12-2018

Spatial Synthetic Cell-Free Biology

Sarah Elizabeth Norred
University of Tennessee

To the Graduate Council:

I am submitting herewith a dissertation written by Sarah Elizabeth Norred entitled "Spatial Synthetic Cell-Free Biology." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Energy Science and Engineering.

Michael L. Simpson, Major Professor

We have read this dissertation and recommend its acceptance:

Steven M. Abel, Eric T. Boder, C. Patrick Collier, Mitchel J. Doktycz

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Spatial Synthetic Cell-Free Biology

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Sarah Elizabeth Norred
December 2018

DEDICATION

This work is dedicated to my family.
Their love and dedication made my education possible.

And to my husband, Jerel.
Who makes sure I don’t waste their efforts.

ACKNOWLEDGEMENTS

I would primarily like to express my gratitude for the guidance and mentorship of my advisor, Dr. Michael Simpson. I would like to thank Dr. Mitch Doktycz and Dr. Pat Collier for their teaching and advice, both in the lab and out, and for serving on my thesis committee. I'd like to thank Dr. Steve Abel for his insightful contributions to computational models presented here. I am grateful for both his and Dr. Eric Boder's teaching in the classroom and their willingness to serve on my thesis committee. I would also like to acknowledge Dr. Scott Retterer, Dr. Jennifer Morrell-Falvey, Dr. Amber Bible, Dr. David Karig, and Dr. Sukanya Iyer for helpful conversations, criticisms, training, and essential expertise for my experimental work.

I would like to acknowledge and thank my fellow graduate students for their advice and support, especially Patrick Caveney, Dr. Charles Chin, Peter Shankles, and Gaurav Chauhan. I would also like to acknowledge Dr. Brandon Razooky and Dr. Jonathan Borekyo for their contributions and advice included in this thesis. I am immensely grateful for the work and dedication of Lauren Collier and Rosemary Dabbs, who performed many of the experiments included here during their internships.

I would like to thank all the members of the Biological Nanoscale Systems Group, especially Carmen Foster and Cathy Gaudreau. I would also like to thank all the members of the Nanofabrication Research Laboratory, especially Jessica Garner and Lisa Goins. I would also like to thank the staff of the Bredesen Center, including and especially Wanda Davis and Dr. Lee Riedinger.

I would like to acknowledge the financial support of the Bredesen Center for Interdisciplinary Research and Education and the Center for Nanophase Materials Science.

ABSTRACT

The U.S. has the biomass production potential to dramatically offset yearly petroleum consumption, but many efficiency barriers remain for developing enduring bioenergy sources. Synthetic biology allows researchers to redesign energy-relevant organisms to increase the efficiency and lower the cost of bioenergy technologies. However, developing complex gene circuit behavior in new organisms or networks can result in unexpected complications and off-target effects. Since cellular structure and scale can affect gene expression dynamics, understanding how gene expression operates within the physiological context of the cell becomes important for developing robust gene circuits. Gene expression occurs in a highly crowded and confined (from about 1 fL to several pL) environment. Macromolecules occupy 5-40% of the intracellular environment, effecting changes in molecular transport, association, and reaction rates associated with gene expression. Gene expression also exhibits “bursty” patterns of expression, characterized by episodic periods of high activity between periods of low activity. These bursting patterns are shaped not only by molecular mechanisms but also by the global availability of resources within the expression environment, both of which may be further modulated by physical effects, like crowding and confinement. Since manipulating the physical conditions surrounding gene expression can be difficult to achieve in cells, cell-free systems are used to directly probe gene expression reactions. In this work, gene expression reactions in cell-free systems are modified to mimic physiological levels of crowding and confinement, revealing information about the interplay between expression bursting, resource sharing, and spatial ordering in transcription and translation. These results explore how confined reactions alter bursting patterns and distribute limited expression resources, as well as how crowding-induced spatial inhomogeneities in transcription can affect bursting patterns in translation. The cell-free platform described here also demonstrates spatial organization of gene expression similar to that seen in cells, providing a useful technique for exploring the mechanisms of cellular self-organization in gene expression and developing spatial control over transcription and translation reactions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Bioenergy's Future in Synthetic Biology

World energy needs are increasing dramatically. It is estimated between 2012 and 2040, world energy demand will increase 48%.[1] Over 80% of primary power in 2015 came from fossil fuel sources,[2,3] which are implicated in concerns about climate change, global energy security, and numerous other social and economic costs. Driven by the need to offset increase in expected energy consumption and limit the negative impacts of fossil fuels, renewables now account for about 10% of energy consumption in the United States.[4] Though large-scale fossil fuel consumption in the US may seem inevitable due to the availability and established infrastructure, the US has an untapped energy resource that may compete with this. The Department of Energy's 2016 Billion ton report concludes that the US has the potential to produce a billion tons of dry biomass on an annual basis— which would offset about 30% of 2005's yearly petroleum consumption without affecting other agricultural production.[5]

As of 2016, the largest single source of renewable energy is biomass, which accounts for about half of all renewable energy consumption.[3,4] Biomass includes agricultural and forestry resources (e.g. energy crops or logging residues), municipal solid waste, and more recently, algae. The use of biomass feedstock in the United States is commonly associated with ethanol and biodiesel production, largely due to policy incentives and mandates such as the 2005 Renewable Fuel Standard, but biomass resources can also be used for other renewable bioproducts, such as biogas or value-added chemical compounds. In the biorefinery, microorganisms are often used to ferment sugar intermediates into fuel blendstocks and valuable chemicals. The most common current ethanol production processes use yeast fermentation of maize feedstock. This method uses the natural homo-ethanol pathways found in *Saccharomyces cerevisiae* or other microbes[6], however this method is not necessarily the most scalable or efficient commodity source. Though harnessing the native functions of biological resources is useful, a redesign of the organisms that contribute to biofuel production could significantly increase the efficiency and lower the cost of bioenergy applications.[7-9]

Synthetic biology uses engineering design principles, like rapid prototyping and testing, to manipulate the function of energy-relevant organisms. Using tools such as standardized gene "parts" and advanced DNA synthesis and sequencing methods, specific cellular functions and pathways can be repurposed or altered outside of their native function. In the context of bioenergy, the advent of synthetic biology has allowed for manipulation of metabolic pathways for bioproduct formation and offers a powerful method for reducing the expected cost of renewable fuels and chemicals. Driven mainly from the desire to alter chemical pathways or the release of lignocellulosic sugars[6], possible synthetic biology solutions range from engineering plants to produce more accessible carbohydrates to using protein engineering to develop more effective digestive enzymes.[7] Genetic parts and chemical processes can similarly be reassigned to more tractable organisms, for example, *E. coli* was manipulated to deliver homo-ethanol production by integrating the production of ethanol (PET) pathway from *Zymomonas mobilis* into the host chromosome.[6] *E. coli* has also been manipulated to perform complex intracellular reactions like in-vivo transesterification, which may reduce the impact of costly or inefficient chemical synthesis.[10] Cells can also be modified to be very specific utilizers of feedstocks or be manipulated to directly produce complex high-value industrial chemicals[11], materials[12], or pharmaceuticals[13]. Synthetic biology tools allow for the potential to finely tune energetic constraints and pathway manipulation—all of which are essential to producing cost-competitive yields.

Bioproduct formation through manipulating existing parts and pathways, however, only captures one aspect of the transformative potential for synthetic biology in energy science. The ability to manipulate organisms via genetic circuit design introduces advanced aspects of computation and decision-making into the cell.[14-16] Regulation of gene expression can provide more complex operational tools for the cell, such as selectively turning on expression in a specific sequence or under certain conditions. Though gene regulation has been used for improving chemical production, more advanced systems could control more advanced decision-making behaviors, including complex modules derived from control theory.[14] Robust and tightly controlled circuits are necessary for the "second wave" of synthetic biology control—movement into organisms, such as bioenergy crops, or involvement in intricate microbial communities.[7,8]

1.2 Gene Circuit Design in Cellular Context

Robust gene circuits that operate effectively in multiple contexts can be difficult and time-consuming to develop. Gene circuit development is not limited simply by the amount of available parts for circuit design. Inserting even well-designed genes into new organisms or networks can result in unexpected complications and off-target effects—the desired effect of the circuit may be diminished, altered, or absent entirely if not designed with the molecular context of the host cell in mind[14]. Problems related to the operation of a gene circuit in the host cell might include host overload, molecule queueing, and retroactivity[14], but since cellular structure and scale can affect gene expression dynamics, understanding how gene expression operates within the physical context of the cell is important for developing robust gene circuits. In order to further elucidate the impact of host systems on gene circuits and networks, further analysis must be done to characterize these circuits in the physical context of the systems in which they will operate.

At the most basic level, cellular context is defined by the physical constraints of confinement and macromolecular crowding. Cell volumes are measured on the picoliter[17] and femtoliter[18] scale. Within those confined volumes, cellular reactions are often compartmentalized or sequestered into distinct subregions for specialized purposes, primarily seen in eukaryotic organelles or bacterial microcompartments[19,20]. Confinement of reactions can not only improve molecular localization and association, but also provide a more favorable chemical environment for a reaction or process, isolated from other processes in the cell. The individuation of different physical subregions and conditions adds to the physical heterogeneity in the cell interior, where 5-40% of the total volume is crowded with proteins, nucleic acids, and other large molecules[21]. A single *E. coli* is approximately one femtoliter in volume, and between 30-40% of this volume is occupied by macromolecules.[22] Macromolecular crowding, which may be thought of as nonspecific steric exclusion, can dramatically affect reaction dynamics by altering the diffusion, orientation, and effective concentration of relevant species.[23] The effects of macromolecular crowding are sensitive to the size and shape of the constituent crowding molecules, as well as the distribution of molecules within the confined space[21,24]. Both confinement and macromolecular crowding together can have a dramatic effect on complex reactions in cells, like gene expression, which require the

localization and coordination of many molecular components of many sizes in order to proceed.[22,25,26]

In order to characterize gene circuits in a physiological context, physical constraints like crowding and confinement must be characterized in terms of their impact on gene expression.[26] As local confinement or crowding can inhibit or alter the diffusion or positioning of molecules relevant to gene expression, the influences of physical molecular conditions and an individual gene's expression behavior are innately entwined. Macromolecular crowding may, for example, alter the positioning of constituent components in a reaction such that the components are increasingly co-localized and thus more likely to react. Conversely, crowding may also limit the mobility of reactants such that constituent components cannot freely interact, reducing effective reaction rates. The effects of both diffusion-limitation and increased molecular association can be seen in gene expression reactions in the re-initiation of reaction components like polymerases, for example, which can temporally alter expression rates through high periods of transcriptional activity.

1.3 Noise in Gene Expression

Gene expression is a noisy process that occurs in bursts characterized by periods of high and low (or no) activity. Gene expression bursting is a phenomenon observed across many different systems, from complex eukaryotes to viral expression. The noise caused by transcriptional bursting, in particular, is implicated in many decision making-processes, including cell differentiation[27,28], the *Bacillus subtillus* decision between competence and sporulation[27], and the HIV decision between active replication and proviral latency[29]. Bursting patterns are shaped not only by molecular mechanisms, like supercoiling[30], promoter architecture[31,32], or chromatin remodeling[33], but also by the global availability of resources within the expression environment[34]. Individual genes are expressed in the context of global gene expression from a finite pool of shared resources, so the expression behavior of individual genes is coupled to the behavior of all genes within a confined expression environment. Since expression bursts draw shared resources (like polymerases, cofactors, or ribosomes) in a time-variant way, even genes with no direct regulatory relationship are affected by the communal use of expression machinery. The use of resources among different genes is further modulated

by the molecular availability of gene expression components, altered in part by crowding and confinement.

The stochastic fluctuations in molecular populations related to noisy gene expression are an important feature of the cellular context. The molecules that regulate gene expression – from small molecule inducers to the polymerases and ribosomes – are present in small populations where the inherent fluctuations are large compared to mean population levels[35-41]. As described above, gene expression reactions occur in a highly crowded and confined environment, which may further modulate the availability of the reactants. These contributing factors to stochasticity in molecular populations have demonstrable consequences for gene regulation as fluctuations in the inputs or controlling mechanisms of gene circuits can alter the operation of a circuit.[26,35,42] The HIV decision between active replication and latency is a notable example of how stochastic fluctuations in molecular populations can affect the behavior of a gene circuit[29]. In this case, the selection between two possible fates is mediated by fluctuations in the population of the Tat protein. Examining the fluctuations in gene expression over time, or the "noise" of the circuit, reveals the temporal variations associated with bursting behavior and provides a way to analyze the expression patterns of the gene circuit.

The simplest way to model gene expression is the constitutive expression model, where mRNA is produced from a gene in a Poissonian process at rate α and translates in a Poissonian process at rate k_p to make protein, p (Figure 1.1A) such that

$$\langle \boldsymbol{m} \rangle = \frac{\boldsymbol{\alpha}}{\boldsymbol{\gamma_r}} \tag{1.1}$$

$$\langle \boldsymbol{p} \rangle = \frac{\boldsymbol{\alpha k_p}}{\boldsymbol{\gamma_r \gamma_p}} = \frac{\boldsymbol{\alpha b}}{\boldsymbol{\gamma_p}}, \tag{1.2}$$

where rates γ_r and γ_p describe exponentially distributed decay of the mRNA and protein populations and the term b is the translational burst size (average number of protein molecules made during the lifetime of an individual mRNA).

