ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Leveraging Active Learning to Establish Efficient In Vitro Transcription and Translation from Bacterial Chromosomal DNA

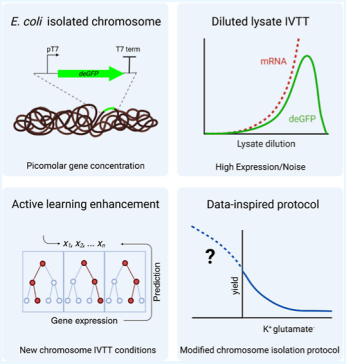
Leonardo Morini, Andrei Sakai, Mahesh A. Vibhute, Zef Koch, Margo Voss, Ludo L. J. Schoenmakers, and Wilhelm T. S. Huck*

Cite This: *ACS Omega* 2024, 9, 19227−19235

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Gene expression is a fundamental aspect in the construction of a minimal synthetic cell, and the use of chromosomes will be crucial for the integration and regulation of complex modules. Expression from chromosomes in vitro transcription and translation (IVTT) systems presents limitations, as their large size and low concentration make them far less suitable for standard IVTT reactions. Here, we addressed these challenges by optimizing lysate-based IVTT systems at low template concentrations. We then applied an active learning tool to adapt IVTT to chromosomes as template DNA. Further insights into the dynamic data set led us to adjust the previous protocol for chromosome isolation and revealed unforeseen trends pointing at limiting transcription kinetics in our system. The resulting IVTT conditions allowed a high template DNA efficiency for the chromosomes. In conclusion, our system shows a protein-to-chromosome ratio that moves closer to in vivo biology and represents an advancement toward chromosome-based synthetic cells.



INTRODUCTION

Recent advancements in the quest for the minimal synthetic cell have been substantial, especially in critical areas such as physicochemical homeostasis,[1−4] DNA replication,[5−7] membrane growth,[8,9] and membrane scission.[5,10−12] The emerging synthetic cell will ultimately harness its own gene expression machinery to regulate these complex features. In vitro transcription-translation (in vitro transcription and translation (IVTT)) is a well-established technology for the synthesis of proteins from plasmids and linear DNA encoding for a small number of genes.[13−19] Previous reports on phage genomes and large plasmids have shown that it is possible to express in the order of tens to a few hundred genes using DNA templates longer than 100 kbp.[6,20] Whereas these examples are excitingly close to a formerly postulated minimal genome (113 kbp with 151 genes[21]), top-down experimental work on *Mycoplasma* established that the current minimal synthetic genome requires at least 493 genes.[22,23] From a bottom-up perspective, it is conceivable that the construction of synthetic cells requires similarly sized genomes, and these will need to be expressed in vitro. Presently, we do not know if there are limits to the size of template DNA that can be used in IVTT and whether IVTT will function efficiently from entire chromosomes that can be a thousand times larger than typically used plasmids.[24] Previous work has shown that gene expression from bacterial chromosomal DNA is possible, although expression levels of single genes were exceedingly low compared to the ones achieved with short templates.[25,26] Such low expression levels would prevent chromosome-based synthetic cells from expressing observable phenotypes and, in the long run, from sustaining a replicating system. Since current IVTT has been developed on abundant and short DNA templates, we hypothesize that there is room for optimization for gene expression from chromosomal DNA.[27−29]

Here, we aim to find optimized conditions for IVTT from chromosomal DNA, using both rational analysis and active learning tools.[28] Our results show that high concentrations of cell lysate and template DNA are not necessary to reach high gene expression. Finally, we developed an IVTT system for chromosomal DNA with expression levels comparable to those obtained with plasmids. Our findings overcome the bottlenecks of working with picomolar DNA concentrations in lysate-based systems and expand the applications of IVTT to megabase-pair chromosomes.

RESULTS AND DISCUSSION

Diluted Cell-Free Extracts Increase Protein Synthesis and mRNA Levels at Low Template Concentrations. In our study, an *Escherichia coli* cell lysate is used to carry out IVTT reactions from a construct expressing *deGFP* under a T7 promoter. The orthogonal promoter allows the independent tuning of transcription and translation by the external addition

Received: January 4, 2024
Revised: March 29, 2024
Accepted: April 3, 2024
Published: April 16, 2024

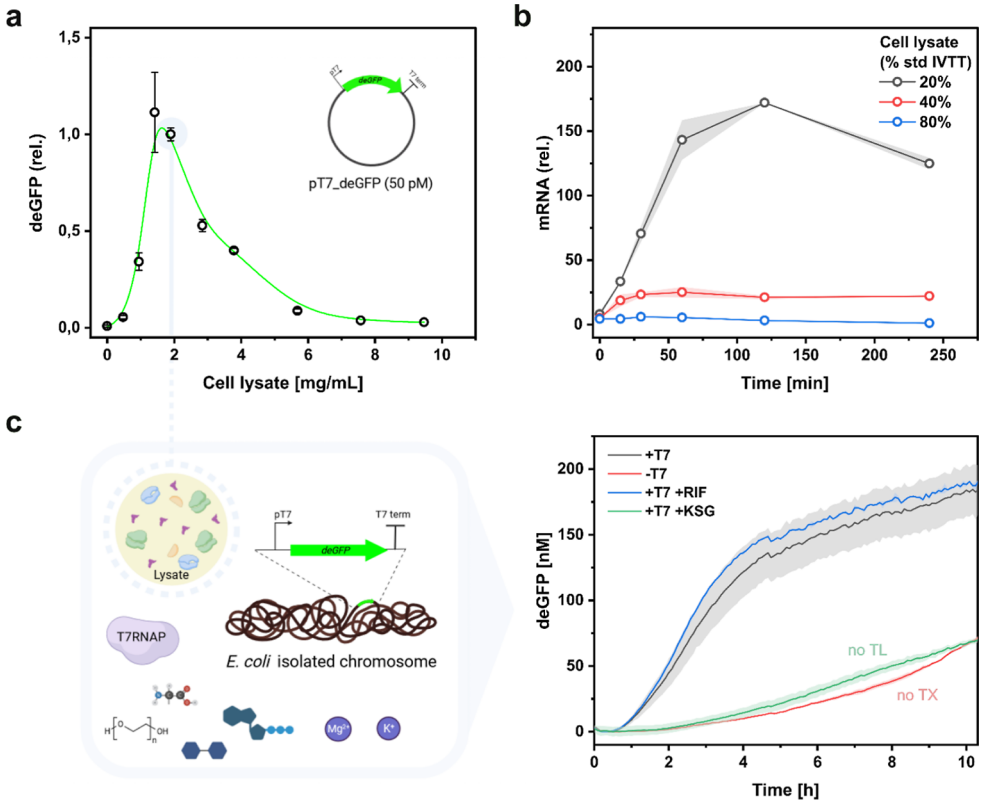


Figure 1. Diluted cell-free extracts increase protein synthesis and mRNA levels at low template concentrations. (a) Relative deGFP expression in IVTT reactions with 50 pM plasmid DNA using different concentrations of lysate. The data points are taken from 2 different experiments, both normalized by the expression levels of the shared condition with 1.9 mg/mL lysate (see Supporting Information). (b) Relative *deGFP* mRNA quantification with qRT-PCR in IVTT reactions from 50 pM plasmid DNA and different concentrations of lysate. Standard IVTT: 9.47 mg/mL lysate. (c) Schematic drawing with the rationale behind the new optimal lysate concentration adopted in IVTT from chromosomal DNA and the deGFP time course from IVTT reactions using 50 pM chromosomes and 1.9 mg/mL lysate with negative controls. T7: T7 RNA polymerase; RIF: rifampicin; KSG: kasugamycin; +T7: full IVTT; −T7: no transcription; +T7 + KSG: no translation; +T7 + RIF: antibiotic control. Created with BioRender.com.

of T7 RNA polymerase (T7RNAP). The construct is first expressed from plasmids and then from the chromosome of a modified *E. coli* strain that carries it as a chromosomal integration.

