.

■ REFERENCES

(1) Guzman, L. M.; Belin, D.; Carson, M. J.; Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **1995**, *177*, 4121−4130.

(2) Inouye, S.; Inouye, M. Up-promoter mutations in the lpp gene of *Escherichia coli*. *Nucleic Acids Res.* **1985**, *13*, 3101−3110.

(3) Park, Y.; Espah Borujeni, A.; Gorochowski, T. E.; Shin, J.; Voigt, C. A. Precision design of stable genetic circuits carried in highly-insulated *E. coli* genomic landing pads. *Mol. Syst. Biol.* **2020**, *16*, No. e9584.

(4) Fontanarrosa, P.; Doosthosseini, H.; Borujeni, A. E.; Dorfan, Y.; Voigt, C. A.; Myers, C. Genetic Circuit Dynamics: Hazard and Glitch Analysis. *ACS Synth. Biol.* **2020**, *9*, 2324−2338.

(5) Shin, J.; Zhang, S.; Der, B. S.; Nielsen, A. A. K.; Voigt, C. A. (2020) Programming *Escherichia coli* to function as a digital display. *Mol. Syst. Biol.* **2020**, No. e9401.

(6) Brophy, J. A. N.; Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **2014**, *11*, 508−520.

(7) Moore, S. J.; MacDonald, J. T.; Wienecke, S.; Ishwarbhai, A.; Tsipa, A.; Aw, R.; Kylilis, N.; Bell, D. J.; McClymont, D. W.; Jensen, K.; Polizzi, K. M.; Biedendieck, R.; Freemont, P. S. Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from nonmodel bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E4340−E4349.
(8) Kelwick, R.; Webb, A. J.; MacDonald, J. T.; Freemont, P. S. Development of a *Bacillus subtilis* cell-free transcription-translation system for prototyping regulatory elements. *Metab. Eng.* **2016**, *38*, 370−381.
(9) Wang, H.; Li, J.; Jewett, M. C. Development of a Pseudomonas putida cell-free protein synthesis platform for rapid screening of gene regulatory elements. *Synth. Biol.* **2018**, *3*, No. ysy003.
(10) Des Soye, B. J.; Davidson, S. R.; Weinstock, M. T.; Gibson, D. G.; Jewett, M. C. Establishing a High-Yielding Cell-Free Protein Synthesis Platform Derived from *Vibrio natriegens*. *ACS Synth. Biol.* **2018**, *7*, 2245−2255.
(11) Wiegand, D. J.; Lee, H. H.; Ostrov, N.; Church, G. M. (2019) Cell-free Protein Expression Using the Rapidly Growing Bacterium *Vibrio natriegens*. *J. Visualized Exp.* **2019**, No. e59495.
(12) Jewett, M. C.; Swartz, J. R. Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **2004**, *86*, 19−26.
(13) Bogart, J. W.; Cabezas, M. D.; Vögeli, B.; Wong, D. A.; Karim, A. S.; Jewett, M. C. Cell-Free Exploration of the Natural Product Chemical Space. *Chembiochem* **2021**, *22*, 84−91.
(14) Rasor, B. J.; Vögeli, B.; Landwehr, G. M.; Bogart, J. W.; Karim, A. S.; Jewett, M. C. Toward sustainable, cell-free biomanufacturing. *Curr. Opin. Biotechnol.* **2021**, *69*, 136−144.
(15) Rasor, B. J.; Yi, X.; Brown, H.; Alper, H. S.; Jewett, M. C. An integrated in vivo/in vitro framework to enhance cell-free biosynthesis with metabolically rewired yeast extracts. *Nat. Commun.* **2021**, *12*, No. 5139.
(16) Rasor, B. J.; Vögeli, B.; Jewett, M. C.; Karim, A. S. Cell-Free Protein Synthesis for High-Throughput Biosynthetic Pathway Prototyping. *Methods Mol. Biol.* **2022**, *2433*, 199−215.
(17) 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.
(18) Dudley, Q. M.; Karim, A. S.; Nash, C. J.; Jewett, M. C. In vitro prototyping of limonene biosynthesis using cell-free protein synthesis. *Metab. Eng.* **2020**, *61*, 251−260.
(19) Liew, F. E.; Nogle, R.; Abdalla, T.; Rasor, B. J.; Canter, C.; Jensen, R. O.; Wang, L.; Strutz, J.; Chirania, P.; Tissera, S.; de Mueller, A. P.; Ruan, Z.; Gao, A.; Tran, L.; Engle, N. L.; Bromley, J. C.; Daniell, J.; Conrado, R.; Tschaplinski, T. J.; Giannone, R. J.; Hettich, R. L.; Karim, A. S.; Simpson, S. D.; Brown, S. D.; Leang, C.; Jewett, M. C.; Köpke, M. Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat. Biotechnol.* **2022**, *40*, 335−344.
(20) Karim, A. S.; Dudley, Q. M.; Juminaga, A.; Yuan, Y.; Crowe, S. A.; Heggestad, J. T.; Garg, S.; Abdalla, T.; Grubbe, W. S.; Rasor, B. J.; Coar, D. N.; Torculas, M.; Krein, M.; Liew, F. E.; Quattlebaum, A.; Jensen, R. O.; Stuart, J. A.; Simpson, S. D.; Köpke, M.; Jewett, M. C. In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design. *Nat. Chem. Biol.* **2020**, *16*, 912−919.