As described above, transcription typically occurs in bursts, and a more realistic model of gene expression involves transcription switching between active and inactive states. This expression motif is typically analyzed using the Random Telegraph, or "two-state", model. In this model, the gene switches from OFF to ON at rate k_{ON} and from ON to OFF at rate k_{OFF} (Figure 1.1B). This expression motif results in mRNA being produced

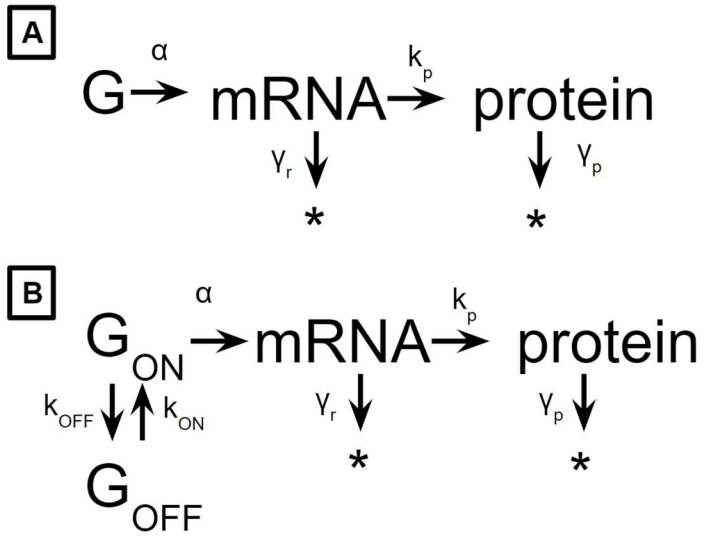


Figure 1.1: Models of Gene Expression A) Constitutive gene expression model B) Random telegraph or "two-state" model

in bursts, with the transcriptional burst size (B) defined as the average number of mRNA molecules made during the lifetime of an ON state. These bursts occur at a burst frequency $f_B = 1/(1/k_{ON} + 1/k_{OFF})$. The constitutive model is a special case of this model, where transcription occurs at B=1 at frequency of $f_B = \alpha$. The overall burst size (B_S) – which includes both transcriptional and translational bursts – is simply the product of the two individual burst sizes (B*b). For the two-state model:

$$\langle \boldsymbol{m} \rangle = \frac{\boldsymbol{B f_B}}{\boldsymbol{\gamma_r}} \quad 1.3$$

$$\langle \boldsymbol{p} \rangle = \frac{\boldsymbol{B b f_B}}{\boldsymbol{\gamma_p}}. \quad 1.4$$

By measuring protein noise in gene expression reactions over time, the burst size and burst frequency may be estimated. Expression data may be used to find the variance, mean, and other information about the expression trace in order to characterize the expression noise. The variances for mRNA and protein in the two-state model are found by[43]:

$$\boldsymbol{\sigma_m^2} = (\boldsymbol{B}+\boldsymbol{1})\langle \boldsymbol{m} \rangle \approx \boldsymbol{B}\langle \boldsymbol{m} \rangle \quad 1.5$$

$$\boldsymbol{\sigma_p^2} = (\boldsymbol{B}+\boldsymbol{1})(\boldsymbol{b}+\boldsymbol{1})\langle \boldsymbol{p} \rangle \approx \boldsymbol{B b}\langle \boldsymbol{p} \rangle \quad 1.6$$

Noise magnitude may be represented using the coefficient of variation squared, CV^2, or the variance of the noise divided by the square of the mean ($CV^2=\sigma^2/\mu^2$). Since

$$CV_m^2 = \frac{\sigma_m^2}{\langle m \rangle^2} \quad 1.7$$

$$CV_p^2 = \frac{\sigma_p^2}{\langle p \rangle^2} \approx \frac{\gamma_p}{f_B}, \quad 1.8$$

plotting the CV^2 against abundance, or mean expression, reveals distinctions between how burst size and burst frequency change between experimental conditions. While abundance (of either mRNA or protein) may change because of a change in burst size or burst frequency, CV^2 changes only with burst frequency. Examining how CV^2 changes with respect to abundance in this space reveals how burst size and burst frequency changes between experiments. For example, horizontal movement of a point in this space, or a large change in abundance without a corresponding large change in CV^2, indicates that burst size, but not burst frequency, is changing. In contrast, a CV^2 inversely proportional to abundance indicates a change in burst frequency but little change in burst size (Figure 1.2) Thus plotting expression noise in this space allows the extraction of burst parameters, as well as the ability to analyze how these parameters change for different experimental conditions.

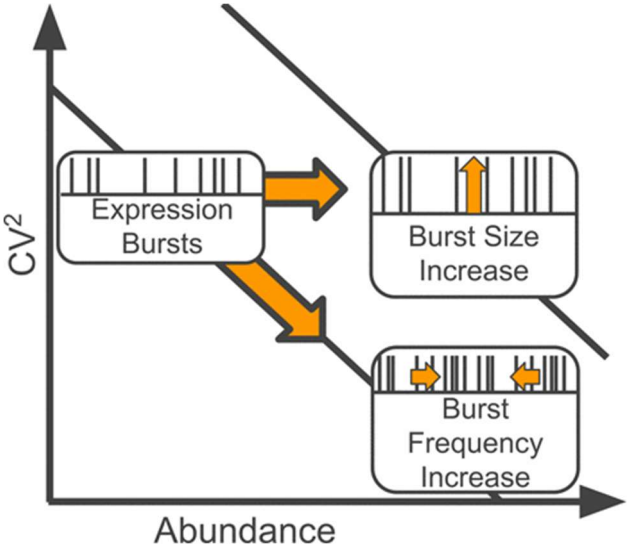


Figure 1.2: CV^2 vs Abundance Plot for Noise Analysis. Changes in bursting parameters can be understood by observing changes in positioning in this space.

1.4 Cell-free Expression

Though gene expression noise has traditionally been investigated in cellular systems, cell-free platforms are increasingly being used as a platform for investigating questions in noise biology.[34,44,45] Made from either purified cell extracts or reconstituted proteins, cell-free systems provide a flexible platform for isolating phenomena contributing to noise in gene circuits. In an effort to diagnose sources of failure or maladaptation in gene circuits, there has been considerable interest in developing cell-free methods for "debugging" problems in gene circuits.[15,46-49] These efforts include the development of minimal cell technology or in cell-free "breadboard" extracts, methods whereby the confounding interactions of a target host cell are minimized or eliminated. These methods concentrate all resources within the system to the operation of the genetic circuit, in theory allowing the researcher to isolate problems innate to the circuit's function.[49] However, even within these simplified systems there exist regulatory problems, such as molecular bottlenecks and crosstalk, which can change the topology of the circuit.[14,15,50] Moreover, transcriptional bursting patterns may be observed in a cell-free context, absent of the usual cell-attributed causes of chromatin remodeling or supercoiling.[30,34,44,51,52] These findings highlight the importance of examining cell-free gene expression and expression noise within the context of physiological conditions, especially since many physical aspects of the intracellular environment, such as confinement and macromolecular crowding, can have large roles in reaction dynamics.[23,26]

Cell-free systems allow the probing of physical and chemical conditions which may be difficult, or even impossible, to rigorously alter in a cellular system. Physiological confinement has been mimicked in cell-free gene expression by using microfluidic structures[44,45,53,54], liposomes[51,55-59], water-in-oil droplets[52,60], and hydrogel structures[61,62]. Many of these confinement methods provide a viable way to acquire timescale information about gene expression within individual reactions— providing insights to the relationship between confinement and gene expression rates[63] but also providing necessary data for noise analysis. Cell-free platforms also provide a method for examining physiological macromolecular crowding conditions. Though cell-free expression systems alone tend to be very dilute in comparison to cells[9,64], artificial molecular crowders, such as PEG, Dextran, or Ficoll[24,65] may be added directly to cell-

free expression systems. Crowding has dramatic effects on gene expression reaction rates and yield in microliter-volume “batch” reactions[65]. However, examining physiological levels of crowding on gene expression in addition to confinement can produce physical effects not seen with crowding and confinement alone[52,61,66-68].

1.5 Scope of Dissertation

The aim of this thesis is to characterize the response of gene expression bursting to spatial constraints, like macromolecular crowding and confinement. This work uses Cell-free Protein Synthesis (CFPS) gene expression systems to probe these conditions experimentally. Confinement is examined using both microfluidic and lipid vesicle encapsulation methods. Macromolecular crowding is examined by supplementing gene expression reactions with Ficoll-70. Both transcriptional bursting parameters and total bursting parameters are examined simultaneously by using a gene construct allowing fluorescence tracking of both mRNA and protein signals. The experimental results shown here are supported by computational simulations modelling gene expression and spatial localization of nucleic acids in confined spaces. Finally, confocal microscopy is used to visualize both crowding and confined reactions in lipid vesicles, revealing spatial heterogeneity in the distribution of mRNA over the volume of the vesicle. This spatial noise affects the temporal noise of the protein expression and is reflective of heterogeneity in cellular gene expression influenced by macromolecular crowding. This work explores the individual and combined effects of crowding and confinement on cell free reactions and discusses the self-organization of cell-free gene expression reactions in comparison to membrane-less spatial organization in cells.

1.6 Organization of Dissertation

Chapter 2 is adapted from the article “Sealable Femtoliter Chamber Arrays for Cell-free Biology” originally published in the *Journal of Visualized Experiments* in 2015. This article is focused on the experimental methods for encapsulating cell-free reactants in microfluidic chambers. Chapter 3 is adapted from the article “Resource Sharing Controls Gene Expression Bursting” originally published in *ACS Synthetic Biology* in 2016. The results of this study are derived from the methods described in Chapter 2. This chapter describes the effect of confinement on cell-free gene expression, demonstrating how

resource use affects bursting patterns. Chapter 4 is adapted from the article "Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns" originally published in *ACS Synthetic Biology* in 2018. This chapter describes the effects of macromolecular crowding on bursting patterns in unconfined gene expression reactions. This chapter also introduces the simultaneous measurement of mRNA and protein, and reveals that transcription and translation have divergent bursting patterns with increased crowding. Chapter 5 is adapted from the article "Synergistic Interactions Between Confinement and Macromolecular Crowding Spatially Order Transcription and Translation in Cell-Free Expression" which is available on *bioRxiv*. This chapter examines both confinement and crowding of cell-free gene expression reactions and reveals how these physical effects control the spatial organization of gene expression reactions in a cell-free context. Where available, the Supplementary Information associated with each article is included in each respective chapter Appendix. References for all chapters are listed all together at the end of the document.

2 SEALABLE FEMTOLITER CHAMBER ARRAYS FOR CELL-FREE BIOLOGY

A version of this chapter was originally published by S. Elizabeth Norred, Patrick M. Caveney, Scott T. Retterer, Jonathan B. Borekyo, Jason D. Fowlkes, C. Patrick Collier, and Michael L Simpson:

Norred, S. E., Caveney, P. M., Retterer, S. T., Boreyko, J. B., Fowlkes, J. D., Collier, C. P., Simpson, M. L. Sealable Femtoliter Chamber Arrays for Cell-free Biology. *J. Vis. Exp.* (97), e52616, doi:10.3791/52616 (2015).

This chapter has been adapted from its published format to accommodate new Figure, Table, and Equation enumeration. All references are located at the end of the document. The article presented in this chapter is methods-based, and includes an instructional protocol in a listed format. A video presentation of this article, featuring practical demonstrations of the methods used, was produced by the *Journal of Visualized Experiments* and is available open-access at the following URL: https://www.jove.com/video/52616/sealable-femtoliter-chamber-arrays-for-cell-free-biology

SEN, PMC, STR, JBB, JDF, CPC, and MLS conceived and planned the experiments. Experiments were adapted for cell-free expression and performed by SEN, PMC, JBB, and CPC. The microfluidic device was originally designed by CPC and JDF for use in a previous publication[69]. SEN, PMC, and CPC performed image analysis. All authors contributed to data analysis and the final manuscript. The authors acknowledge and thank Dr. Sukanya Iyer for constructing the Pet3a-EGFP plasmid used in these experiments.

2.1 Abstract

Cell-free systems provide a flexible platform for probing specific networks of biological reactions isolated from the complex resource sharing (*e.g.,* global gene expression, cell division) encountered within living cells. However, such systems, used in conventional macro-scale bulk reactors, often fail to exhibit the dynamic behaviors and efficiencies characteristic of their living micro-scale counterparts. Understanding the impact of internal cell structure and scale on reaction dynamics is crucial to understanding complex gene networks. Here we report a microfabricated device that confines cell-free reactions in cellular scale volumes while allowing flexible

characterization of the enclosed molecular system. This multilayered poly(dimethylsiloxane) (PDMS) device contains femtoliter-scale reaction chambers on an elastomeric membrane which can be actuated (open and closed). When actuated, the chambers confine Cell-Free Protein Synthesis (CFPS) reactions expressing a fluorescent protein, allowing for the visualization of the reaction kinetics over time using time-lapse fluorescent microscopy. Here we demonstrate how this device may be used to measure the noise structure of CFPS reactions in a manner that is directly analogous to those used to characterize cellular systems, thereby enabling the use of noise biology techniques used in cellular systems to characterize CFPS gene circuits and their interactions with the cell-free environment.

2.2 Introduction

Cell-free systems offer a simplified and flexible platform for viewing biological reactions free from complicating factors such as fitness, division, and mutation that are unavoidable in the study of living cells. Such approaches have been employed to study cellular systems including the characterization of membrane proteins[70], the probing of protein interactions[71], and the exploration of fundamental aspects of translation[47,72-75]. Recently cell-free systems have begun to gain a foothold as viable platforms for synthetic biology[44,76,77]. The appeal of such approaches is that they free synthetic biology from the resource sharing and 'extrinsic noise' that affects reaction dynamics in living cells. However, questions remain as to how the physical environment in which cell-free reactions are embedded affects the progression and outcome of the reaction. Cell-free reaction environments — particularly confined environments that approach cell-relevant volumes — remain poorly characterized. Cell-Free Protein Synthesis (CFPS) is conventionally thought of as being 'scale-free,' exhibiting equivalent kinetics across a range of microliter to liter-scale reaction volumes[78]. Nonetheless, confining reactions to cellular scale volumes has been shown to significantly affect protein expression rates[63].

The stochastic nature of cell-free reactions — especially as these systems approach or even go below femtoliter volumes — may be of particular importance. Noise in gene expression is a property greatly influenced by confinement as small cell volumes and high densities of components force many of the important molecules to very low population levels — for example, *Escherichia coli* confines within a 1 fL volume as many

as 4,300 different polypeptides under the inducible control of several hundred different promoters[79]. This inherent noise has been implicated as a central driving force in numerous biological processes including chemotaxis[80], the HIV decision between active replication and latency[81], the λ phage decision between lysis and lysogeny[27,82], and the *Bacillus subtillus* decision between competence and sporulation[27]. Cell-free synthetic biology then provides both an opportunity to explore the stochastic properties of cellular gene circuits and networks, and manipulate these behaviors to achieve specific technological goals. While the noise behavior of cellular systems has been well-studied[29,35,37,40,83-88], there has been little exploration of the fundamental noise behavior of cell-free systems[44], particularly at the cellular scale.

Here we present a platform for the study of stochastic effects in cell-free synthetic biology. This microfabricated platform contains femtoliter-scale reaction chambers which may be quickly transitioned between open (free diffusion in and out of the chamber) and closed (reactants confined within the chamber) states. In the closed state, we confine Cell-Free Protein Synthesis (CFPS) reactants expressing a green fluorescent protein (GFP), and follow gene expression using time-lapse fluorescence microscopy (Figure 2.1). We characterize this cell-free environment by measuring the structure of the stochastic fluctuations in gene expression in a manner directly analogous to those used to characterize cells[35]. Non-microfabrication methods for confining cell-free reactions include vesicles and liposomes[55,56,89,90], water-in-oil emulsions[63], and porous media[91]. However, while these methods can provide control over the size distribution of the confined volumes[92], microfabrication methods create highly replicable features with tightly specified dimensions, even on the nanoscale. Moreover, these rigid structures can be easily tracked over time without being susceptible to evaporation or changes in the external environment. Microfabricated container designs used in previous work[44,53] cannot quickly seal the reaction chambers following reaction initiation, complicating the clear assignment of the time when the reaction was initiated (time zero). Using the method presented here, only 4-5 min are needed between initiation and visualization of the reaction on the device, thereby providing a well-defined “time zero”. The following protocols describe the methods for fabricating and testing this device, including optical lithography, device assembly, device testing, and methods for image analysis.

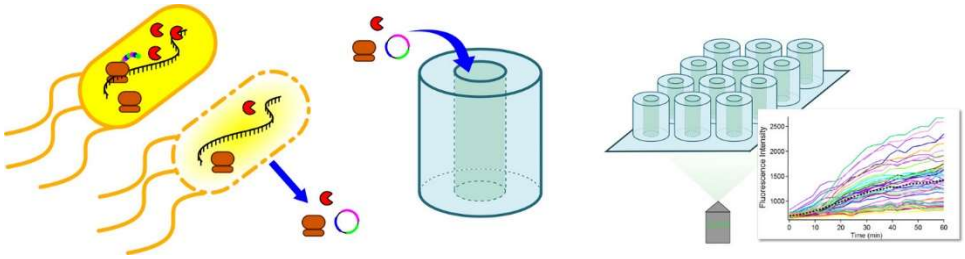


Figure 2.1 Cell-Free Protein Synthesis reactants are confined in femtoliter scale reaction chambers for the purpose of measuring gene expression. Reactants from a commercial cell-free protein expression system are used to constitutively express GFP inside confined PDMS reaction chambers. An array of these chambers may be visualized with time-lapse fluorescence microscopy in order to characterize protein expression and gene expression noise. The fluorescence intensity of each reaction chamber over time may be plotted as an individual trace.