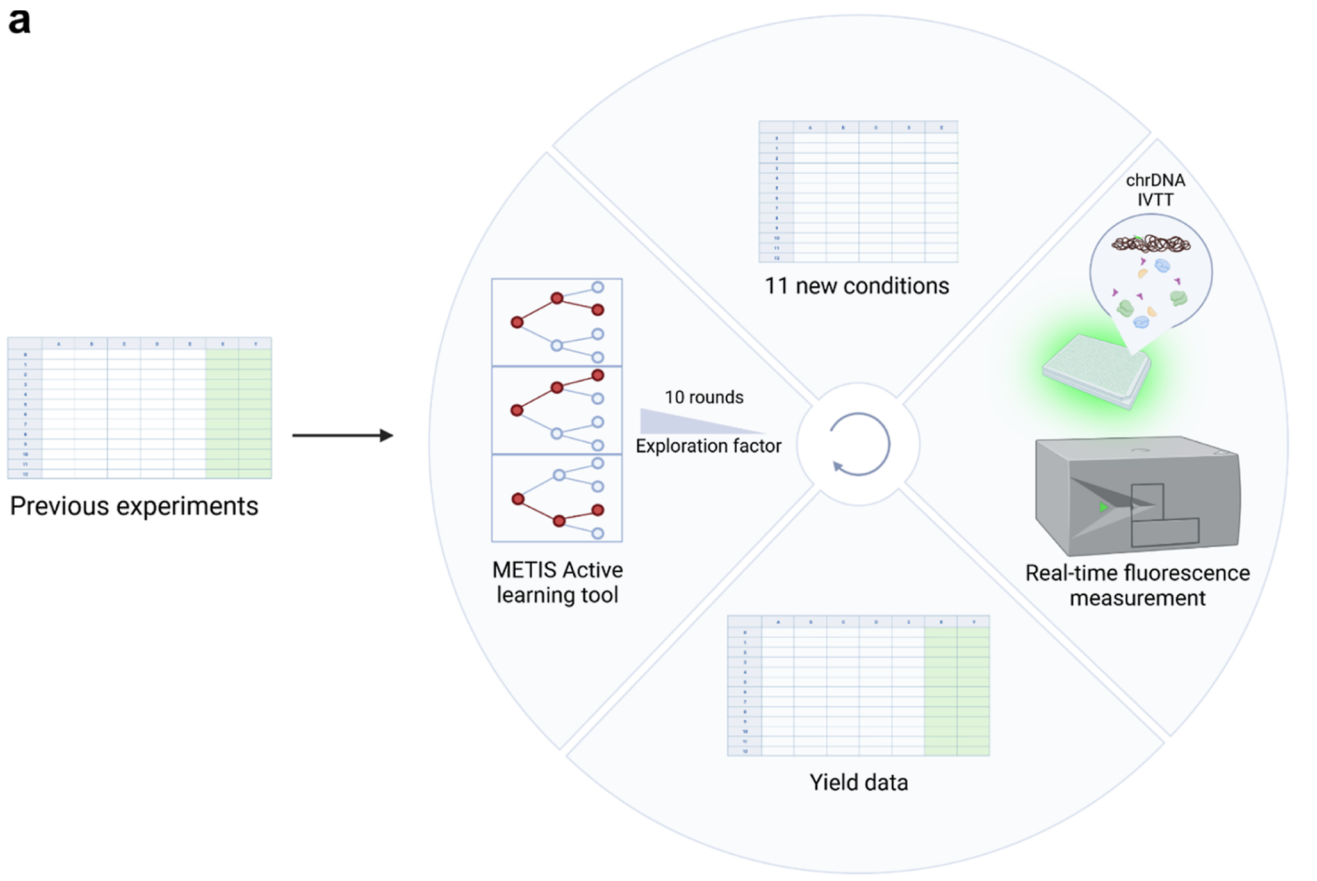
Isolated chromosomes in solution have a higher viscosity than that of plasmids. High viscosity and molecular weight limit the yield of the isolation protocol to picomolar concentrations (around 150 pM). Initial attempts of in vitro transcription from chromosomal DNA showed a clear increase in mRNA levels (Figure S1), but the same template (approximately 50 pM final DNA concentration) did not produce a measurable expression in standard IVTT reactions (not shown).

To optimize the gene expression yields, we first used plasmid DNA as an initial model. We hypothesized that because of the very low gene copy number present when using chromosomal DNA, transcription would become a limiting factor due to competing degradation reactions in the cell lysate. Therefore, we used picomolar concentrations of plasmid DNA to mimic the gene concentration from chromosomes and tested different dilutions of lysate in IVTT. Interestingly, the expression of *deGFP* reached a maximum when only 1.4−1.9 mg/mL of lysate (~15−20% of the concentration used in the standard protocol) was used (Figure 1a). The reaction reached a plateau after 10 h (Figure S2). When mRNA levels were measured over time, a clear correlation between higher concentrations of lysate and lower levels of mRNA was seen (Figure 1b), indicating that the lysate affects total mRNA levels by either increasing their degradation rate or reducing resource competition through metabolic side reactions. The large difference between samples with 20, 40, and 80% lysate does not seem to support a linear model for mRNA kinetics. Limiting degradation rates, cooperativity, or resource competition with mRNA synthesis could play a role in the nonlinear effect. Further experiments at higher dilutions of lysate confirmed that mRNA levels kept increasing when the lysate concentration was reduced, reaching a further 20-fold increase when lysate was completely omitted (Figure S3).

Finally, the condition with 20% lysate was used to express the *deGFP* gene under a T7 promoter from chromosomal DNA. The increase in fluorescence over the first 4 h could be attributed to deGFP expression, as selectively blocking either transcription or translation resulted in near complete suppression of fluorescence increase (Figure 1c).