(21) Rybnicky, G. A.; Dixon, R. A.; Kuhn, R. M.; Karim, A. S.; Jewett, M. C. Development of a Freeze-Dried CRISPR-Cas12 Sensor for Detecting Wolbachia in the Secondary Science Classroom. *ACS Synth. Biol.* **2022**, *11*, 835−842.
(22) Huang, A.; Nguyen, P. Q.; Stark, J. C.; Takahashi, M. K.; Donghia, N.; Ferrante, T.; Dy, A. J.; Hsu, K. J.; Dubner, R. S.; Pardee, K.; Jewett, M. C.; Collins, J. J. (2018) BioBits Explorer: A modular synthetic biology education kit. *Sci. Adv.* **2018**, No. eaat5105.
(23) Stark, J. C.; Huang, A.; Hsu, K. J.; Dubner, R. S.; Forbrook, J.; Marshalla, S.; Rodriguez, F.; Washington, M.; Rybnicky, G. A.; Nguyen, P. Q.; Hasselbacher, B.; Jabri, R.; Kamran, R.; Koralewski, V.; Wightkin, W.; Martinez, T.; Jewett, M. C. BioBits Health: Classroom Activities Exploring Engineering, Biology, and Human Health with Fluorescent Readouts. *ACS Synth. Biol.* **2019**, *8*, 1001−1009.
(24) Stark, J. C.; Huang, A.; Nguyen, P. Q.; Dubner, R. S.; Hsu, K. J.; Ferrante, T. C.; Anderson, M.; Kanapskyte, A.; Mucha, Q.; Packett, J. S.; Patel, P.; Patel, R.; Qaq, D.; Zondor, T.; Burke, J.; Martinez, T.; Miller-Berry, A.; Puppala, A.; Reichert, K.; Schmid, M.; Brand, L.; Hill, L. R.; Chellaswamy, J. F.; Faheem, N.; Fetherling, S.; Gong, E.; Gonzalzles, E. M.; Granito, T.; Koritsaris, J.; Nguyen, B.; Ottman, S.; Palffy, C.; Patel, A.; Skweres, S.; Slaton, A.; Woods, T.; Donghia, N.; Pardee, K.; Collins, J. J.; Jewett, M. C. (2018) BioBits Bright: A fluorescent synthetic biology education kit. *Sci. Adv.* **2018**, No. eaat5107.
(25) Jaroentomeechai, T.; Stark, J. C.; Natarajan, A.; Glasscock, C. J.; Yates, L. E.; Hsu, K. J.; Mrksich, M.; Jewett, M. C.; DeLisa, M. P. Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **2018**, *9*, No. 2686.
(26) Kightlinger, W.; Duncker, K. E.; Ramesh, A.; Thames, A. H.; Natarajan, A.; Stark, J. C.; Yang, A.; Lin, L.; Mrksich, M.; DeLisa, M. P.; Jewett, M. C. A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.* **2019**, *10*, No. 5404.
(27) Hershewe, J. M.; Warfel, K. F.; Iyer, S. M.; Peruzzi, J. A.; Sullivan, C. J.; Roth, E. W.; DeLisa, M. P.; Kamat, N. P.; Jewett, M. C. Improving cell-free glycoprotein synthesis by characterizing and enriching native membrane vesicles. *Nat. Commun.* **2021**, *12*, No. 2363.
(28) Grubbe, W. S.; Rasor, B. J.; Krüger, A.; Jewett, M. C.; Karim, A. S. Cell-free styrene biosynthesis at high titers. *Metab. Eng.* **2020**, *61*, 89−95.
(29) Stark, J. C.; Jaroentomeechai, T.; Moeller, T. D.; Hershewe, J. M.; Warfel, K. F.; Moricz, B. S.; Martini, A. M.; Dubner, R. S.; Hsu, K. J.; Stevenson, T. C.; Jones, B. D.; DeLisa, M. P.; Jewett, M. C. (2021) On-demand biomanufacturing of protective conjugate vaccines. *Sci. Adv.* **2021**, No. eabe9444.
(30) Chappell, J.; Jensen, K.; Freemont, P. S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.* **2013**, *41*, 3471−3481.
(31) Gan, R.; Jewett, M. C. Evolution of translation initiation sequences using in vitro yeast ribosome display. *Biotechnol. Bioeng.* **2016**, *113*, 1777−1786.
(32) Agresti, J. J.; Antipov, E.; Abate, A. R.; Ahn, K.; Rowat, A. C.; Baret, J.-C.; Marquez, M.; Klibanov, A. M.; Griffiths, A. D.; Weitz, D. A. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4004−4009.
(33) 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.
(34) Shembekar, N.; Chaipan, C.; Utharala, R.; Merten, C. A. Droplet-based microfluidics in drug discovery, transcriptomics and high-throughput molecular genetics. *Lab Chip* **2016**, *16*, 1314−1331.
(35) Hori, Y.; Kantak, C.; Murray, R. M.; Abate, A. R. Cell-free extract based optimization of biomolecular circuits with droplet microfluidics. *Lab Chip* **2017**, *17*, 3037−3042.
(36) 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.
(37) Davidson, E. A.; van Blarcom, T.; Levy, M.; Ellington, A. D. Emulsion based selection of T7 promoters of varying activity. *Pac. Symp. Biocomput.* **2010**, 433−443.
(38) 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.
(39) Tschirhart, T.; Shukla, V.; Kelly, E. E.; Schultzhaus, Z.; NewRingeisen, E.; Erickson, J. S.; Wang, Z.; Garcia, W.; Curl, E.;