2.3 Protocol

2.3.1 Optical Lithography of Device Masters

1. Dehydrate clean silicon wafers on a hot plate at ~250 °C for at least 1 hr. NOTE: It is good practice to use more than one wafer when preparing a master, in case of user error.
2. Prepare photoresist aliquots. Prepare aliquots of both SU-8 2015 photoresist and a dilution of SU-8 2015 photoresist in 2:1 ratio using SU-8 thinner as diluent. NOTE: Approximately 1 ml of photoresist is needed for spin-coating one wafer.
3. Prepare three mask patterns for producing these masters. For the Membrane Master, prepare two masks: one patterning the membrane channel and the other patterning the reaction chambers. For the Control Valve master, prepare only one mask pattern. NOTE: For more details on lithographic techniques, including mask patterning, see Ito and Okazaki, 2000[93]. See Fowlkes and Collier, 2013 for a more detailed description of device design[69].
4. Prepare the Membrane Master
    1. Spin-coat 2:1 SU-8 2015 photoresist dilution on wafers at 1,000 rpm for 45 sec.
    2. Soft bake wafers at 95 °C for 2 min. Using a contact aligner, expose wafers with membrane channel pattern for 10 sec, and perform a post-exposure bake for 2 min at 95 °C.

3. Develop wafers in SU-8 developer for 1 min, or until photoresist residue is removed. Rinse wafer with isopropanol, moving from top to bottom. Dry wafer with nitrogen, again moving from top to bottom. Bake wafers at 180 °C for 4 min.
4. Spin-coat patterned wafers again with 2:1 SU-8 dilution at 2,000 rpm for 45 sec.
5. Soft bake patterned wafers for 2 min at 95 °C. Using contact aligner, align patterned wafers with reaction chamber pattern, and expose for 10 sec. Perform post-exposure bake for 2 min at 95 °C.
6. Develop wafers as described in step 1.4.3. After developing and drying wafers, bake wafers at 180 °C for 4 min. NOTE: The wafers may be developed in the same developer that was used in the previous step.

5. Prepare the Control Valve Master
    1. Spin-coat undiluted SU-8 photoresist onto clean wafers at 2,000 rpm for 45 sec.
    2. Soft bake wafer at 95 °C for 6 min. Using a contact aligner, expose wafers with control valve pattern for 10 sec. Perform a post-exposure bake at 95 °C for 6 min.
    3. Develop wafers in SU-8 Developer for 2 min, or until residue is removed. Rinse with isopropanol, moving from top to bottom. Dry wafer with nitrogen and bake at 180 °C for 4 min.

2.3.2 PDMS Device Fabrication

1. Silanize all masters with ~0.2 ml trimethylchlorosilane via vapor deposition.
    1. Quickly enclose the master in an airtight container at RT with a few drops of the silanizing agent. NOTE: Other silanizing protocols may be acceptable[94]. If performed properly, the PDMS will be easy to remove.
2. Mix a commercial poly(dimethylsiloxane) (PDMS) base and curing agent in different ratios for both the membrane and control valve layers of the device, as has been demonstrated in similar multilayer valve designs[95]. Use 20:1 and 5:1 ratios of base:curing agent for the membrane and control valve molds, respectively.

    1. For the membrane mold, mix 10 g of base with 0.5 g of curing agent. NOTE: This volume will be spin-coated onto the membrane master.
    2. For the control valve mold, mix the base and curing agent in a 5:1 ratio. The amount of PDMS necessary to mold the control valve will depend on the container used to hold the control valve master; fill the container such that the master is coated with ~1 cm of PDMS.
3. Thoroughly mix both PDMS preparations, and de-gas them in a vacuum chamber until no air bubbles are visible. Place the control valve master in a heat-resistant container, such as a glass dish. Carefully pour 5:1 ratio PDMS over the master, and de-gas the container a second time.
4. While the control valve PDMS container is being de-gassed, spin-coat the 20:1 ratio PDMS on the membrane master by carefully pouring the PDMS mixture onto the membrane master to minimize air bubble formation, then spin-coating the master at 1,000 rpm for 45 sec.
5. Partially cure both masters in an oven at 80 °C for 6 min for the membrane master and 15 min for the control valve master. NOTE: When partially cured, the PDMS should hold its form, but the material will be slightly tacky. If PDMS is not yet cured, bake again in increments of a few minutes at a time until the material holds its form when pressed.
6. Cut rectangular PDMS molds from control valve master, peeling the molds away gently. Punch inlet holes through the molded component using a 0.75 mm hole punch. NOTE: The hole may be cleaned by inserting a 23 gauge blunt tip needle, and the mold exterior may be cleaned with cellophane tape, if necessary.
7. Using an optical microscope to locate the reaction chambers on the membrane master, align the control valve mold component with the features of the reaction chamber membrane and place the control valve component directly on top of the membrane master. Orient the control valve inlet to the bottom left corner of the device, and ensure that the reaction chambers and channel of the membrane master are visible inside the rectangular control valve.
8. Bake the aligned mold components at 80 °C for 2 hr. NOTE: The membrane and control valve molds will now be sealed together, and manipulated as one mold.

9. Cut the layered PDMS mold away from the membrane master, peeling the mold away from the master very gently so as not to perforate the membrane.
10. Punch inlet and outlet holes for the cell extract input using a 0.75 mm hole punch. Punch holes through both layers, and clean them in the same way as described in step 2.6.
11. Using an inductively-coupled plasma cleaner, plasma treat both the mold (membrane side up) and a No. 0 glass coverslip at 10.5 W for 20 sec. Immediately remove the coverslip and mold from the plasma cleaner and layer the components, membrane side towards the glass, attempting to minimize air pockets between the glass and the mold. Do not press directly on the membrane input channel, or the membrane may anneal to the glass, making it difficult to fill the channel with reactants.
    1. Take special care when handling the assembled devices to avoid breaking the glass layer. Use thin glass coverslips as the device must be imaged through the glass coverslip using high magnification oil-immersion objectives — if the glass is too thick, the device features may not be visible.
12. Finally, cure the completed devices at 80 °C for 2 hr.

2.3.3 Experimental Setup for Cell-free Protein Synthesis Reaction

1. Hydrate a device by boiling it in deionized water for 1 hr. NOTE: Device should have a cloudy appearance when completely hydrated. Device may also be left O/N in sterile water at RT in order to hydrate it.
2. Using an inverted microscope with an incubation chamber, set the ambient temperature to 30 °C. NOTE: This temperature was chosen to optimize expression of GFP with a T7 promoter, so optimal temperatures for other reactions may vary[49].
3. Mount device to microscope stage holder with cellophane tape and wrap edges of device with wet tissue paper in order to maintain local hydration.
4. Use two high precision closed-loop voltage-pressure transducers to modulate nitrogen gas pressure for control valve actuation and reagent input. NOTE: This

protocol has only been tested with low-purity nitrogen, though other inert gases may be used.

    1. Connect the first transducer by 24-gauge PTFE tubing to a water reservoir held in a 4 ml glass vial with a septum lid. Connect the reservoir to the control valve inlet using a second tube terminated by a 23 gauge blunt tip needle. NOTE: Both tubes penetrate the reservoir septum with two sharp 23 gauge needles.
    2. Connect the second transducer by 24-gauge PTFE tubing connected to a male-to-male Luer-lok connector. Attach this to a Luer-lok 23 gauge needle connected by tubing with another 23 gauge blunt tip needle, which is assembled individually for each device. This needle connects to the membrane reaction channel; use it to flush water from the reaction channel and input reagents.

5. Using a cell-free protein expression system, assemble the components for the CFPS reaction on ice, according to manufacturer's instructions. Minimize the time spent holding CFPS reagents on ice and place the reaction into the device immediately after assembly. NOTE: This device has been used with a commercial *E. coli* extract protein expression kit and a plasmid constitutively expressing GFP. The total reaction volume was scaled to 25 µl — it may be possible to use an even lower volume for reactants, if desired. As CFPS reagents tend to be sensitive to freeze-thaw cycles, it may be helpful to make aliquots of the reagents at the appropriate volume prior to the experiment. Other reagents may be added to the reaction mixture, but the reaction must be fully assembled before being applied to the device.
    1. Assemble the reaction, adding the DNA input last. NOTE: Once assembled in an Eppendorf tube, the CFPS reaction will begin if not held on ice. Since the time taken to apply the reagents to the device and begin the experiment may vary, it is helpful to start a timer once the reaction is assembled and mixed — this will keep the timescale between experiments consistent, and aid in troubleshooting.
6. Using the tubing and needle connector described in Step 3.4.2, withdraw the assembled reaction into the tube using a 1 ml syringe. Insert the blunt tip needle

into the reaction chamber inlet. Detach the needle connector from the syringe and attach it to the male-to-male connector used for the reaction chamber transducer.

7. Apply pressure (<10 psi) to the CFPS reactants to fill the channel. Remove the needle when the reaction is filled.
8. Insert the blunt tube from the other transducer into the control valve inlet. Do not pressurize the control valve yet.
9. Place the mounted device on the stage. Using brightfield imaging, locate the reaction chambers with a 100X oil-immersion objective.
10. Actuate the control valve by pressurizing the control valve transducer to 20 psi; a visible change in the membrane will be evident when the control valve is actuated. Focus on the bottoms of the reaction chambers.
11. Begin the image acquisition; growth in fluorescence will be visible in the interior and around the exterior of the reaction chambers, though it will likely not be evident in the early stages of the reaction. Capture images every 1-3 min until the reaction reaches a steady state fluorescence. If an automatically focusing stage is not available, briefly refocus each image prior to the images being taken. NOTE: While some photobleaching will occur, the effects on relative fluorescence due to photobleaching may be accounted for as long as the rate of photobleaching is known. This photobleaching rate may be estimated by exposing a fluorescent standard, such as a known concentration of GFP or a fluorophore mixture, to constant photobleaching over a period of time.
12. Record the time elapsed from the reaction assembly to the first image acquired. NOTE: This typically takes 4-5 min

2.3.4 Image Analysis and Data Processing

1. Using an image analysis software such as ImageJ, select the interior of the reaction chambers as an ROI. Acquire the mean fluorescence intensity value of the ROI for all images. NOTE: This is the raw fluorescence intensity trace.
    1. Perform this task in ImageJ using the Time Series Analyzer and ROI Manager plugins — Use Time Series Analyzer to choose regions of interest around the interior of each reaction chamber. Set

"AutoROIProperties" to an area which corresponds to the interior of each reaction chamber, check "Add On Click", and select each chamber. NOTE: This step may also be done using the ellipse tool to draw an ROI around the fluorescent chamber. This ROI size usually corresponds to a 30 x 30 pixel ellipse for a 10 μm diameter chamber viewed with a 100X objective.

2. Highlight all ROIs in the ROI Manager. Use the "Multi Measure" function to determine the fluorescence intensity mean of each ROI through the entire image stack. NOTE: A plugin named StackReg may be used to align the image stack, if necessary.

2. After acquiring the raw fluorescence intensity traces for all chambers in an experiment, determine the deterministic component of the reaction by taking an inter-experimental average across all traces, and subtracting the average from individual raw traces. Use data analysis software such as IGOR or MS Excel for this analysis. NOTE: This provides noise traces for each reaction chamber.
3. Analyze the gene expression noise from these reaction chambers using the same methods used to analyze gene expression noise derived from cells[35]

2.4 Representative Results

The distinct advantage of this microfabricated platform is in the application of the controllable elastomeric "control valve" which is independently actuated in order to confine CFPS reactions (Figure 2.2A). When the device is actuated, the membrane chambers are pressed against the glass slide to confine fluorescent reagents into an array of reaction chambers below (Figure 2.2C). In order to verify that the chambers reliably confine the reaction through the duration of the experiment, a basic FRAP (Fluorescence Recovery After Photobleaching) test was conducted[69]. A fluorophore (AF 555) was applied to the device, and the control valve was actuated; using the shutter aperture of the microscope, a single well confining the fluorophore was isolated and photobleached individually (Figure 2.2D). The chosen well became dark and did not recover in brightness until the control valve was depressurized 20 min. later, releasing the chamber from the glass. This test verifies that these reaction chambers remain well-sealed for the duration of the experiment.

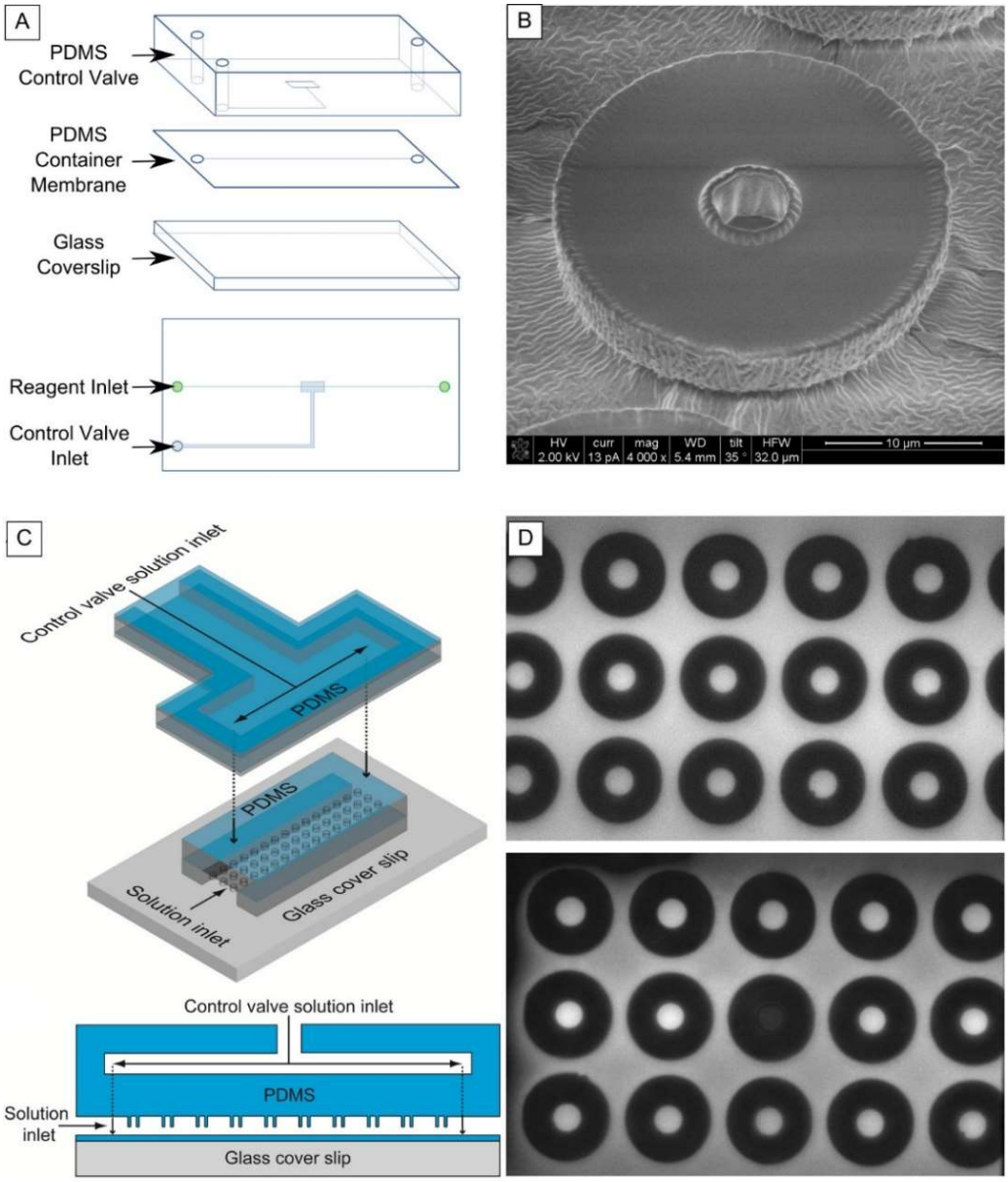


Figure 2.2 Fabrication of two-layer microfluidic device with sealable femtoliter-scale chambers. (A) Layout and exploded view of device layers. The device is composed of two PDMS layers and a glass coverslip. The PDMS membrane, sealed between the glass and control valve layers, holds the reaction chambers. (B) SEM image of PDMS reaction chamber. The interior diameter is 10 µm. (C) Schematic of input channels in device. Cell-Free Protein Synthesis (CFPS) reagents are flown through the reaction channel. Water is pressurized in the control valve to compress the reaction chambers against the glass slide, sealing the chambers. Reproduced from Fowlkes and Collier 2013[69], with permission from The Royal Society of Chemistry. (D) Fluorescence Recovery After Photobleaching (FRAP) test on a single well using FITC indicates chamber is well-sealed against external environment. The fluorophore was captured in the chambers (upper image) and a single well was photobleached (lower image). No fluorescence recovery was seen in the photobleached chamber until the control valve was released.