Active Learning Tool Refines Reaction Buffer for IVTT from Chromosomes. Having established gene expression from low concentrations of chromosomal DNA in diluted cell lysate, we surmise that the very different physical presentation of a gene inside the chromosomes,[24,30,31] as compared to a plasmid or linear DNA fragment, leads to suboptimal gene expression. Reoptimizing IVTT is challenging, as there are many components (including lysate, K^+, Mg^{2+}, and poly-

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235



	Initial	Best performing
PEG8000	1.5 % w/v	2.5% w/v
Cell lysate	1.9 mg/mL	1.9 mg/mL
T7RNAP	1 U/μL	1.7 U/μL
Maltose	10 mM	8.3 mM
Mg-Glut	4.16 mM	8.3 mM
K-Glut	66.7 mM	0 mM
chrDNA (150 pM)	33% v/v	25% v/v

Figure 2. Active learning tool refines the reaction buffer for IVTT from chromosomes. (a) Schematic workflow of the active learning tool adapted from Pandi et al. (2022).[28] Previous experiments ($n \sim 30$) were used as an initial training data set. At every round, new conditions were suggested, tested, and included in the dynamic data set for further training. The exploration factor decreased over the rounds. chrDNA: chromosomal DNA. Created with BioRender.com (b) Boxplot with the yields of the conditions tested at each round. The yield is defined as the $\log_{10}$ of the ratio between each new condition and the initial reference condition (see Materials and Methods). The best-performing condition is highlighted in the red circle and analyzed in the following panel. (c) Comparisons between the initial IVTT condition and the best-performing one. Yellow: parameters that were modified before the start of the active learning process (in the case of the lysate, the active learning maintained the same value); red: parameters that were decreased at the end of the active learning; and green: parameters that were increased at the end of the active learning.

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235

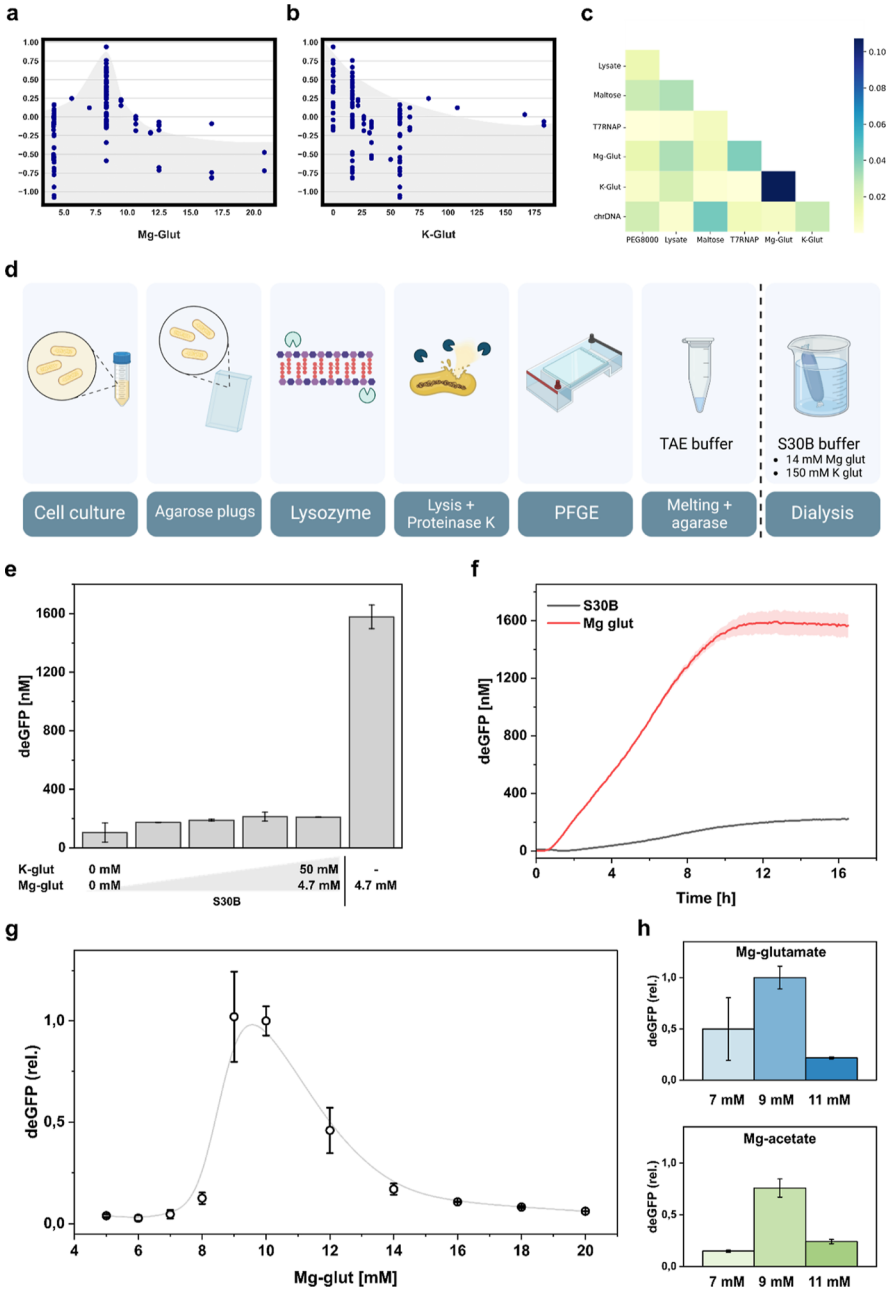
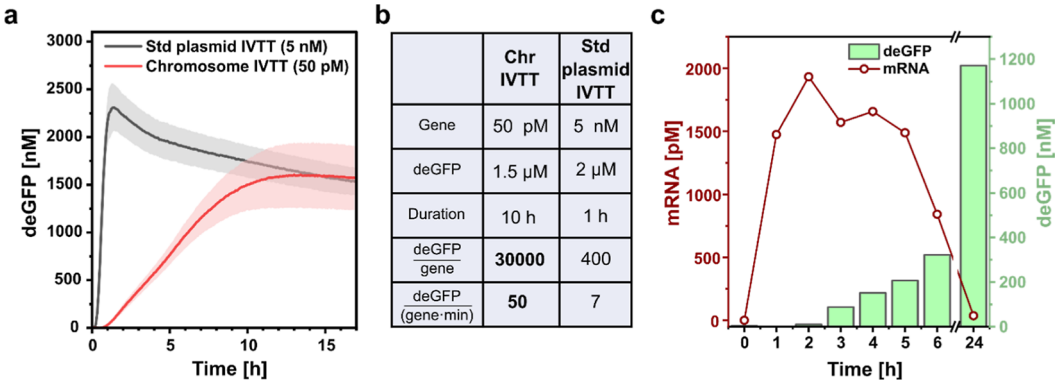


Figure 3. Data-guided adjustment of the chromosome isolation protocol enhances protein synthesis at micromolar levels. (a, b) Results for Each Metabolite function of METIS for magnesium glutamate and potassium glutamate, respectively, after 10 rounds of experiments. The data points represent each tested concentration (x-axis, [mM]) and the achieved yield in gene expression. The gray areas were added manually to highlight the patterns. (c) NonLinear Interactions (function of METIS) between the tested components in the IVTT reaction. (d) Schematic illustration of the chromosome isolation protocol (see Materials and Methods). The dash line separates the omitted dialysis step, which exchanges TAE buffer with S30B buffer, introducing additional magnesium and potassium glutamate. Created with BioRender.com (e) deGFP levels of IVTT from undialyzed chromosomes with increasing concentrations of S30B buffer or, for the last sample, the same amount in magnesium glutamate only. Different pH values (6.5, 7.5, and 8.5) did not affect the reaction significantly (not shown). We cannot exclude the role of the Tris buffer in the inhibitory effect of S30B, as that increases the ionic strength (see Results and Discussion). (f) deGFP time course of IVTT from undialyzed chromosomes, where the same amount of magnesium glutamate was supplemented alone (Mg glut) or as S30B. (g) Relative deGFP levels of IVTT from undialyzed chromosomes at different concentrations of magnesium glutamate. (h) Relative deGFP levels of IVTT from undialyzed chromosomes at different concentrations of magnesium using either glutamate or acetate as the counterion. All values were normalized by the protein levels under the conditions of 9 mM Mg-glutamate.