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Egbert, R. G.; Yeung, E.; Vora, G. J. Synthetic Biology Tools for the Fast-Growing Marine Bacterium *Vibrio natriegens*. *ACS Synth. Biol.* **2019**, *8*, 2069−2079.
(40) Salis, H. M. The ribosome binding site calculator. *Methods Enzymol.* **2011**, *498*, 19−42.
(41) Baret, J.-C.; Miller, O. J.; Taly, V.; Ryckelynck, M.; El-Harrak, A.; Frenz, L.; Rick, C.; Samuels, M. L.; Hutchison, J. B.; Agresti, J. J.; Link, D. R.; Weitz, D. A.; Griffiths, A. D. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **2009**, *9*, 1850−1858.
(42) Hong, S. H.; Ntai, I.; Haimovich, A. D.; Kelleher, N. L.; Isaacs, F. J.; Jewett, M. C. Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.* **2014**, *3*, 398−409.
(43) Kwon, Y.-C.; Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **2015**, *5*, No. 8663.
(44) Martin, R. W.; Des Soye, B. J.; Kwon, Y.-C.; Kay, J.; Davis, R. G.; Thomas, P. M.; Majewska, N. I.; Chen, C. X.; Marcum, R. D.; Weiss, M. G.; Stoddart, A. E.; Amiram, M.; Ranji Charna, A. K.; Patel, J. R.; Isaacs, F. J.; Kelleher, N. L.; Hong, S. H.; Jewett, M. C. Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* **2018**, *9*, No. 1203.

https://doi.org/10.1021/acssynbio.2c00050
ACS Synth. Biol. XXXX, XXX, XXX−XXX