In optimal conditions, a CFPS reaction expressing an easily visualized protein (such as GFP or Luciferase) expresses detectable protein within a few minutes of being applied to this device. Over the lifetime of the reaction, protein synthesis in the interior and exterior of the reaction chambers is imaged and quantified by measuring units of fluorescence intensity within each chamber (Figure 2.3A). Fluorescence intensity, corresponding to protein concentration, may be mapped over time for each reaction chamber (Figure 2.3D).

Gene expression is an inherently stochastic process that introduces fluctuations (noise) at every molecular step (synthesis, degradation, protein-DNA binding, *etc.*)[40]. One branch of noise biology focuses on the probative value of gene circuit noise[96] . Expression in cell-free systems will have extrinsic noise effects that arise from interactions between the molecular machinery of expression and the surfaces that define the boundaries of the reaction vessels. These extrinsic effects will likely become more pronounced as cell-free reactions are confined into even smaller reaction chambers. The ability to perform time-lapse imaging of multiple confined CFPS reactions then enables the careful analysis of noise structure (magnitude and dynamics) in confined cell-free systems in a way directly analogous to methods that have been reported for cellular systems[35]. Figure 2.3C and Figure 2.3D show the time courses of constitutive GFP expression from a T7 promoter in a standard 384-well microplate with a well volume of 15 µL, compared to in PDMS reaction chambers 10 µm in diameter, corresponding to volumes of only about 300 fL, about seven orders of magnitude less. The variability in protein expression rates in the 10 µm reaction chambers is much higher than in the well-plate measurements, approaching those seen in cells.

Multiplexed reactions performed on the device exhibit similar kinetics to CFPS reactions performed in bulk on a microplate reader (Figure 2.3B), where there is a swift increase in fluorescence which plateaus, often assumed to be caused by resource limitation within the reaction volume[46,97]. This deterministic growth behavior, though fluctuating, is generally consistent across all reaction chambers, and between experiments — by averaging traces between chambers across experiments, the deterministic trend may be subtracted from trace values, leaving only the noise components of the reaction (Figure 2.4A). Figure 2.4B shows the GFP expression noise after removal of the deterministic, transient component (top), and the autocorrelation of

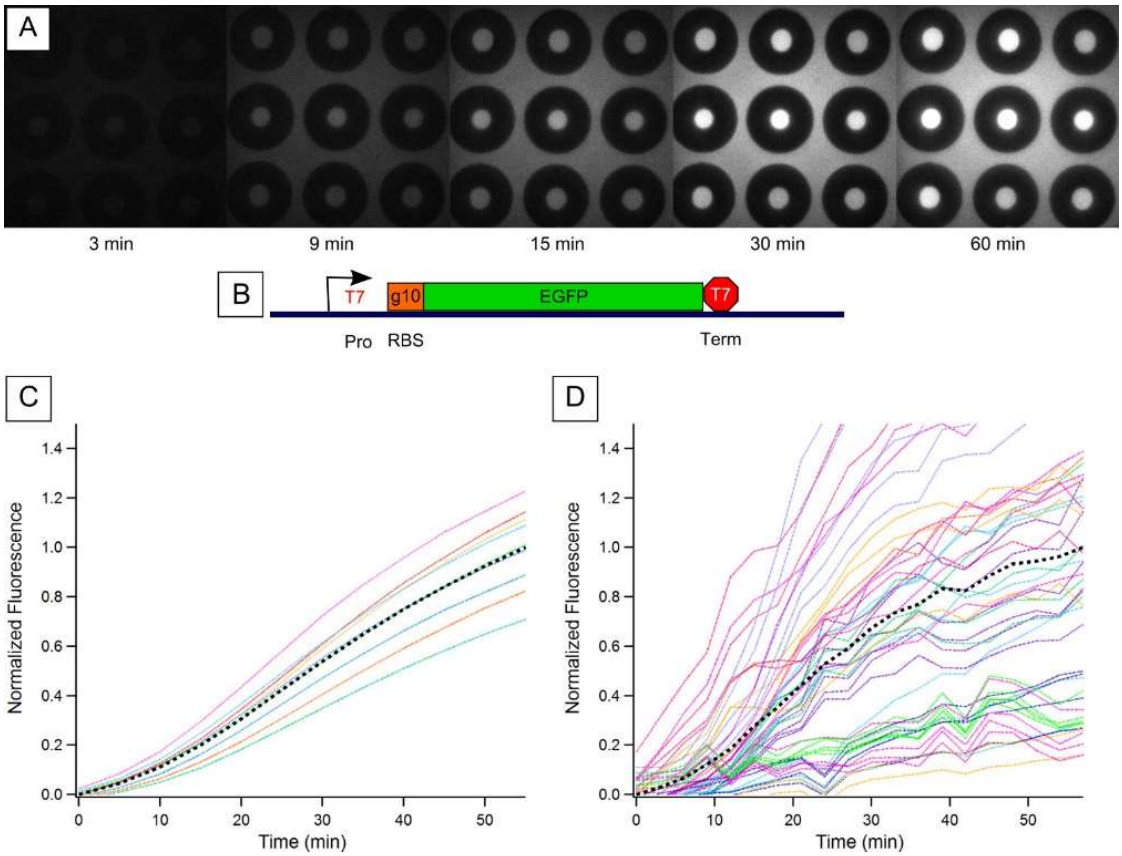


Figure 2.3: EGFP Expression in Confined Cell-Free Reaction. (A) Fluorescence images of sealed reaction chambers at chosen time points in the reaction. Protein production can be seen both inside the reaction chambers and outside the chambers in the main channel. (B) EGFP was cloned into a Pet3a vector, providing a T7 polymerase promoter and terminator and a strong ribosome binding site (RBS). (C) Normalized fluorescence measurements of constitutive expression of EGFP in a bulk cell-free reaction performed in a microplate reader. CFPS reactions usually produce protein quickly before slowing to a 'steady state' fluorescence — this is associated with resource limitation. Black dashes indicate the average trace. (D) Normalized fluorescence of 51 raw fluorescence intensity traces read from 51 reaction chambers over several experiments. Black dashes indicate the average trace over several experiments, which illustrate the deterministic component of the protein expression.

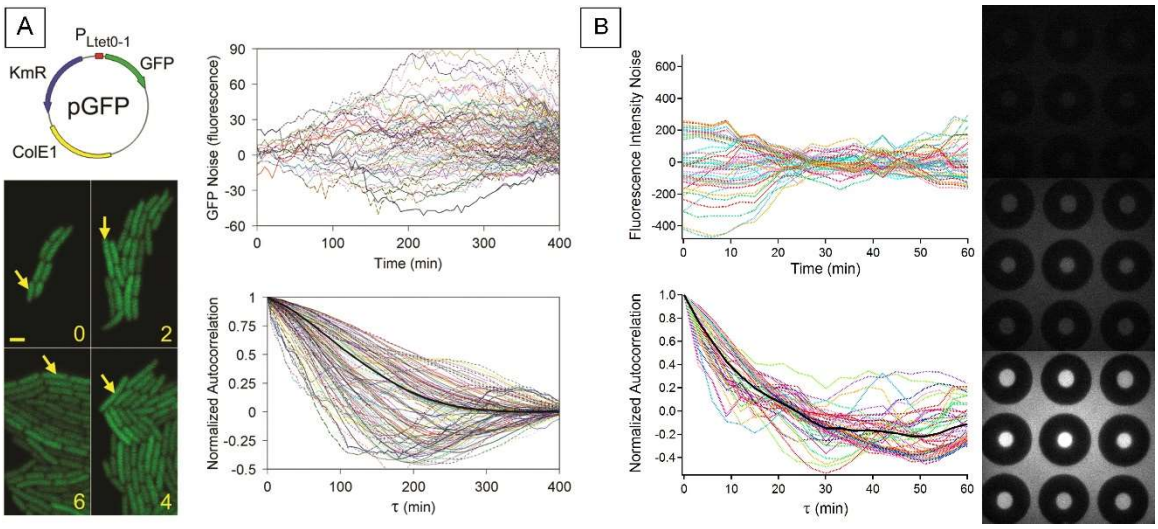


Figure 2.4: Individual Noise Traces and Noise Autocorrelation of a Cellular and Cell-Free System. (A) From Austin et al., 2006. Noise in GFP expression (top) and normalized autocorrelation functions (bottom) acquired from tracking GFP production in living bacteria. Reprinted by permission from Macmillan Publishers Ltd: [Nature] 25 (Vol. 439), copyright (2006). (B) Noise in GFP expression (top) and normalized autocorrelation functions (bottom) acquired from GFP production in cell-free system, tracked in microfluidic device reaction chambers.

the noise (bottom), while Figure 2.4B shows the corresponding traces in the 10 μm reaction chambers. The distribution in the half-times of the autocorrelation traces gives the frequency dependence of the noise while the zero lag time of the autocorrelation traces gives the magnitudes of the noise, as the variance.

2.5 Discussion

Gene expression in cells is inherently noisy due to small cellular volumes and low copy numbers of important reactants. Noise biology often focuses on the sources, processing, and biological consequences of fluctuations in the populations, concentrations, positions, or states of molecules that control gene circuits and networks[98]. The vast majority of this work has been performed in cellular systems, which has the advantage of viewing the noise of a gene circuit within the natural context of the genetic networks within the cell. However, cell-free systems allow the characterization of the intrinsic fluctuations of an individual gene circuit without the confounding extrinsic effects[83] that cannot be avoided in cellular systems. Analysis of noise can offer important physical insights into how genetic circuits are structured and how they function, and has

been used in cellular systems to characterize negative[35] and positive[29] autoregulation, extrinsic and intrinsic contributions to expression noise[83], and transcriptional bursting[99,100]. Here we describe the study of a cell-free expression system in microfluidic devices that enable the simultaneous control of reactor size and reaction initiation times, in order to better understand the roles that confinement and crowding[60,66] have on intrinsic protein expression noise without the complications associated with living cells.

The key enabling feature of the design is the integration of arrays of femtoliter-volume (micron-scale) reaction chambers used for confining the reactants of a cell-free protein expression system, with an elastomeric “control valve” membrane in PDMS that traps the reactants at a well-defined, “time zero” for reaction initiation (Figure 2.1). This control allows the kinetics of the reactions involved in protein synthesis to be followed in real time with high precision. As such, it is important to manage cell-free reactants so that inter-experimental variability is minimized as much as possible. This control allows us to evaluate noise structure of cell-free genetic circuits in a manner that is analogous to techniques previously used to evaluate gene expression in living cells.

As reactants used in CFPS systems can be sensitive to freeze-thaw cycles, it is important to keep the reactants cold and minimize the time the reactants spend thawing on ice. It is good practice to periodically test the expression of the CFPS system in bulk in order to identify changes in expression levels over time — this may be done in a 10-15 µl reaction in an Eppendorf tube, or in a device like a microplate reader, which performs multiple reads over time to capture reaction kinetics. Noting the age and thaw times of the reactants for every experiment will help when troubleshooting low expression levels. Furthermore, when assembling CFPS reagents, it is important to note that the reaction will begin once it is fully assembled and removed from the ice. In order to maintain a consistent “time zero”, it is helpful to record the time following the initiation of the CFPS reaction after the final addition of the DNA input, and to apply the reaction as quickly as possible to the incubated device. This process should take about 4-5 min, and fluorescence should not yet be visible within the reaction chambers. This control assures that the time available to visualize the growth portion of the reaction curve is maximized.

Before running CFPS reactions on the device, it is advisable to run quality-control tests to verify there is no leakage from the chambers. A FRAP test can be performed (as

in Figure 2.2D) by applying a fluorophore to the device and exposing an individual well until the well is completely bleached. If the chambers are well-sealed, no recovery should be visible inside the well — there should be a stark contrast between the walls of the compartment and the interior and exterior spaces. If fluorescence recovery is apparent or the walls of the reaction chamber are not well defined, the pressure on the control valve should be increased or the device should be checked for leakage or delamination from the glass slide.

This protocol has been tested with CFPS reagents from a commercial *E. coli* cell-free protein expression kit (scaled to 25 µl), though other robust CFPS systems may be used. It is possible to use volumes much lower than 25 µl when applying reactions to the device, which may be helpful when reagent cost is a limiting factor in experiments. Once reactants are added to the device and the reaction chambers are sealed, it is not possible to add reactants to the solution without de-actuating the control valve — thus this device is not suitable for reactions which require the addition of reagents during the course of the reaction. This device is also not optimized for observing CFPS reactions which may run longer than 3 hr — the effects of dehydration and drying of the device after this time period have not been evaluated. If longer reaction times are desired, these effects may be mitigated by sealing the device to prevent evaporation, changing the incubation temperature, or by using a humidity chamber. Modifications to the device design, such as nanoporous structures in the chamber walls[54,101] or the inclusion of a porous membrane layer, represent a few methods which could allow reagent exchange and thus lengthen reaction timescales.

Microfabricated reaction compartments of uniform volume are valuable for maintaining consistent dimensions across experiments and highly suitable for investigation into “side reactions” with the compartment walls. Unlike methods using non-microfabricated techniques, these reactions must be evaluated in small numbers, and do not provide dimensional flexibility during experiments. However, the controllable design for these reaction chambers is highly suitable for time-lapse microscopy, and may be an illuminating complement to a high-throughput method of confinement.

3 RESOURCE SHARING CONTROLS GENE EXPRESSION BURSTING

A version of this chapter was originally published by Patrick M. Caveney, S. Elizabeth Norred, Charles W. Chin, Jonathan B. Borekyo, Brandon S. Razooky, Scott T. Retterer, C. Patrick Collier, and Michael L Simpson:

Caveney, P. M. *et al.* Resource Sharing Controls Gene Expression Bursting. *Acs Synth Biol*, doi:10.1021/acssynbio.6b00189 (2016).