(ethylene glycol)) that interact with each other. Instead of varying concentrations of individual parameters, recently developed machine learning tools allow for a more effective probing of a high-dimensional parameter space. In particular, active learning algorithms can be trained on preexisting or dynamic data sets and predict the combined effects of multiple parameters on a user-defined function. This strategy speeds up an otherwise cumbersome and ultimately inferior analysis of parameters that have interconnected effects. We used the active learning tool METIS[28] to refine the IVTT conditions for

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235



	Chr IVTT	Std plasmid IVTT
Gene	50 pM	5 nM
deGFP	1.5 μM	2 μM
Duration	10 h	1 h
deGFP / gene	**30000**	400
deGFP / (gene·min)	**50**	7

Figure 4. High template efficiency of chromosomes in IVTT. (a) deGFP time course of standard IVTT from 5 nM plasmids and new IVTT reaction from 50 pM chromosomes. (b) Comparison of template efficiencies between IVTT reactions from chromosomes and plasmids from panel a. Full calculations are available in the Supporting Information. Chr: chromosome (c) Time course of mRNA (measured with qRT-PCR) and deGFP levels (fluorescence measurement) in IVTT from chromosomal DNA.

chromosomes by simultaneously varying the concentrations of the following main components: chromosomal DNA, T7 RNA polymerase, lysate, magnesium glutamate, potassium glutamate, PEG8000, and maltose. A preassembled feeding buffer (FB) containing NTPs, amino acids, cofactors, and other small molecules was kept constant (see Materials and Methods). The active learning was run for 10 rounds with 11 new conditions each and a decreasing exploration factor, thus yielding more focused combinations by the end of the process after some less biased initial rounds (Figure 2a). The active learning approach yielded an 8.6-fold increase in gene expression, and the best-performing condition was analyzed (Figure 2b).

Briefly, at the end of the 10 rounds, the suggested concentrations for PEG8000, T7RNAP, and magnesium glutamate were higher than the reference values, whereas the ones for chromosomes and potassium glutamate were lower (Figure 2c).

Interestingly, the suggested concentration for cell lysate had not changed from the optimal value that was found using plasmid DNA, suggesting that the optimal lysate concentration is not affected by the size of the template DNA. Moreover, METIS suggested that the T7 RNA polymerase concentration should be increased, indicating that a lack of mRNA was limiting gene expression. Counterintuitively, the algorithm allowed the chromosomal DNA concentration to decrease, indicating that the low template concentration might not be the most limiting parameter in our system. Finally, potassium glutamate was completely omitted from the suggested reaction buffer, even though it was not expected to have inhibitory effects on the IVTT reaction (and hence, there would be no need for elimination).

Data-Guided Adjustment of the Chromosome Isolation Protocol Enhances Protein Synthesis at Micromolar Levels. Magnesium and potassium are fundamental cations in IVTT as they stabilize active sites and macromolecular structures.[32−38] The complete absence of additional potassium glutamate under the best-performing condition prompted us to look deeper in the active learning model. We analyzed the yields of all the tested conditions as a function of the concentration of magnesium and potassium glutamate. The distribution for magnesium glutamate shows a sharp peak at the final concentration of 8.3 mM (Figure 3a). On the other hand, the pattern for potassium glutamate has its maximum at 0 mM, but the trend suggests a peak at negative values (Figure 3b). To the best of our knowledge, this is a rare example of an IVTT system that would give high expression yields in the absence of potassium glutamate. We then looked into the interactions between components, and we noticed that magnesium glutamate and potassium glutamate had the strongest interaction (Figure 3c). With this observation, we hypothesized two scenarios. First, glutamate could have an inhibitory effect on the system. Glutamate is known to have a role in the metabolism of IVTT reactions, and since it is the counterion of both magnesium and potassium, the two salts could cooperate in the inhibitory effect. However, the active model reduced only potassium glutamate because the IVTT system is much more sensitive to variations in magnesium concentration. In the second scenario, magnesium and potassium glutamate could interact in the model via their co-occurrence in the S30B buffer. The S30B buffer contains 14 mM magnesium glutamate and 150 mM potassium glutamate (pH 8 with Tris) and is used for dialyzing the cell lysate and the chromosomes. In the diluted lysate conditions, chromosomes are the main secondary source of potassium glutamate. This could also explain why the active learning model allowed the chromosomal DNA concentration to decrease. In the chromosome isolation protocol, the dialysis exchanges TAE buffer with S30B, removing the enzyme agarase and the agarose subunits after the digestion of the agarose plug (Figure 3d). Therefore, we decided to omit the dialysis in the S30B buffer from the chromosome isolation protocol. First, TAE, agarase, and digested agarose were confirmed not to interfere with IVTT (not shown). Then, the dialysis step was omitted, and undialyzed chromosomes were used in IVTT using either S30B or pure magnesium glutamate to restore the optimal magnesium concentration. Strikingly, gene expression reached more than 1.5 μM (above 40 mg/L) in the first 10 h when S30B was replaced by magnesium glutamate only (Figure 3e,f). We ruled out unspecific fluorescence from other sources (i.e., cell lysate and gene expression from endogenous promoters) by detecting a single native fluorescent band in PAGE under partially denaturing conditions (Figure S4). We further confirmed the expected size of deGFP (~27 kDa) in a fully denaturing PAGE, labeling the IVTT reaction with a fluorescent tRNALys (FluoroTect GreenLys) (Figure S5a,b). Next, different concentrations of magnesium glutamate were tested, and a sharp pattern with a peak around 10 mM could be seen, underlining the high sensitivity of the system to

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235

magnesium glutamate (Figure 3g). A similar pattern was observed when magnesium glutamate was replaced with magnesium acetate (Figure 3h). This excludes a critical role for glutamate in the inhibitory effect predicted by the active learning model. The results point out a narrower parameter space for IVTT from chromosomes compared with standard IVTT reactions. We speculate that the selection for lower concentrations of potassium glutamate is a shift of the reaction buffer toward more favorable conditions for T7 RNA polymerase activity and the synthesis of mRNAs. Noticeably, established protocols for in vitro transcription reactions based on T7 RNA polymerase use buffers that avoid potassium (and other monovalent cations) since high ionic strength reduces the polymerase affinity for its DNA template.[39,40] In standard IVTT reactions, potassium is included as it is fundamental for different processes, including ribosome stability,[34] but its deleterious effect on transcription is marginal in systems doped with high concentrations of template DNA. In our situation, instead, the reaction conditions suggested by active learning seem to reduce potassium in order to improve transcription in a system with a low template DNA concentration. Recent work reported a similar behavior when in vitro systems were optimized for different reactions, showing that translation was favored by higher potassium concentrations compared to transcription from T7 RNA polymerase.[41]

High Template Efficiency of Chromosomes in IVTT. Above, we established robust expression from low concentrations of chromosomal DNA under conditions that substantially differ from those typically used in plasmid-based IVTT systems. How do these systems compare in terms of efficiency? To obtain some kind of comparison, we determined the number of proteins produced per copy of gene between different IVTT systems as a measure of template efficiency (Figure 4a,b). Assuming a negligible degradation rate for deGFP, chromosomes can express in the order of ∼50 protein/(gene*min) for a final yield of ∼30,000 protein/gene. These values are orders of magnitude higher than the corresponding ones in standard IVTT from plasmids, where ∼7 proteins/(gene*min) are produced for a final yield of ∼400 protein/gene (Figure 4b).

Finally, in vivo studies have shown protein synthesis rates of ∼1 protein per mRNA molecule per second in a bacterial cell.[42] While it is challenging to measure this in our cell-free system, we can estimate protein synthesis rates for both the chromosomal and plasmid IVTT systems. As with plasmid DNA (Figure 1b), the mRNA levels in IVTT from chromosomes reached a steady state in the first hour of the reaction, ranging around 1.5 nM (Figure 4c). After the second hour, deGFP levels started to increase at an approximate rate of 75 nM/h, yielding a ratio of approximately 0.014 proteins/mRNA*s. These rates are comparable to protein synthesis rates obtained for plasmid DNA ∼ 0.006−0.033 protein/(mRNA*s).[14,43] We observed similar trends in mRNA levels, but not in protein levels when the reaction was carried out at 25 and 37 °C (Figure S6). Please note that we use only the initial time window when mRNA levels are at a steady state to compare our systems; protein synthesis was sustained for much longer, considering the final deGFP concentration after 24 h.

CONCLUSIONS

IVTT is a well-established and heavily optimized method for cell-free protein synthesis from plasmid DNA, but it has rarely been used to express proteins from chromosomal DNA. Future work on synthetic cells requires significantly improved protein synthesis yields in IVTT systems. We found that diluting the cell lysate promoted gene expression from low template concentrations. In particular, diluted lysate allows for higher mRNA levels without suppressing the translation activity. This is in contrast with current IVTT protocols that tend to work at a high concentration of both DNA template and lysate. Using an active learning workflow, we confirmed the beneficial effects of diluted cell lysate at low template concentrations and highlighted a potential inhibitory effect of potassium glutamate. With this information, we readapted the protocol for chromosome isolation in order to minimize potassium glutamate in our system, and we reached micromolar protein synthesis levels, pointing at a template efficiency orders of magnitude higher than what is typically achieved in standard IVTT from plasmids. It is important to note that our results are based on the expression of a gene under a T7 promoter, and therefore, we would expect different optimal conditions for gene expression that are dependent on the activity of the endogenous RNA polymerase from *E. coli*. In conclusion, we have shown how an active learning method is not only a black box route to improve yields but can also provide detailed insights into the complex biochemistry of cell-free expression systems. Our results prove that high gene expression can be achieved at lower concentrations of template DNA and cell lysate. This narrows the gap between the amount of proteins required to carry out gene expression and the amount that is synthesized *de novo*, an aspect that will be crucial for building a sustainable replicating system in vitro.[44] However, it is evident that there is much room for improvement of the protein synthesis efficiency in cell-free systems. An intuitive way to achieve this would be to avoid synthesizing excess mRNA, but our experiments show that transcription is already limiting in our system, and hence, this cannot be the solution. Therefore, improvement must be sought in the more efficient translation of mRNA (possibly by avoiding mRNA degradation or other routes that lead to mRNA not being translated, such as the formation of folded mRNA structures). In conclusion, the new IVTT regime and its underlying mechanisms define the principles of gene expression from chromosomal DNA and will prove valuable for the future development of synthetic cells.

MATERIALS AND METHODS

Strains and Plasmids. *E. coli* BL21 (DE3) STAR (pRARE) was used for cell lysate production (see the following section). For plasmid amplification, *E. coli* XL-1 Blue with pRSET5d *T7-deGFP-T7t His tag AmpR* (pMY9) was used, and the plasmids were purified using QIAprep Spin Miniprep Kit (QIAGEN), following the instructions from the manufacturer. Plasmid concentration was measured with a Nanodrop. For the isolated chromosomes, *E. coli* strain eGFP16 was obtained with the editing procedure based on the two-plasmid system from Jiang et al.[45] Briefly, pCas was transformed into chemically competent E. coli *B*L21(DE3). The strain was cultured in the presence of 10 mM L-arabinose to induce λ Red expression and grown to an OD600 of 0.5. Cells were made electrocompetent by repeated washings in ice-cold Milli-Q water. Next, pTF (100 ng) and donor DNA (400 ng) were coelectroporated into the competent *E. coli* BL21(DE3) pCas cells. Transformed cells were plated and grown at 30 °C overnight. Successful insertion was checked by genomic colony PCR against the new insert and IPTG

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235

induction with a fluorescent readout under an SD microscope and plate reader.

Cell Lysate Preparation. *E. coli* BL21 (DE3) STAR with pRARE plasmid for the expression of rare tRNAs was used for lysate production following the previously described protocol based on cell press lysis.[16,46] Briefly, *E. coli* was grown in liquid LB medium with 34 μg/mL chloramphenicol (Chl) at 37 °C overnight and scaled up in an intermediate culture in 2xYT (potassium phosphate buffer, pH 7) + Chl until an OD600 of 1.7−1.8 was reached. Then, the culture was scaled up in 2xYT, omitting Chl, and incubated until an OD600 of 1.7−1.8 was reached. Cells were pelleted and washed in S30A buffer (14 mM magnesium glutamate, 60 mM potassium glutamate, and 50 mM Tris−HCl, pH 8.2). The cell pellet was resuspended in S30A (0.9x pellet weight) and passed through the cell press 3 times for cell lysis. The extract was cleared by centrifugation at 12,000× g for 10 min at 4 °C, and the supernatant was aliquoted in microcentrifuge tubes and incubated at 37 °C and 220 rpm for 80 min with caps open to facilitate endogenous mRNA and DNA degradation. The extract was cleared by centrifugation as described above, and the supernatant was dialyzed in S30B buffer (14 mM magnesium glutamate, 150 mM potassium glutamate, 50 mM Tris−HCl, pH 8.2) using 10 kD MWCO cassettes. The dialyzed extract was cleared by centrifugation (as above), homogenized by inversion, aliquoted, flash frozen, and stored at −80 °C for further use. Protein concentration was measured with Pierce BCA Protein Assay Kits (ThermoScientific).