This chapter has been adapted from its published format to accommodate new Figure, Table, and Equation enumeration. The Supplementary Information associated with this work may be found in the Appendix; all references are located at the end of the document. The work presented in this chapter relies heavily upon the experimental design and methods presented in the previous chapter[45]. PMC, SEN, CWC, JBB, BSR, STR, CPC, and MLS conceived and planned the experiments. Experiments were performed by PMC, BSR, and SEN. The microfluidic device was originally designed by CPC and Dr. Jason Fowlkes for use in a previous publication[69]. BSR adapted the vesicle fabrication method from previous work by Nishimura et al. 2015[51]. SEN, PMC, and CWC performed image analysis. CWC programmed Gillespie simulations and wrote analytical scripts. SEN's specific contributions included preparing experimental reagents, fabricating microfluidic devices, performing experiments, maintaining laboratory setups, performing image analysis and data acquisition, and participating in group meetings and data analysis. All authors participated in data analysis and contributed to the final manuscript. The authors acknowledge and thank Dr. Sukanya Iyer for constructing the Pet3a-EGFP plasmid used in these experiments.

3.1 Abstract

Episodic gene expression, with periods of high expression separated by periods of no expression, is a pervasive biological phenomenon. This bursty pattern of expression draws from a finite reservoir of expression machinery in a highly time variant way, i.e., requiring no resources most of the time but drawing heavily on them during short intense bursts, that intimately links expression bursting and resource sharing. Yet, most recent investigations have focused on specific molecular mechanisms intrinsic to the bursty behavior of individual genes, while little is known about the interplay between resource sharing and global expression bursting behavior. Here, we confine *Escherichia coli* cell

extract in both cell-sized microfluidic chambers and lipid-based vesicles to explore how resource sharing influences expression bursting. Interestingly, expression burst size, but not burst frequency, is highly sensitive to the size of the shared transcription and translation resource pools. The intriguing implication of these results is that expression bursts are more readily amplified than initiated, suggesting that burst formation occurs through positive feedback or cooperativity. When extrapolated to prokaryotic cells, these results suggest that large translational bursts may be correlated with large transcriptional bursts. This correlation is supported by recently reported transcription and translation bursting studies in *E. coli*. The results reported here demonstrate a strong intimate link between global expression burst patterns and resource sharing, and they suggest that bursting plays an important role in optimizing the use of limited, shared expression resources.

3.2 Introduction

Bursty or episodic gene expression—periods of high expression separated by periods of very low or no expression—is a widespread phenomenon observed across biological domains[31,37,43,102-105]. The common gene expression burst pattern (Figure 3.1A) consists of short intense periods of expression separated by relatively long periods without expression[40,43,100]. This bursty pattern of expression draws from a finite reservoir of reusable expression machinery, e.g., polymerases and ribosomes, in a highly time variant way. Thus, the majority of expressed genes require no resources most of the time, yet these genes draw heavily on them during short intense bursts. The reservoir of expression machinery is common to all genes in the organism, and many studies have addressed how, in a time-averaged way, expression resources are shared among genes[106-108]. Thus, genes with no direct regulatory relationships still interact through expression resource sharing[46,107,109,110]. However, little is understood about the relationships (if any) that exist between expression bursting patterns and resource sharing.

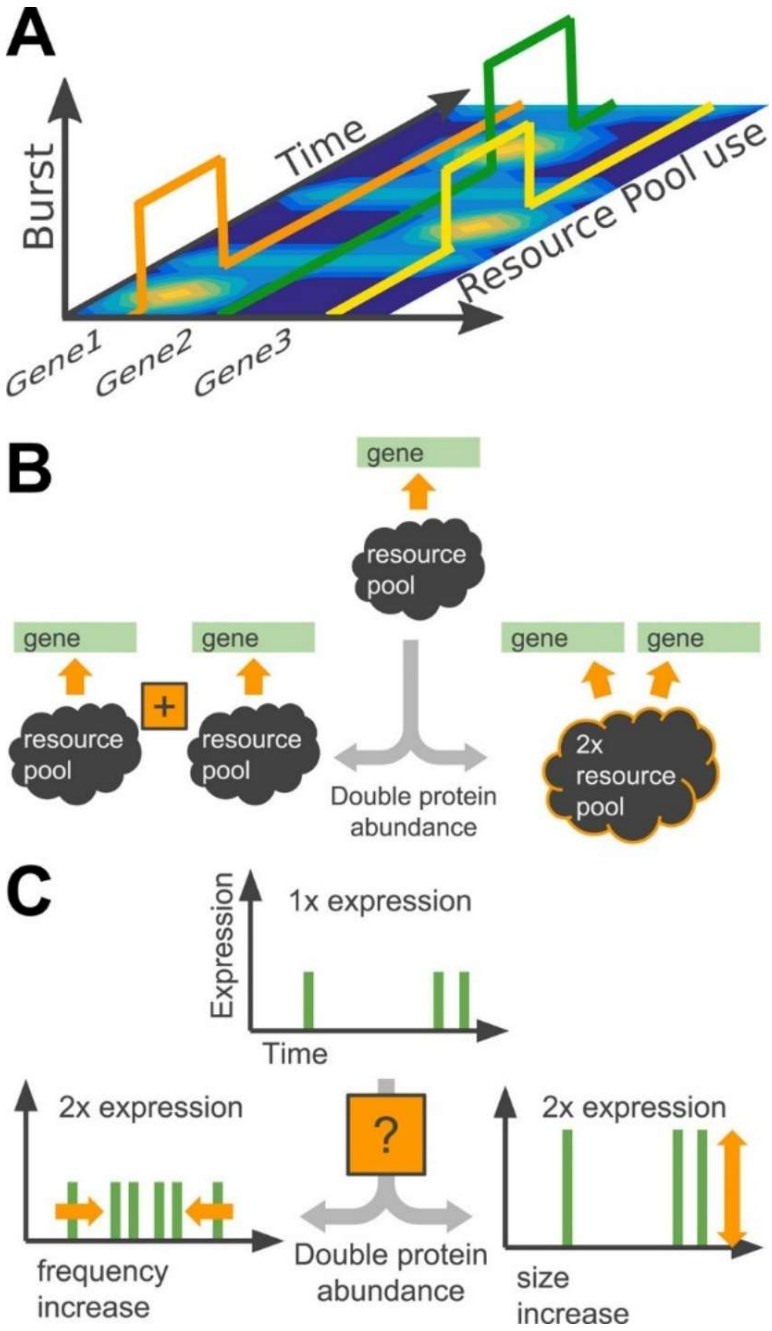


Figure 3.1: Bursty gene expression impact on global resource utilization. (a) Bursty gene expression draws heavily from shared global resource pools but only for limited durations. (b) Increasing protein abundance by increasing the number of genes and the amount of expression resources. The larger resource pool may be shared by all of the genes (right), or the sharing of resources may be enforced by compartmentalization (left). (c) Protein abundance changes may be driven by an increased expression burst size (right) or bust frequency (left). Does the resource sharing scenario affect the expression burst pattern?

Gene expression bursting studies have most often focused on molecular processes that are not directly related to resource sharing. Translational bursting, occurring when many proteins are synthesized from reading the same mRNA molecule, is initiated by the birth of an mRNA molecule and terminated by its decay. Transcriptional bursting has been shown, at least in part, to be controlled by molecular processes such as transcription factor kinetics[111-113], promoter architecture[32,114-116], chromatin remodeling[115,117], supercoiling[30], and transcriptional reinitiation[37,118,119]. Such a view sees expression bursting primarily as an intrinsic property of individual gene circuits. However, given the demands that an expression burst places on the common cellular pool of resources, this intrinsic view of bursting may be overly limiting. Instead, it seems likely that changes in the size of the common reservoir of expression machinery or in the number of genes drawing upon these resources (Figure 3.1B) will globally impact expression burst patterns (Figure 3.1C).

Studies of gene expression patterns have been carried out using various experimental techniques[32,36,43,120,121] in cells or in cell-free systems that were not confined to cellular-scale reaction chambers. Cell-based platforms provide the important advantage of viewing function within its natural context, but it is difficult to manipulate specific parameters, such as confinement, when they are isolated from all the other cellular processes, such as growth, cell division, and global gene expression. Conversely, *in vitro* reaction chambers are especially suited for isolating the effects of specific mechanisms from confounding cellular processes[48,54,101,122,123], and cell-free protein synthesis (CFPS) systems have been successfully used to observe gene expression bursting[30]. Recently, arrays of microfabricated cellular-scale reaction chambers have been demonstrated to be a viable way to confine CFPS reactions to study gene expression, in particular the noise in expression[44,45]. Bursting and noise are inseparably linked as bursting is often the dominant contributor to expression noise[43,100,119,124,125], and noise measurements are often used to understand the underlying dynamics of gene expression *in vivo*[42,43,100,126]. In combination, microfabricated cell-scale reactors and gene expression noise measurements provide a unique platform to explore gene expression bursting and resource sharing in well-controlled and easily manipulated environments.

Here, we study cell-free gene expression in synthetic reaction chambers under different resource sharing scenarios. Specifically, we measure gene expression burst patterns as the number of genes and size of the resource pool are increased (i.e., the volume of the reaction chamber is increased) either by summing together discrete individual chambers (discrete resources; Figure 3.1B, left) or by making one larger chamber (shared resources; Figure 3.1B, right). As expected for both cases, total protein production and production rate scaled linearly with the amount of DNA and expression resources. However, while the discrete resources case (i.e., summed smaller chambers) generated higher protein abundance through more frequent bursts (Figure 3.1C, left), the shared resource case (i.e., individual larger chambers) drove increased protein production by increased burst sizes (Figure 3.1C, right). Surprisingly the divergent bursting behavior was found even though a constant ratio between expression resources and DNA was maintained for both scenarios, showing that resource sharing and expression bursts are directly coupled. For transient expression in cell-free expression chambers, we present a model that suggests this behavior emerges from the timing of mRNA production and size of the available resource pool. The mRNA molecules produced early consume most of the translational resources and make many proteins, whereas mRNAs produced later are created in a resource poor environment and make few proteins. As a result, in all cases the same number of mRNA molecules is responsible for the majority of protein production, but in the large chambers, those few mRNA molecules experience a very large translational burst size. This model of self-reinforcement of bursts may explain the robust positive correlation observed between transcriptional and translational burst sizes in *Escherichia coli*[127] and suggests that burst size control is the principle mechanism driving protein abundance changes.

3.3 Results

To confine cell-free expression reactions, we fabricated actuatable polydimethylsiloxane (PDMS) cylindrical chambers on membranes suspended above microfluidic channels using soft lithography as described earlier[45]. All chambers were 5 μm deep but ranged in diameter from 2 to 10 μm and in volume from ~15 to ~400 fL, respectively. A 25 μL commercial raw extract CFPS reaction was mixed with 500 ng of enhanced green fluorescent protein (EGFP)-coding pET3a plasmid and was confined

within the chambers through a two-step process described previously[45]. First, the cell-free mixture was loaded into the microfluidic channel using <10 psi of pressurized nitrogen, and then the membrane was actuated with ~20 psi of DI water to seal the reaction chambers (Figure 3.2A). Since the chambers were actuated and imaging began very soon, ~4 min, after plasmids were added to the CFPS mixture, this experimental platform provided a well-defined t = 0, i.e., the time when expression began, thereby allowing for the direct comparison of results from experiments performed on different days (See Appendix, Figure 3.7). Additionally, through microfabrication techniques, the reaction chamber size could be easily and accurately defined.

The time course of protein expression was characterized by measuring total fluorescence of EGFP within individual chambers every 3 min for 1 h (Figure 3.2B; Methods). Time courses averaged across all 119 individual chambers were similar to those observed in bulk reactions, although as reported elsewhere[55,63] confined reactions did proceed at a slightly increased rate (See Appendix, Figure 3.8). The fluorescence transients exhibited a relatively rapid increase in protein expression initially, followed by a much slower rate of GFP accumulation. This two-phase expression profile is consistent with resource limitations and not equilibrium between protein decay (e.g., photobleaching; See Appendix, Figure 3.9 and Figure 3.10) and synthesis[44,47]. Similar to cellular experiments[35], there was considerable chamber-to-chamber variation in the final fluorescence levels, yet there was a striking uniformity to the shape of the transient response between experiments and chamber sizes (See Appendix, Figure 3.7 and Figure 3.11).

This uniformity in transient response allowed the use of a previously described method[29,43,44] to extract the noise from each individual trace (Figure 3.2D; Methods). Briefly, the deterministic transient response was removed from each trace. The remaining signals were assumed to be due only to the stochastic fluctuations in the gene expression process, i.e., the expression noise. The magnitude of the expression noise within an individual chamber was quantified using the square of the coefficient of variation (CV^2; variance/[final fluorescent abundance]^2). The CV^2 of individual chambers and the composite CV^2 of all chambers of the same size were plotted versus their final fluorescence abundance (Figure 3.2D). Similar to cellular experiments[35,43], the CV^2 values of individual chambers were scattered around the composite CV^2

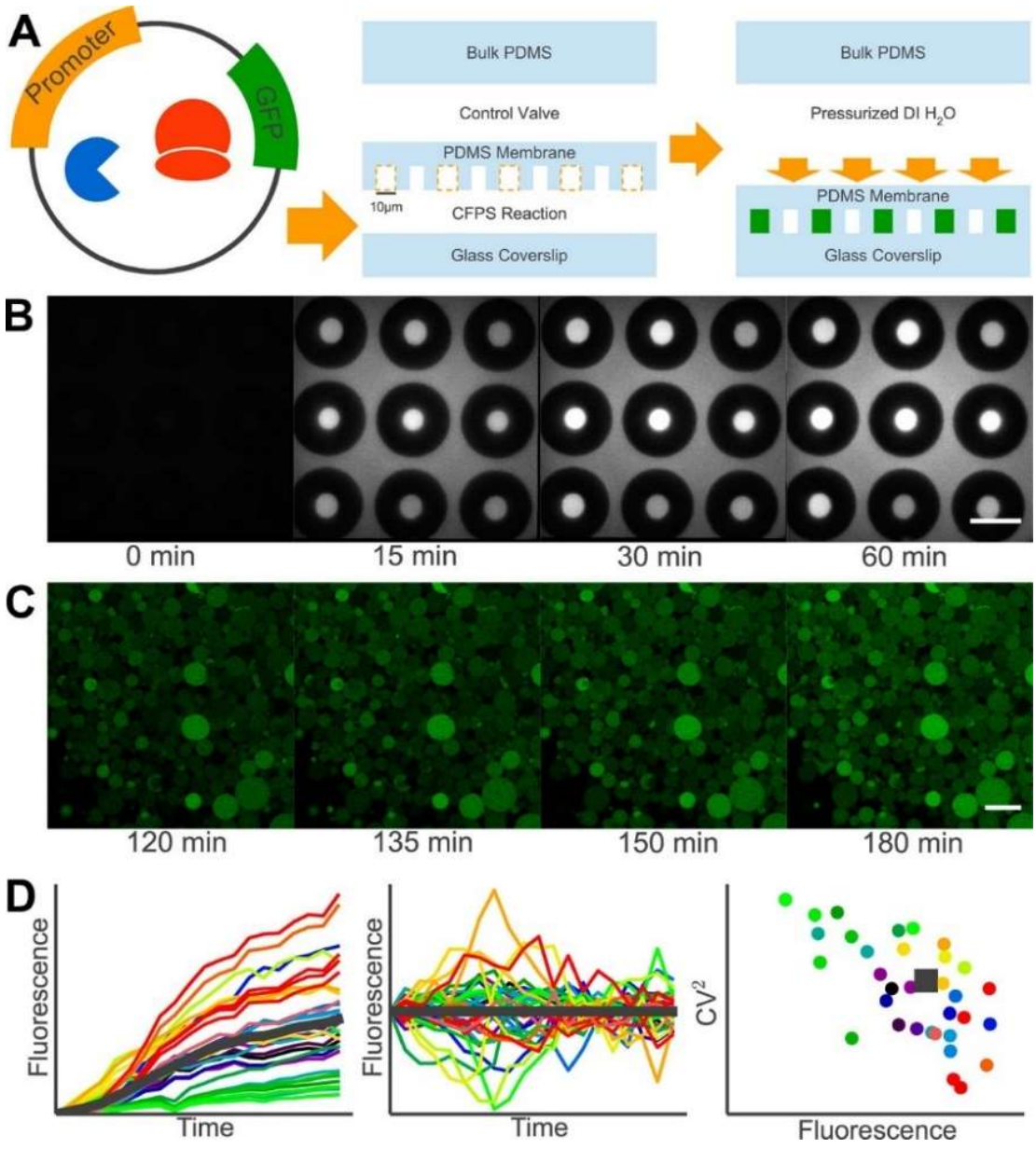


Figure 3.2: Confined cell-free gene expression and noise measurements. (a) Cell-free protein synthesis (CFPS) reactions expressing EGFP were isolated within microfabricated chambers. (b) Time-lapse fluorescence microscopy was used to image the confined reactions every 3 min for 1 h. Images from an expression experiment performed in 10 μm diameter reaction chambers show fluorescence intensity increasing over time. Scale bar, 20 μm. (c) Representative z-slice of POPC vesicles expressing EGFP. Imaged every 3 min for 1 h. Scale bar, 20 μm. (d) (left) Time history of the growth of the protein population collected for each chamber. (middle) Gene expression noise found by removing the deterministic general trend from each expression transient. (right) CV^2 and final fluorescence level (protein abundance) for individual chambers (colored circles) and for the average of all chambers (gray square).

oriented along a line inversely proportional to final fluorescence intensity.

The baseline expression noise vs abundance relationship (Figure 3.3A) of this experimental system was established using the average behavior of the 2 μm chambers (large filled orange triangle). To study the effects of resource sharing on gene expression noise and bursting, the volume of the reaction was increased from this baseline in two ways. First, composite chambers were created by summing fluorescence signals from between two and six individual 2 μm chambers (open orange triangles Figure 3.3A). These composite chambers allowed for the total reaction volume to be varied while ensuring that expression resources were shared exactly as they were in the individual 2 μm chambers. Average final fluorescence levels in these composite chambers scaled linearly with volume, and as expected[36,38,43,104], CV^2 scaled linearly with the inverse of abundance (Figure 3.3A dotted line). This behavior indicated that expression within each of the chambers was similar to, but statistically independent of, the other chambers included in the sum. Said differently, the number of expression bursts, or burst frequency, increased linearly (Figure 3.4) with the number of chambers included in the composite.

Reaction chamber volume was also increased by fabricating larger (5 and 10 μm diameter) individual chambers, allowing one individual expression resource pool to be shared by all of the genes. These larger chambers had proportionally more plasmids and resources, and once again the final protein abundance scaled linearly with volume (Figure 3.3A inset). Yet, in striking contrast to the composite 2 μm chambers, the CV^2 values of these shared resource chambers (large filled symbols in Figure 3.3A) were insensitive to abundance variation driven by changes in chamber volume (Figure 3.3A solid line; greater than a factor 25 change in abundance with less than a factor 3 change in CV^2). Notably, a composite of six individual 2 μm chambers, which is nearly equal in volume and final protein abundance to a single 5 μm chamber, produced a CV^2 approximately a factor of 5 lower than single 5 μm chambers (red box, Figure 3.3A). This strikingly different noise behavior is not only apparent in the time histories of the expression experiments but also is seen in the distribution of final protein abundances seen across the populations of individual and composite reaction chambers (Figure 3.3B).

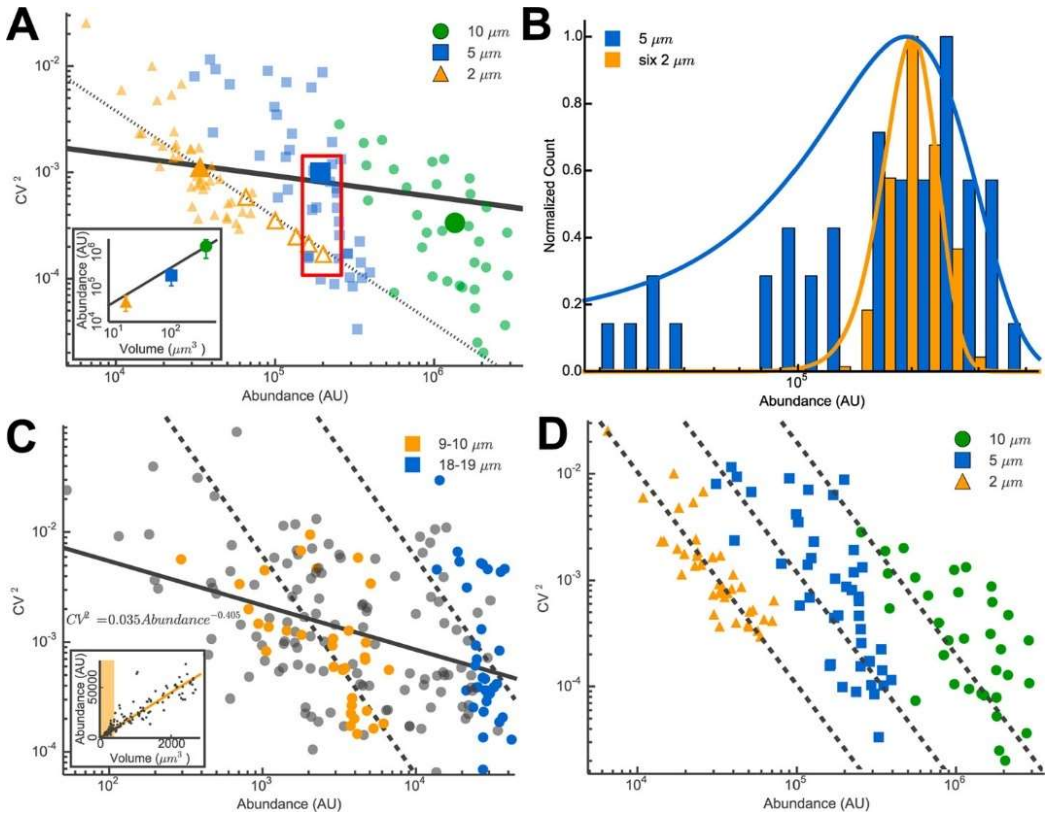


Figure 3.3: Effects of resource pool size and configuration on gene expression noise in both microfluidic chambers and vesicles. (a) CV^2 vs protein abundance for 2, 5, and 10 μm diameter chambers. The small filled data points represent individual chambers, and the large filled data points show the mean behaviors for all chambers of a given size. The dotted line is of the form *a*/abundance, and where *a* is a constant that is calibrated so the line passes through the mean of the 2 μm diameter chambers (large filled orange triangle). The open orange triangles show combinations of 2 μm chambers. The left most open triangle shows the average behavior of sums of two individual 2 μm chambers, and the right most open triangle shows the average behavior of sums of six individual 2 μm chambers. CV^2 for these combinations of 2 μm chambers closely follow the a/abundance trend. In contrast, the individual 5 μm chambers deviate strongly (solid line) from the a/abundance trend even though their volume and protein abundance are about equal to six 2 μm chambers (red box). The inset shows that protein abundance scales approximately linearly with volume. (b) Histograms of protein abundance across the ensemble of individual 5 μm chambers (blue) and the ensemble of combinations of six 2 μm chambers (orange). Histograms are normalized (i.e., frequency of most likely protein abundance is set to 1) and fit with normal distributions (solid lines). (c) CV^2 vs protein abundance for vesicles ranging in diameter from 4 to 19 μm. Each data point (gray or colored) represents an individual vesicle. The orange points are vesicles with diameters of 9–10 μm, and the blue points have diameters of 18–19 μm. The solid line is a power law fit to all points. While abundance varies by 3 orders of magnitude, CV^2 values decrease only by about 1 order of magnitude. Dashed lines show fits to individual volumes (orange and blue), where CV^2 goes as 1/abundance^2. The inset shows that protein abundance scales linearly with vesicle volume. The shaded region on the inset corresponds to the volume range explored using the chambers. (d) Same data in (a) without means. Dashed lines are fits to each size chamber where CV^2 goes as 1/abundance^2.

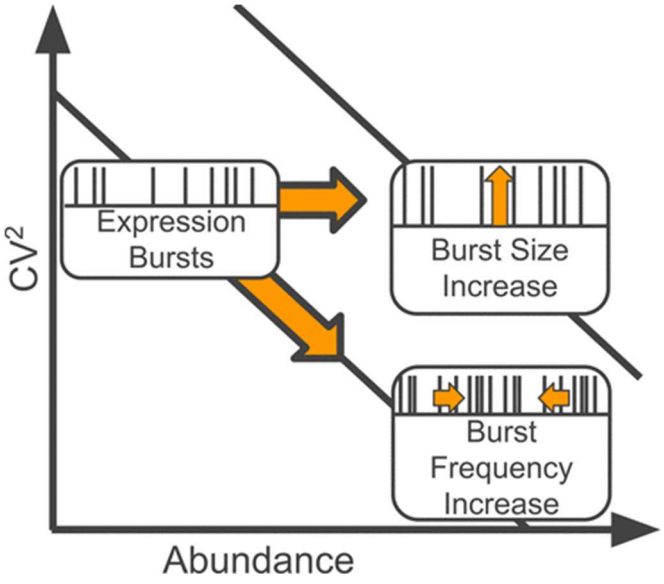


Figure 3.4: Graphical Description of "Noise Space"

To confirm that this flat CV^2 trend across volumes was not unique to PDMS chambers, perhaps due to surface charge or molecular adsorption or absorption, we encapsulated PURE cell-free reactions expressing EGFP in more biologically similar POPC water-in-water vesicles (Methods) and imaged them with confocal microscopy (Figure 3.2C). Vesicles ranged from about 4–19 μm in diameter (~65–3500 fL). Just as in the PDMS chambers, abundance scaled linearly with volume (Figure 3.3C inset), and CV^2 was only modestly sensitive to abundance changes across the range of volumes (e.g., fluorescent abundance increased by 3 orders of magnitude while CV^2 decreased by only 1 order of magnitude; solid line in Figure 3.3C).

While insensitive to systematic changes in protein abundance driven by changes in the reaction volume, CV^2 was hypersensitive to random fluctuations in protein abundance across a population of same-sized reaction chambers. Final protein abundance across the population of individual 2 μm chambers varied less than 1 order of magnitude, from 10^4 to 8×10^4 AU, but CV^2 varied more than an order of magnitude, from 10^{-2} to 3×10^{-4}. Similar behavior was observed across populations of 5 and 10 μm chambers as well (Figure 3.3D).

3.3.1 Box 1

In bursty protein synthesis[45]

$$\langle \boldsymbol{P} \rangle = \frac{\boldsymbol{B f_B}}{\boldsymbol{\gamma_P}} \quad 3.1$$

$$\boldsymbol{CV^2} \approx \frac{\boldsymbol{\gamma_P}}{\boldsymbol{f_B}} \quad 3.2$$

where $\langle P \rangle$ is the amount of protein produced and γ_P is the protein decay rate. The term f_B is often called the burst frequency, and it is a measure of how often a burst occurs. In the case of a single bursty gene, f_B is simply a frequency (Figure 3.1C) and is the inverse of the time period between adjacent burst events. If there are multiple copies of a gene, f_B is the average number of these genes that are active at any given time. B is the size of an expression burst, i.e., the average number of protein molecules produced in one expression burst.

Although protein abundance may be changed by either burst size or burst frequency, CV^2 is sensitive only to changes in burst frequency. As a result, CV^2 vs protein abundance plots reveal if abundance changes are driven primarily by changes in burst size or in burst frequency (Figure 3.4). In the cell-free experiments reported here, systematic protein abundance changes were induced by changes in reaction chamber volume (i.e., by changes in the number of copies of the gene and the associated expression resources). Changes in protein abundance that induced little or no changes in CV^2 were indicative of changes in burst size with little or no change in burst frequency (Figure 3.4). In contrast, changes in protein abundance where CV^2 varied inversely with protein abundance were indicative of changes in burst frequency with little or no change in burst size (Figure 3.4).

3.4 Discussion

The most important implication of the results is that resource sharing and expression bursting are intimately linked. Sums of small discrete pools of resources achieved much lower expression noise than large shared pools, even with a constant ratio between DNA and expression resources (Figure 3.3A). These results lead to the inference that expression occurring in the large shared resource pool environments displays larger bursts (Figure 3.4) than equal volume sums of discrete resource pools. The intriguing result is that instead of frequently consuming a small fraction of the

available resources, individual genes are more apt to infrequently consume a large fraction of total resources. In a large shared volume, when both the pool of resources and the number of genes increased proportionally, the increased resources were drawn into making bursts larger (Figure 3.1C, right), not more frequent (Figure 3.1C, left). It seems that, given additional resources, expression bursts are more readily made bigger rather than made more often.

We investigated expression in the reaction chambers using a random telegraph[128,129] model of transcription from a group of genes competing for a shared population of translational resources (modeled here just as ribosomes; Figure 3.5A). In this model, genes switched between an OFF state with no transcription and an ON state where they produce uniquely identifiable mRNA molecules. Ribosomes diffused between a global pool and being bound to mRNA. Once bound to mRNA molecules, the ribosomes were less likely to rerandomize by diffusion back into the global pool, i.e., $k_b \gg k_r$. To correspond with the experiments, we performed Gillespie simulations of this model and measured the CV^2 of the resulting protein population using exactly the same algorithms used for the experimental data (Methods). Small reaction chambers were modeled with a small number of genes drawing from a small pool of ribosomes, whereas larger reaction chambers were modeled as larger numbers of genes drawing from a proportionally larger pool of ribosomes. In agreement with the experimental data, larger reaction chambers led to a proportional increase in protein abundance, yet the CV^2 of this population remained flat (Figure 3.5B). The invariance of the CV^2 to protein abundance indicated that burst size, not frequency, was responsible for increasing protein abundance (Figure 3.4).