Chromosome Isolation. The chromosome isolation protocol was adapted from different works.[25,47−51] *E. coli* eGFP16 was grown in 25 mL of liquid LB medium until an OD600 of 1.2−1.5, and cells were pelleted. Pellets were recovered with 200 μL of TAE buffer (40 mM Tris-Acetate, 1 mM EDTA, pH 7) and stored in microcentrifuge tubes at −80 °C for further use. A 2% low-melting-point (LMP) agarose solution in TAE buffer was prepared, molten at 70 °C, and cooled to 50 °C. Thawed cell pellets were equilibrated at 50 °C, and equal volumes of resuspended *E. coli* pellet and 2% LMP agarose were mixed and cast into several 90−100 μL flat agarose plugs (1.5 mm thick) in a custom-made mold. The plugs were solidified at 4 °C. For the rupture of the cell wall, the plugs were incubated for 2 h at room temperature in lysozyme buffer (10 mM Tris pH 8, 10 mM EDTA, 100 mM NaCl, 0.4 mg/mL lysozyme from chicken egg white from Sigma-Aldrich) with mild inversion. For the cell lysis, the plugs were washed in TAE and incubated overnight at 50 °C in proteinase K buffer (100 mM EDTA pH 8, 0.2% w/v sodium deoxycholate, 1% w/v sodium lauroylsarcosine, and >600 U/mL Proteinase K from Fisher Scientific, corresponding to 40 μL per 1 mL of buffer) with mild shaking. Next, plugs were washed in TAE and loaded in a custom agarose gel (1% agarose in 0.5× TAE with a single full-width well) on a Pippin Pulse (Sage Science) device for Pulse Field Gel Electrophoresis (PFGE) to get rid of DNA fragments. The PFGE was run for 24 h with the following parameters: voltage = 65 (V), forward time at start = 300 ms (A), reverse time at start = 100 ms (B), increment to A = 255 ms (C), increment to B = 100 ms (D), increment to C = 0 ms (E), increment to D = 0 ms (F), number of steps per cycle = 250 (G). The plugs were stored in TAE for further use. Next, the plugs were molten at 70 °C, and the solution was cooled down to 42 °C. The agarose was digested, adding 1 U β-agarase I (New England Biolabs) per 100 μL of plugs. To remove TAE, agarose subunits, and β-agarase I, the isolated chromosome solution was dialyzed in S30B buffer (14 mM magnesium glutamate, 150 mM potassium glutamate, 50 mM Tris−HCl, pH 8.2) using a Spectra/Por 300kD MWCO dialysis tubing. Isolated chromosomes in solution could be stored for up to 1 week at 4 °C.

IVTT Reactions. IVTT reactions were carried out at 30 °C under the following conditions: 9.8 mg/mL lysate, 1 U/μL T7 RNA polymerase, 0.05−5 nM plasmid DNA, 8 mM magnesium glutamate, 60−100 mM potassium glutamate, 10 mM maltose, 2% v/v PEG8000 and a preassembled FB providing 50 mM HEPES (pH 8), 33 mM 3-PGA, 1.56 mM ATP, 1.56 mM GTP, 0.94 mM UTP, 0.94 mM CTP, 1.5 mM of each amino acid, 1 mM spermidine, 1 mM dithiothreitol, 0.2 mg/mL tRNAs, 260 μM CoA, 340 μM NAD, 780 μM cAMP, and 70 μM folinic acid. Intermediate changes of the protocol for gene expression from isolated chromosomes are described in the Results and Discussion. The final condition for IVTT from isolated chromosomes is the following: 1.9 mg/mL lysate, 1.7 U/μL T7RNAP, 37−50 pM chromosomal DNA (corresponding to 3 μL, in TAE buffer), 10 mM magnesium glutamate, 0 mM potassium glutamate, 8.3 mM maltose, 2.5% v/v PEG8000, and the preassembled FB. For IVTT from isolated chromosomes, particular care was taken when handling genomic material. In particular, pipet tips were cut to reduce shear damage, and the IVTT mix was preassembled and thoroughly mixed before the addition to isolated chromosomes.

Fluorescence Measurement. Fluorescence measurements were carried out on 12 μL IVTT reactions under controlled temperature using a 384-well plate with a flat bottom in a Tecan Infinite M200 Plate Reader or a Tecan Spark Plate Reader with shaking and readings every 5 min. The respective calibration curves were obtained with a purified deGFP of a known concentration. For the fluorescent measurements, all IVTT conditions were tested in duplicates, with the exception of Figures 4c and S6 (n = 1).

qRT-PCR. For mRNA quantification, 1 μL of IVTT reaction was diluted 1:10 in a 10% RNasecure (ThermoFisher) solution in Milli-Q water and incubated for 10 min at 60 °C to inactivate the RNases from the lysate. The chromosomal DNA was removed with a TURBO DNA-free Kit (ThermoFisher), following the instructions from the manufacturer. The samples were used as templates for the qRT-PCR reaction with the CAPITAL 1-Step qRT-PCR Green Master Mix (BiotechRabbit) and the primers FW and RV. The reaction was carried out with qTOWER^{3}G (AnalitikJena).

METIS Active Learning Algorithm. The active learning tool was run by readapting the code for optimizing IVTT systems from the original paper (Pandi et al., 2022).[28] METIS Optimization Notebook was run with the following parameters: number of combinations per round = 11, minimum drop size = 0.2 (μL), final reaction volume = 24 (μL), fixed parts: FB (20% v/v), days = 10, exploration factor = [1.41, 1.41, 1.41, 1, 1, 1, 1, 0.5, 0.5, 0.5]. The input metabolites used were PEG8000, Lysate, Maltose, T7RNAP, Mg-Glut, K-Glut, and gDNA (chromosomes). A reference sample (Figure 2c, "initial") was included for every round. All the conditions were assembled by hand-pipetting. To minimize the technical errors, the empty reaction volume (Milli-Q water) was used by diluting the stock solutions and increasing the pipetting volumes. For the same purpose, a common reaction mix (excluding chromosomal DNA) was prepared and split before adding individual specific components. The reactions were run

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235

in duplicates, and following the original paper, the yield function was defined as the log10 ratio between the fluorescence value of each tested condition and the reference at the plateau (>10 h in our case). Background initial fluorescence was not subtracted to avoid working with values close to 0 and mitigate the effect of random noise in the log scale (especially for inactive conditions) that could have biased the learning process. The active learning algorithm was launched with a preexisting data set of around 30 tested conditions (Results_0) from previous experiments. The functions *Results as BoxPlot*, *Results For Each Metabolite* and *NonLinear Interactions* were run after the 10 learning cycles.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c00111.

Additional experimental details, calculations, materials, and DNA sequences and methods (PDF)
(PDF)
(PDF)

■ AUTHOR INFORMATION

Corresponding Author

Wilhelm T. S. Huck − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands;* orcid.org/0000-0003-4222-5411; Email: w.huck@science.ru.nl

Authors

Leonardo Morini − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands;* orcid.org/0000-0002-3862-6055

Andrei Sakai − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands;* orcid.org/0000-0002-2700-8059

Mahesh A. Vibhute − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands;* orcid.org/0000-0002-5230-5290

Zef Koch − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands; HAN University of Applied Sciences, Nijmegen 6503GL, The Netherlands*

Margo Voss − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands*

Ludo L. J. Schoenmakers − *Institute for Molecules and Materials, Radboud University, Nijmegen 6525 AJ, The Netherlands; Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg 3400, Austria;* orcid.org/0000-0002-6882-6840

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c00111

Author Contributions

Leonardo Morini: Conceptualization, Formal analysis, Investigation, Writing−original draft. Andrei Sakai, Mahesh A. Vibhute: Conceptualization, Investigation. Ludo L. J. Schoenmakers: Conceptualization, Bacterial strain engineering. Zef Koch: Investigation, Isolation of bacterial chromosomes. Margo Voss: Investigation, Application of active learning workflow. Wilhelm T. S. Huck: Conceptualization, Writing−original draft, Supervision.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Mart Bartelds, Dr. Bob van Sluijs, Dr. Roel Maas, Dr. Maike M. K. Hansen, Dr. Hans A. Heus Óscar García Blay, and Carmen Grandi for the fruitful discussions and suggestions and Ing. Frank H. T. Nelissen, Ing. Aafke J. Jonker, Nicole van der Zanden and Emer Slaa for the technical support, and Michael Kwint for sharing the equipment for PFGE. This project has received funding from The Netherlands Organization for Scientific Research (NWO/OCW) via the "BaSyC-Building a Synthetic Cell" Gravitation grant (024.003.019).