Interestingly, examination of the different simulations showed that regardless of reaction chamber size a similar number of genes, those that burst ON early, captured a disproportionate percentage of the available translational resources. Conversely, genes that turned ON late captured very few translational resources. The net result was that in both small and large reaction chambers just a few mRNA dominated protein synthesis (Figure 3.5C). The dominant mRNA molecules in the bigger reaction chambers drew from a much larger pool of available translational resources, so in effect each of these mRNA molecules experienced larger translational bursts than dominant mRNA molecules in the smaller reactions. This model predicts that the larger protein

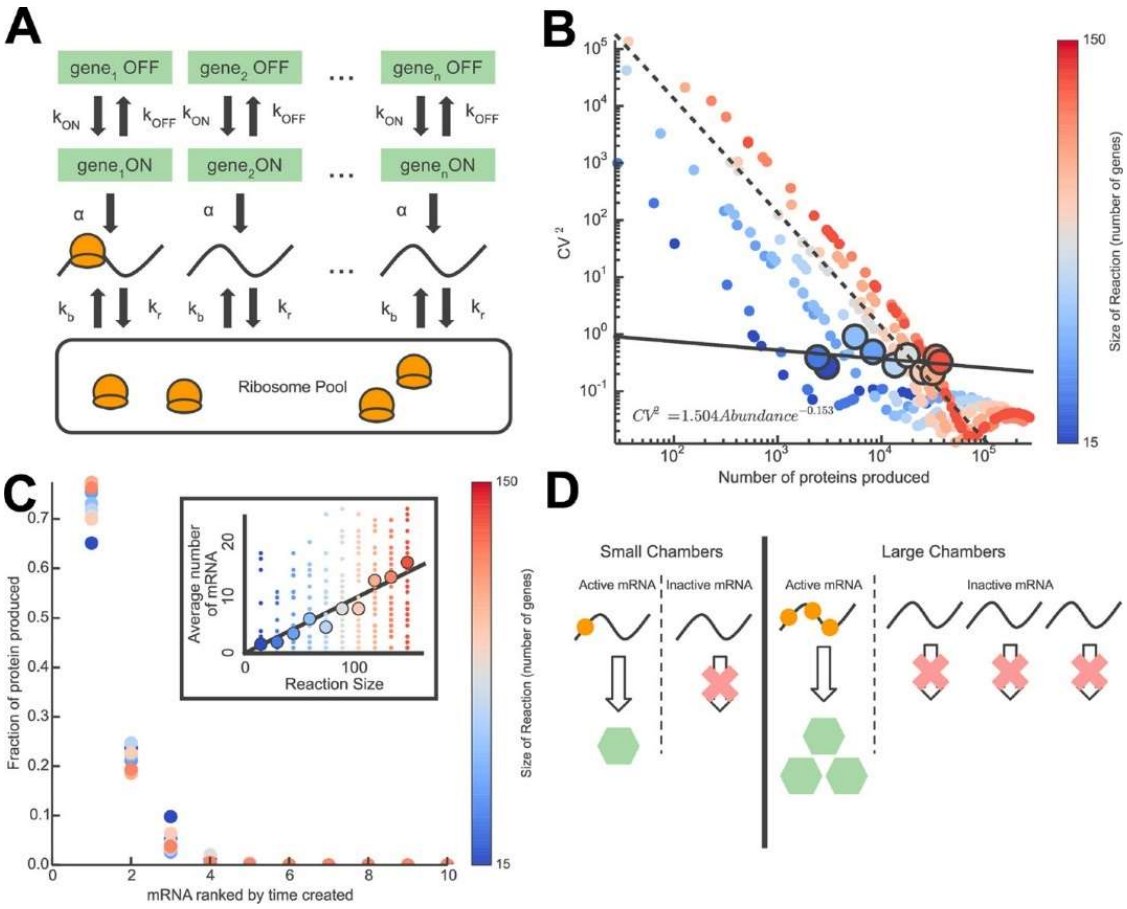


Figure 3.5: Model of the effects of resource pool size on expression bursting. (a) In silico model of resource sharing includes a resource pool of a limited number of reusable molecules, e.g., ribosomes, that associate with one of n genes (rate constant k_b) and return to the resource pool with rate constant k_r. The number of resources is proportional to the number of genes in each reaction. Genes burst ON and OFF at rates k_{ON} and k_{OFF}, respectively. Molecules of mRNA are created at rate α only in the ON state. (b) CV^2 vs protein abundance from the model described in (a). Colors represent the size of the reaction from 15 to 150 genes. Large points are means of multiple runs of the same reaction size. The solid line is a power law fit to all mean data points, and the dashed line is of the form a/abundance^2, where a is selected so that the line passes through the mean of a 105 gene reaction. As in the experimental chambers and vesicle data, CV^2 from this model is relatively insensitive to increases in abundance driven by changes in reaction sizes, but it is highly sensitive to increases in abundance that occur within a single reaction size. (c) mRNA molecules are ranked in the order of the time they were created. The y-axis shows the fraction of the total protein translated from each mRNA molecule. Points are colored by the reaction size (small reaction sizes are more blue, larger ones are more red). mRNA molecules made early, regardless of the reaction size, collected a disproportionate amount of resources and made a disproportionate amount of the total protein. The inset shows that mRNA abundance scales linearly with reaction size. (d) Schematic showing large reactions produced proportionally more mRNA molecules than small reactions, but most of this mRNA was inactive.

populations found in larger reaction chambers resulted from the translational amplification of burst sizes, not the initiation of more bursts. In this model, larger chambers did indeed produce proportionally larger mRNA populations (Figure 3.5C inset), yet much of this mRNA was translationally inactive because earlier produced mRNA molecules had already sequestered translational resources (Figure 3.5D).

Although the model predicts that CV^2 is insensitive to abundance changes driven by volume increases, it predicts hypersensitivity (CV^2 α 1/abundnace^2; dashed line Figure 3.5B) to abundance variations that occur across chambers with the same volume. Examination of the simulation results showed that this strong relationship between CV^2 and abundance arose from the natural variability in the number of genes that initiated transcription early enough to effectively compete for ribosomes. Some runs of the simulation naturally showed a larger than average number of early turn-on genes, so these simulation trials exhibited a larger than average burst frequency. However, in these trials, a fixed population of ribosomes was distributed across this larger number of bursts, resulting in a reduction in the burst size. The prediction is a distribution of expression burst patterns across an ensemble of same-sized chambers where higher burst frequency is correlated with lower burst size (See Appendix, Figure 3.12). The net result is a distribution of final protein abundances and a CV^2 that declines sharply as protein abundance increases. In good agreement with this model prediction, CV^2 is highly sensitive to abundances changes across ensembles of same-sized PDMS chambers and POPC vesicles, and ensembles are well-fit by CV^2 α 1/abundance^2 (dashed lines in Figure 3.3 C,D and Appendix Figure 3.13).

The results presented here suggest that expression bursts are self-reinforcing and that available translational resources are readily drawn into active transcriptional regions. This leads to the intriguing idea that, at least within prokaryotic cells, large translational bursts may be the direct result of large transcriptional bursts (Figure 3.6A). Recent work has shown that in *E. coli* large mRNA populations are strongly correlated with large transcriptional burst sizes[100] and that large protein populations are strongly correlated with larger translational burst sizes[47]. Taken together, these data demonstrate a strong correlation between transcriptional (B) and translational (b) burst sizes, with translational burst size increasing sharply ($b = 0.25 \times B^{4.77}$) with increased transcriptional burst size (Figure 3.6B). Although these data do not prove causation, this correlation

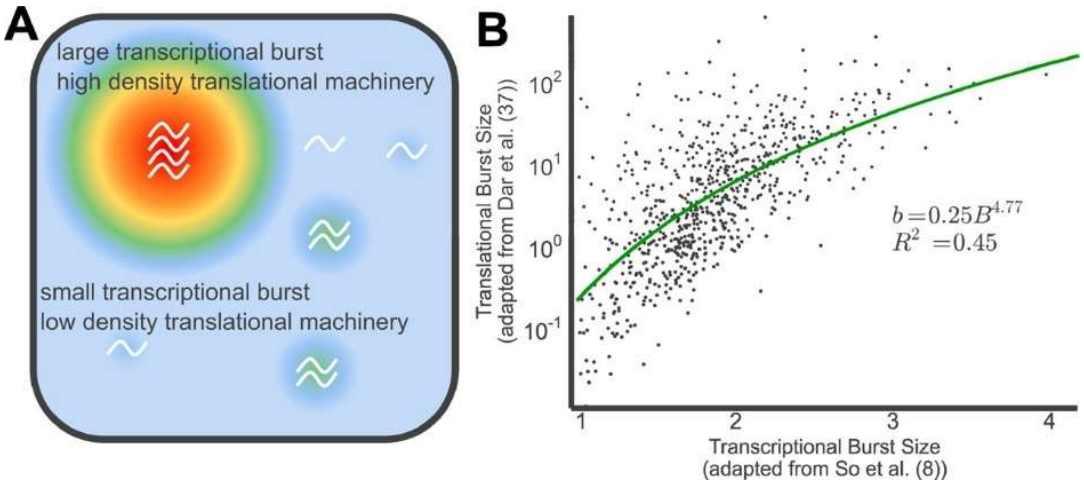


Figure 3.6: Correlation between transcriptional and translational burst sizes in *E. coli*. (a) Schematic of a large transcriptional burst sequestering a disproportionate amount of resources (heat map) leading to a large translational burst. (b) *E. coli* translational (adapted from Norred et al. 2015) vs transcriptional (adapted from So et al. 2011) burst sizes. Each data point is for an individual *E. coli* gene. The solid line is a power law fit as given by the equation in the graph.

does suggest strong cooperativity between the transcriptional and translational components of expression bursting. While the results presented here highlight the idea of transcriptional events controlling translational burst size, it is possible that there is mutual feedback, i.e., that a large transcriptional burst encourages a large translational burst, which in turn encourages an increase of the transcriptional burst size. Such mutual feedback would likely involve spatial effects such as beneficially crowding RNAP[65] or other crowding-enhanced localization of the ~100 components necessary for expression[130].

Although it is well-known that expression bursting is a ubiquitous phenomenon, little is known about the possible benefits of bursting. However, organism-scale gene expression presents a classic problem of optimizing the utility of limited shared resources. As the results presented here illustrate, expression bursting and resource sharing may be intimately linked phenomena. Expression bursting constrains a gene to draw heavily from the common resource pool over limited periods, yet it draws no resources the majority of the time. This pattern of resource sharing is reminiscent of packet mode communication[131], which allows the capacity of a shared network to be efficiently divided across a variable number of messages. Although expression bursting

is noisier than constitutive expression, it is certainly conceivable that it provides efficiency in the allocation of limited shared resources. Part of this efficiency may arise from the propensity—illustrated by the results presented here—to extend existing bursts instead of initiating new bursts. The initiation of new bursts requires the nucleation and assembly of expression machinery, while in contrast the extension of a burst leverages resources already in place. This preference for resource sharing through burst size, not burst frequency, modulation may explain recent results and analyses that have noticed burst frequency saturation in many cell types[41,43,127]. Indeed, other recent results have shown increased burst sizes in response to increased cell volume[132] or increased crowding in vesicles[52]. The picture that is emerging across multiple studies is one where protein abundance changes are preferentially mediated by modulation of burst size, not burst frequency. The results reported here clarify this picture by showing that expression resources are more readily pulled into existing bursts instead of going into the nucleation of more frequent bursts. Furthermore, these results suggest the intriguing idea that bursting is not just an artifact but instead may play a critical role in the optimal use of limited shared resources.

3.5 Methods

3.5.1 Device Fabrication

A detailed description of device design and fabrication can be found elsewhere[45,69]; briefly, clean silicon wafers were coated in SU-8 photoresist. These were exposed to two patterns: one for the control valve and one for the channel. The patterns were developed, and extra SU-8 was washed off with isopropanol. The masters were silanized with trimethylchlorosilane. PDMS base and curing agent were thoroughly mixed at 20:1 and 5:1 ratios for the membrane and control valve, respectively, and degassed under vacuum. The membrane PDMS was spin-coated at 1000 rpm for 1 min over the control valve master. The channel PDMS was poured on the channel master, and both masters were partially cured at 80 °C. Channel masters were cut out, the control valve inlet hole was punched, and the channel master was aligned and bonded to the control valve. Aligned devices were baked for 2 h at 80 °C. Devices were cut out,

inlet and outlet holes were punched, and the devices were plasma bonded to No. 0 coverslips.

3.5.2 Experimental Procedure (Chambers)

A detailed description of the experimental procedure can be found in Norred et al. 2015[45]; briefly, devices were boiled in DI water for 1 h to hydrate the PDMS. Cell-free kits were mixed according to the manufacturer's instructions (Promega S30 T7 high-yield protein expression system). One 25 µL reaction was prepared with 500 ng of pet3a T7 EGFP plasmid, 10 µL of S30 premix, and 9 µL of T7 S30 extract and was filled to 25 µL with nuclease-free water. Reagents were mixed just before the experiment, and time was recorded when reagents were mixed. The reaction was then loaded into the hydrated device with <10 psi of low-grade nitrogen. The control valve was then pressurized with ~20 psi of low-grade nitrogen. This sealed the individual chambers. The time between mixing reagents and sealing the chambers was about 4 min. The device was then placed on a Nikon Instruments Eclipse TE 300 inverted microscope and imaged through a Nikon N.A. 1.4, 100× oil objective with a Roper Scientific CoolSNAP-HQ CCD. Metamorph (Universal Imaging Corp., version 7.8.3.0) was used to capture images.

3.5.3 Experimental Procedure (Vesicles)

Vesicle preparation was adapted from Nishimura et al. 2015[51]; briefly, an inner and outer solution were prepared. The inner solution contained 1 µL of Alexa 647 (15 ng of Alexa 647 transferrin dissolved in 3 mL of water), 10 µL of PURE Solution A, 7.5 µL of PURE Solution B, 400 ng of pet3a T7 EGFP plasmid, 0.125 µL of RNAsin (40 U/µL), and 5 µL of sucrose (1 M) and was filled to 30 µL with nuclease-free water. The outer solution contained 3.6 µL of amino acid mix (50 mM), 4.9 µL of ATP (460 mM), 3.0 µL of GTP (500 mM), 1.5 µL of CTP (500 mM), 1.5 µL of UTP (500 mM), 3.6 µL of spermidine (250 mM), 7.5 µL of creatine phosphate (1 M), 9 µL of DTT (100 mM), 1.5 µL of folinic acid (4 mg/mL), 168 µL of potassium glutamate (1 M), 22.6 µL of magnesium acetate (0.5 M), 60 µL of HEPES (1 M), and 120 µL of glucose (1 M) and was filled to 600 µL with water. POPC (11.3 mg) was dissolved in 113 µL of chloroform. Of this mix, 30 µL was combined with 330 µL of paraffin oil and heated at 80 °C for 30 min. The

POPC/paraffin oil mix was removed from heat, and the inner solution was added. The mixture was vortexed for 30 s to create an emulsion. The emulsion was layered on top of the outer solution and centrifuged for 20 min at 13 200 g and 4 °C. The bottom 100 µL of outer solution and vesicle pellet was pipetted onto a Petri dish with a No. 1.5 coverslip bottom.

3.5.4 Resource Sharing Model

The resource sharing model was simulated using a Gillespie algorithm. The model consisted of a fixed pool of available ribosomes equal to 100 times the number of genes in the simulation (roughly corresponding to the number of ribosomes per plasmid in the cell-free reactions[130]). Genes stochastically bursted[100] ON and OFF with rates k_{ON} and k_{OFF} (0.0002 min^{-1}, 0.2 min^{-1}), respectively. While in the ON state, genes produced mRNA at rate α (1 min^{-1}). Ribosomes bound mRNA molecules with rate constant k_b (1 min^{-1}) and returned to the pool with rate constant k_r(0.0001 min^{-1}). The rate of protein production per bound ribosome, k_p, decayed with time, $e^{-0.05t}$, to capture the decay in synthesis capacity observed in cell-free reactions[44,47].