■ ABBREVIATIONS

IVTT in vitro transcription translation
T7RNAP T7 RNA polymerase
K-glut potassium glutamate
Mg-glut magnesium glutamate
PEG8000 polyethylene glycol 8000
FB feeding buffer
Chr chromosomal DNA

■ REFERENCES

(1) Pols, T.; Sikkema, H. R.; Gaastra, B. F.; Frallicciardi, J.; Śmigiel, W. M.; Singh, S.; Poolman, B. A Synthetic Metabolic Network for Physicochemical Homeostasis. *Nat. Commun.* **2019**, *10* (1), 4239.

(2) Sikkema, H. R.; Gaastra, B. F.; Pols, T.; Poolman, B. Cell Fuelling and Metabolic Energy Conservation in Synthetic Cells. *ChemBioChem* **2019**, *20*, 2581−2592.

(3) Bailoni, E.; Partipilo, M.; Coenradij, J.; Grundel, D. A. J.; Slotboom, D. J.; Poolman, B. Minimal Out-of-Equilibrium Metabolism for Synthetic Cells: A Membrane Perspective. In *ACS Synthetic Biology*; American Chemical Society, 2023; pp 922−946.

(4) Partipilo, M.; Ewins, E. J.; Frallicciardi, J.; Robinson, T.; Poolman, B.; Slotboom, D. J. Minimal Pathway for the Regeneration of Redox Cofactors. *JACS Au* **2021**, *1* (12), 2280−2293.

(5) 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* (1), 1583.

(6) Libicher, K.; Hornberger, R.; Heymann, M.; Mutschler, H. In Vitro Self-Replication and Multicistronic Expression of Large Synthetic Genomes. *Nat. Commun.* **2020**, *11* (1), 904.

(7) Hagino, K.; Ichihashi, N. In Vitro Transcription/Translation-Coupled DNA Replication through Partial Regeneration of 20 Aminoacyl-TRNA Synthetases. *ACS Synth. Biol.* **2023**, *12* (4), 1252−1263.

(8) Blanken, D.; Foschepoth, D.; Serrão, A. C.; Danelon, C. Genetically Controlled Membrane Synthesis in Liposomes. *Nat. Commun.* **2020**, *11* (1), 4317.

(9) Scott, A.; Noga, M. J.; De Graaf, P.; Westerlaken, I.; Yildirim, E.; Danelon, C. Cell-Free Phospholipid Biosynthesis by Gene-Encoded Enzymes Reconstituted in Liposomes. *PLoS One* **2016**, *11* (10), No. e0163058.

(10) Godino, E.; Danelon, C. Gene-Directed FtsZ Ring Assembly Generates Constricted Liposomes with Stable Membrane Necks. *Adv. Biol.* **2023**, *7*, 2200172.

(11) Godino, E.; López, J. N.; Zarguit, I.; Doerr, A.; Jimenez, M.; Rivas, G.; Danelon, C. Cell-Free Biogenesis of Bacterial Division Proto-Rings That Can. Constrict Liposomes. *Commun. Biol.* **2020**, *3* (1), 539.

(12) De Franceschi, N.; Barth, R.; Meindlhumer, S.; Fragasso, A.; Dekker, C. Dynamin A as a One-Component Division Machinery for Synthetic Cells. *Nat. Nanotechnol.* **2024**, *19*, 70−76.

(13) Garenne, D.; Haines, M. C.; Romantseva, E. F.; Freemont, P.; Strychalski, E. A.; Noireaux, V. Cell-Free Gene Expression. In *Nature Reviews Methods Primers*; Springer Nature, December 1, 2021.
(14) Marshall, R.; Noireaux, V. Quantitative Modeling of Transcription and Translation of an All-E. Coli Cell-Free System. *Sci. Rep.* **2019**, *9* (1), 11980.
(15) Garamella, J.; Garenne, D.; Noireaux, V. TXTL-Based Approach to Synthetic Cells. In *Methods in Enzymology*; Academic Press Inc., 2019; Vol. *617*, pp 217−239.
(16) Sun, Z. Z.; Hayes, C. A.; Shin, J.; Caschera, F.; Murray, R. M.; Noireaux, V. Protocols for Implementing an Escherichia Coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Visualized Exp.* **2013**, *79*, 1−14.
(17) Bartelds, M. W.; García-Blay, Ó.; Verhagen, P. G. A.; Wubbolts, E. J.; van Sluijs, B.; Heus, H. A.; de Greef, T. F. A.; Huck, W. T. S.; Hansen, M. M. K. Noise Minimization in Cell-Free Gene Expression. *ACS Synth. Biol.* **2023**, *12*, 2217−2225.
(18) Sakai, A.; Jonker, A. J.; Nelissen, F. H. T.; Kalb, E. M.; van Sluijs, B.; Heus, H. A.; Adamala, K. P.; Glass, J. I.; Huck, W. T. S. Cell-Free Expression System Derived from a Near-Minimal Synthetic Bacterium. *ACS Synth. Biol.* **2023**, *12*, 1616−1623.
(19) van Sluijs, B.; Maas, R. J. M.; van der Linden, A. J.; de Greef, T. F. A.; Huck, W. T. S. A Microfluidic Optimal Experimental Design Platform for Forward Design of Cell-Free Genetic Networks. *Nat. Commun.* **2022**, *13* (1), 3626.
(20) Rustad, M.; Eastlund, A.; Jardine, P.; Noireaux, V. Cell-Free TXTL Synthesis of Infectious Bacteriophage T4 in a Single Test Tube Reaction. *Synth. Biol.* **2018**, *3* (1), ysy002.
(21) Forster, A. C.; Church, G. M. Towards Synthesis of a Minimal Cell. In *Molecular Systems Biology*; Nature Publishing Group, 2006.
(22) Pelletier, J. F.; Sun, L.; Wise, K. S.; Assad-Garcia, N.; Karas, B. J.; Deerinck, T. J.; Ellisman, M. H.; Mershin, A.; Gershenfeld, N.; Chuang, R. Y.; Glass, J. I.; Strychalski, E. A. Genetic Requirements for Cell Division in a Genomically Minimal Cell. *Cell* **2021**, *184* (9), 2430−2440.e16.
(23) Hutchison, C. A.; Chuang, R. Y.; Noskov, V. N.; Assad-Garcia, N.; Deerinck, T. J.; Ellisman, M. H.; Gill, J.; Kannan, K.; Karas, B. J.; Ma, L.; Pelletier, J. F.; Qi, Z. Q.; Richter, R. A.; Strychalski, E. A.; Sun, L.; Suzuki, Y.; Tsvetanova, B.; Wise, K. S.; Smith, H. O.; Glass, J. I.; Merryman, C.; Gibson, D. G.; Venter, J. C. Design and Synthesis of a Minimal Bacterial Genome. *Science* **2016**, *351* (6280), aad6253.
(24) Verma, S. C.; Qian, Z.; Adhya, S. L. Architecture of the Escherichia Coli Nucleoid. *PLoS Genet.* **2019**, *15* (12), No. e1008456.
(25) Deyama, T.; Matsui, Y.; Chadani, Y.; Sekine, Y.; Doi, N.; Fujiwara, K. Transcription-Translation of the Escherichia Coli Genome within Artificial Cells. *Chem. Commun.* **2021**, *57* (80), 10367−10370.
(26) Fujiwara, K.; Sawamura, T.; Niwa, T.; Deyama, T.; Nomura, S. I. M.; Taguchi, H.; Doi, N. In Vitro Transcription-Translation Using Bacterial Genome as a Template to Reconstitute Intracellular Profile. *Nucleic Acids Res.* **2017**, *45* (19), 11449−11458.
(27) Borkowski, O.; Koch, M.; Zettor, A.; Pandi, A.; Batista, A. C.; Soudier, P.; Faulon, J. L. Large Scale Active-Learning-Guided Exploration for in Vitro Protein Production Optimization. *Nat. Commun.* **2020**, *11* (1), 1872.
(28) Pandi, A.; Diehl, C.; Kharrazi, A. Y.; Faure, L.; Scholz, S. A.; Nattermann, M.; Adam, D.; Chapin, N.; Foroughijabbari, Y.; Moritz, C.; Paczia, N.; Cortina, N. S.; et al. A Versatile Active Learning Workflow for Optimization of Genetic and Metabolic Networks. *Nat. Commun.* **2022**, *13*, 3876.
(29) Banks, A. M.; Whitfield, C. J.; Brown, S. R.; Fulton, D. A.; Goodchild, S. A.; Grant, C.; Love, J.; Lendrem, D. W.; Fieldsend, J. E.; Howard, T. P. Key Reaction Components Affect the Kinetics and Performance Robustness of Cell-Free Protein Synthesis Reactions. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 218−229.
(30) Forquet, R.; Pineau, M.; Nasser, W.; Reverchon, S.; Meyer, S. Role of the Discriminator Sequence in the Supercoiling Sensitivity of Bacterial Promoters. *mSystems* **2021**, *6* (4), No. e0097821.
(31) Dame, R. T.; Rashid, F. Z. M.; Grainger, D. C. Chromosome Organization in Bacteria: Mechanistic Insights into Genome Structure and Function. In *Nature Reviews Genetics*; Nature Research, 2020; pp 227−242.
(32) Yamagami, R.; Sieg, J. P.; Bevilacqua, P. C. Functional Roles of Chelated Magnesium Ions in RNA Folding and Function. *Biochemistry* **2021**, *60* (31), 2374−2386.
(33) Kartje, Z. J.; Janis, H. I.; Mukhopadhyay, S.; Gagnon, K. T. Revisiting T7 RNA Polymerase Transcription in Vitro with the Broccoli RNA Aptamer as a Simplified Real-Time Fluorescent Reporter. *J. Biol. Chem.* **2021**, *296*, 100175.
(34) Rozov, A.; Khusainov, I.; El Omari, K.; Duman, R.; Mykhaylyk, V.; Yusupov, M.; Westhof, E.; Wagner, A.; Yusupova, G. Importance of Potassium Ions for Ribosome Structure and Function Revealed by Long-Wavelength X-Ray Diffraction. *Nat. Commun.* **2019**, *10* (1), 2519.
(35) Lykke-Andersen, J. The C-Terminal Carboxy Group of T7 RNA Polymerase Ensures Efficient Magnesium Ion-Dependent Catalysis. *Nucleic Acids Res.* **1998**, *26* (24), 5630−5635.
(36) Fuchs, E.; Millette, R. L.; Zillig, W.; Walter, G. Influence of Salts on RNA Synthesis by DNA-Dependent RNA-Polymerase from *Escherichia Coli*. *Eur. J. Biochem.* **1967**, *3* (2), 183−193.
(37) Nierhaus, K. H. Mg^{2+}, K^{+}, and the Ribosome. *J. Bacteriol.* **2014**, *196* (22), 3817−3819.
(38) Guo, X.; Zhu, Y.; Bai, L.; Yang, D. The Protection Role of Magnesium Ions on Coupled Transcription and Translation in Lyophilized Cell-Free System. *ACS Synth. Biol.* **2020**, *9* (4), 856−863.
(39) Malagodapathiranage, K.; Cavac, E.; Chen, T. H.; Roy, B.; Martin, C. T. High-Salt Transcription from Enzymatically Gapped Promoters Nets Higher Yields and Purity of Transcribed RNAs. *Nucleic Acids Res.* **2023**, *51* (6), No. E36.
(40) Thomen, P.; Lopez, P. J.; Bockelmann, U.; Guillerez, J.; Dreyfus, M.; Heslot, F. T7 RNA Polymerase Studied by Force Measurements Varying Cofactor Concentration. *Biophys. J.* **2008**, *95* (5), 2423−2433.
(41) Seo, K.; Ichihashi, N. Investigation of Compatibility between DNA Replication, Transcription, and Translation for in Vitro Central Dogma. *ACS Synth. Biol.* **2023**, *12*, 1813−1822.
(42) Milo, R.; Phillips, R. *Cell Biology by the Numbers*; Garland Science, 2015.
(43) 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* (4), 048104.
(44) De Capitani, J.; Mutschler, H. The Long Road to a Synthetic Self-Replicating Central Dogma. *Biochemistry* **2023**, *62*, 1221−1232.
(45) Jiang, Y.; Chen, B.; Duan, C.; Sun, B.; Yang, J.; Yang, S. Multigene Editing in the Escherichia Coli Genome via the CRISPR-Cas9 System. *Appl. Environ. Microbiol.* **2015**, *81* (7), 2506−2514.
(46) Caschera, F.; Noireaux, V. Synthesis of 2.3 Mg/Ml of Protein with an All Escherichia Coli Cell-Free Transcription-Translation System. *Biochimie* **2014**, *99* (1), 162−168.
(47) Holub, M.; Birnie, A.; Japaridze, A.; Van Der Torre, J.; den Ridder, M.; de Ram, C.; Pabst, M.; Dekker, C. Extracting and Characterizing Protein-Free Megabasepair DNA for in Vitro Experiments. *bioRxiv* **2022**, *06*, 497140.
(48) Pelletier, J.; Halvorsen, K.; Ha, B.-Y.; Paparcone, R.; Sandler, S. J.; Woldringh, C. L.; Wong, W. P.; Jun, S. Physical Manipulation of the Escherichia Coli Chromosome Reveals Its Soft Nature. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, No. E2649.
(49) Pelletier, J.; Jun, S. Isolation and Characterization of Bacterial Nucleoids in Microfluidic Devices. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. *1624*, pp 311−322.
(50) Cunha, S.; Odijk, T.; Süleymanoglu, E.; Woldringh, C. L. Isolation of the Escherichia Coli Nucleoid. *Biochimie* **2001**, *83*, 149−154.
(51) Lartigue, C.; Glass, J. I.; Alperovich, N.; Pieper, R.; Parmar, P. P.; Hutchison, C. A.; Smith, H. O.; Venter, J. C. Genome Transplantation in Bacteria: Changing One Species to Another. *Science* **2007**, *317* (5838), 632−638.

https://doi.org/10.1021/acsomega.4c00111
ACS Omega 2024, 9, 19227−19235