3.5.5 Data Acquisition and Analysis

Metamorph (Universal Imaging Corp., version 7.8.3.0) recorded images as .tif files. These files were read with Fiji (Fiji is just ImageJ, version 2.0.0-rc-14/1.49g). Images were captured by hand, so the chambers moved frame-to-frame. Images were aligned with the Fiji plugin StackReg. To make region-of-interest (ROI) placement easier, aligned stacks were averaged. ROI centers were located using an automated Hough circle finding algorithm. The averaged image was used to determine the center of the individual reaction chambers for the 5 and 10 µm chambers. For the 2 µm chambers, the exterior edges of the dark chamber walls were estimated, and the center was estimated by finding the highest local intensity value within a 30 pixel radius of the center of the found circle. Centers found were fed into a MATLAB script, which summed the intensity values of all pixels within a given radius of the center. ROI radii for every chamber in each of the three defined chamber sizes, 10, 5, and 2 µm, were 23, 10, and 5 pixels, respectively.

3.5.6 Image Processing (Vesicles)

The vesicles settled to the coverslip and were imaged at 26 °C for 3 h with a 63× confocal oil objective on a Zeiss LSM710 confocal scanning microscope. Frames were recorded every 3 min as a *z*-stack of between 25 and 35 slices that were 1 μm thick each. The *z*-stack time series files were loaded into FIJI. Vesicles were found with the plugin TrackMate[133]. Spots were filtered with an estimated diameter of 10 μm, a signal-to-noise ratio >0, and a contrast >0. Trackmate stitched together found vesicles into traces. Traces were filtered with duration >90 min, no gaps, jumps between frames <5 μm, and a total track displacement <11.9 μm. Traces that remained, and existed for the entire third hour of the experiment, were analyzed by the same noise extraction method described below.

3.5.7 Noise Extraction

The method for noise extraction was adapted from Weinberger et al. 2008[29], where it is explained in more detail. Experiments were sorted by chamber size (indexed by s = 2, 5, and 10) and day of experiment (indexed as d = 1, 2, 3, 4, and 5). General trends of fluorescence signals ($A_{s,d}$) were calculated as

$$A_{s,d}(kT) = \frac{\sum_{m=1}^{M} I_{m,s,d}(kT)}{M} \quad 3.3$$

where T is the time interval between measurements of fluorescent intensity and k = 0, 1, 2, ... K is the sample number; m = 1, ..., M represents each of the M individual chambers of a given size imaged during a given day; and $I_{m,s,d}(kT)$ is the time-dependent fluorescent intensity of an individual reaction chamber as measured by the procedure above.

Noise ($N_{m,s,d}(kT)$) was defined as

$$\boldsymbol{N_{m,s,d}(kT) \equiv I_{m,s,d}(kT) - g_{m,s,d} \cdot A_{s,d}(kT)} \quad 3.4$$

where $g_{m,s,d}$ is a gain factor that describes the extent to which the general trend coupled into each individual noise trajectory. The $g_{m,s,d}$ values were selected to minimize the cross-correlation(45)between $N_{m,s,d}(kT)$ and $A_{s,d}(kT)$. CV^2 was calculated as

$$CV^2_{m,s,d} = \frac{\sigma^2_{N_{m,s,d}}}{I^2_{m,s,d}(final)} \quad 3.5$$

where $I_{m,s,d}$(final) is the final fluorescence level measured at the end point of the experiments.

3.5.8 Chamber Combination Analysis

Two-hundred composite noise traces were created by randomly combining without replacement between two and six of the 45 individual 2 μm chambers imaged in the experiments. Composite chambers had no more than three individual 2 μm chambers in common with any other composite chamber. The fluorescent abundances of the composite chambers were found by summing the abundances of each individual chamber in the composite chamber. The variances of the composite chambers were found using the sums of the extracted noise of each individual chamber in the composite. The CV^2 of each composite chamber was defined as the composite variance divided by the composite abundance squared. The volume of composite chambers was found as the sum of the individual chambers in the composite.

3.5.9 Calculating mRNA Contributions

The reaction size of the model was varied by changing the number of genes in the system from 15 to 150 genes in increments of 15 (indexed by g = 15, 30, 45, ..., 150). Fifty trajectories (indexed as c = 1, 2, 3, ..., 50) were simulated for each reaction size. mRNA molecules were created and indexed by l = 1, 2, 3, ..., L in the order in which they were created (i.e., the first mRNA made was ranked l = 1). The number of ribosomes bound to an mRNA molecule was $R_{c,g,l,k}$, where k was the sample number (k = 0, 1, 2, ..., K). The protein production rate decayed exponentially, $k_p = e^{-0.05(kT)}$, over the duration of the experiment. The decay modeled the loss of expression capacity observed in cell-free reactions[44,47]. The total number of protein produced, $P_{c,g,l}$, from each mRNA at each sample number was calculated by summing the product of the number of bound ribosomes and the protein production rate for each minute of the simulated experiment.

$$\boldsymbol{P_{c,g,l} = \sum_{k=0}^{K} R_{c,g,l,k} e^{-0.05(kT)}} \quad 3.6$$

where T is the interval between samples. The average number of protein produced by mRNA molecules of each rank was calculated as

$$\boldsymbol{P_{g,l} = \frac{1}{C} \sum_{c=1}^{C} P_{c,g,l}} \quad 3.7$$

The average protein population associated with an individual mRNA molecule was normalized by the total amount of protein produced and was calculated as

$$\boldsymbol{P_{NORM_{g,l}} = \frac{P_{g,l}}{\sum_{l=1}^{L} P_{g,l}}} \quad 3.8$$

This normalization, $P_{\mathrm{NORM}g,l}$, when plotted against mRNA rank, j, illustrated the relative influence of mRNA rank on the final protein population.

3.6 Appendix

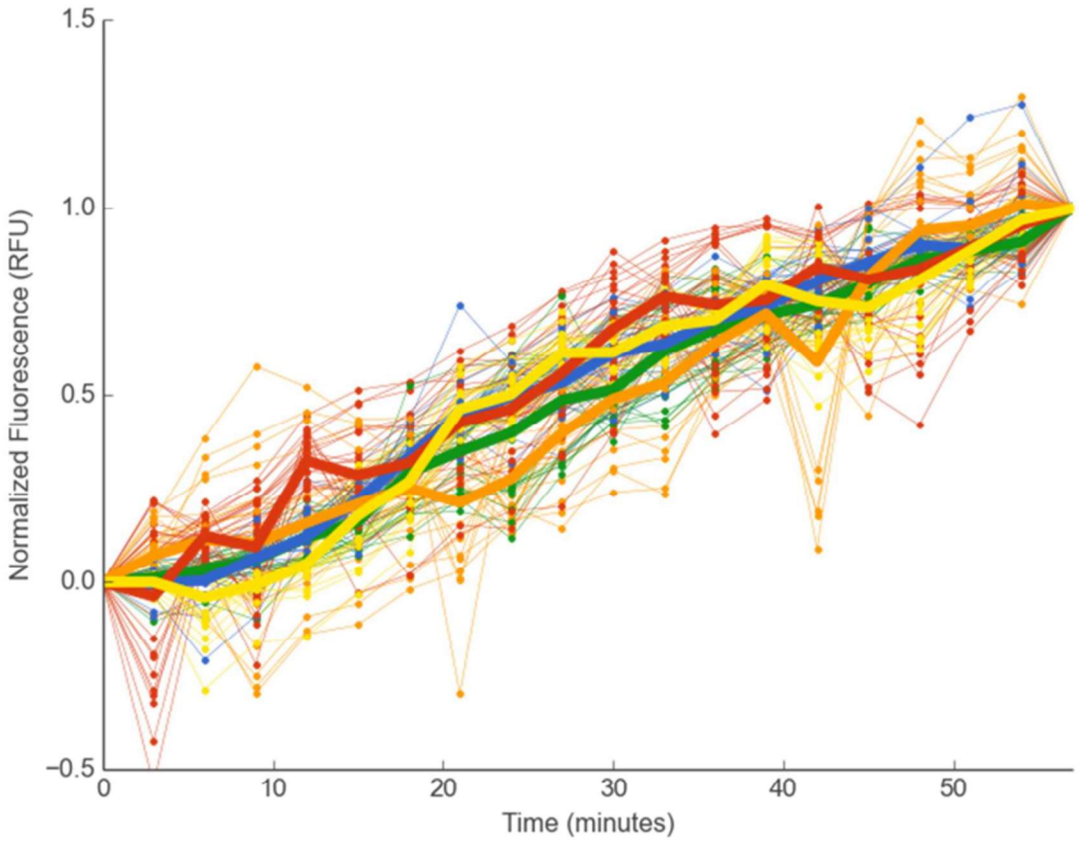


Figure 3.7: Comparison of transient behavior between different experimental days. Each color shows the average normalized fluorescence transient on a particular experimental day. Thin lines are individual chambers while thick lines are averages of chambers of all sizes acquired in the same experimental session. Each experiment had a similar 50% rise time (~25 minutes) and GFP production rate.

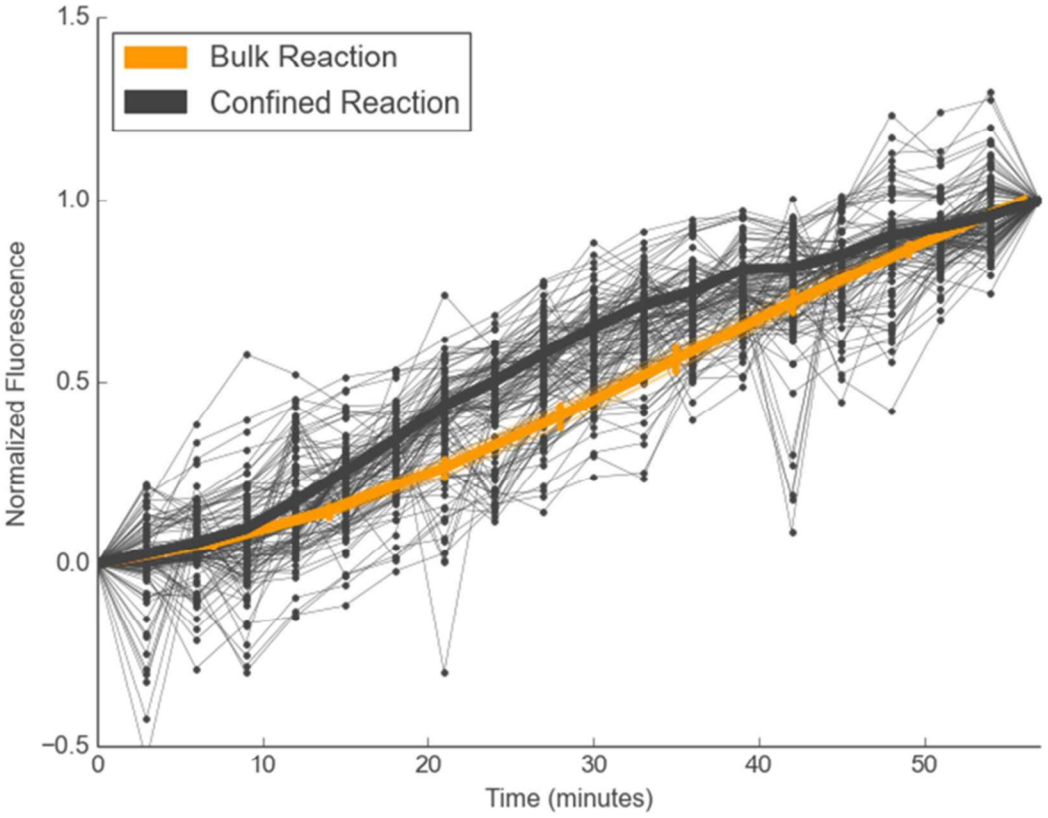


Figure 3.8: Comparison of fluorescence transients in unconfined (i.e. bulk) and confined chambers. Bulk reactions were run in a BioTek Synergy 2 plate reader. 25 µL, commercial, raw extract, cell-free protein synthesis (CFPS) reactions were mixed with 500ng of Enhanced Green Fluorescent Protein (EGFP) coding pET3a plasmid, and pipetted into a 384 well plate. Reactions were covered with 10 µL of mineral oil to prevent evaporation. Excitation was at 485 nm and fluorescence was measured at 528 nm every 7 minutes for 1 hour.

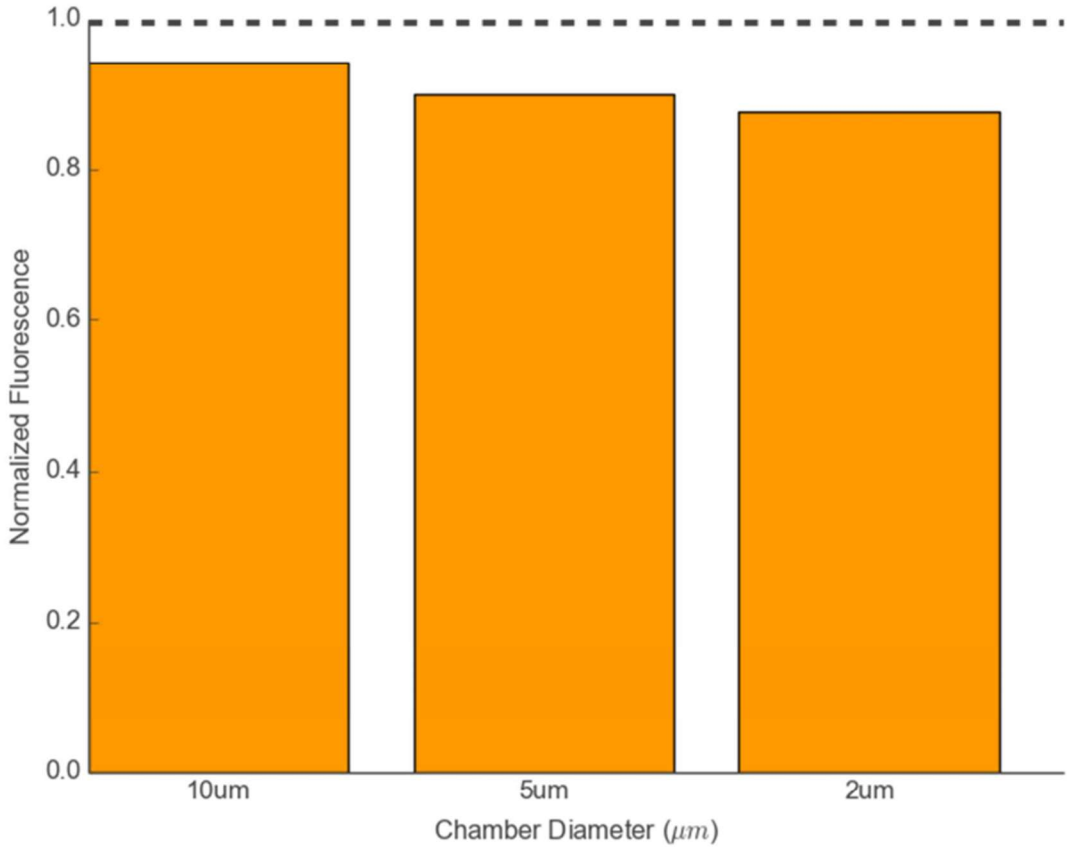


Figure 3.9: Photobleaching during imaging. To measure the effect of photobleaching occurring during experiments we ran a standard experiment for one hour with the aperture narrowed to only the 9 chambers in the field of view. At the end of the experiment we compared the intensity of the 9 chambers in the field of view to the intensity of 9 chambers of the same size, on the same chip, during the same reaction that had not been exposed. Fluorescence intensity was only slightly reduced (by about 10%, dashed gray line vs. bars) and was reduced by similar amounts for all chamber sizes.

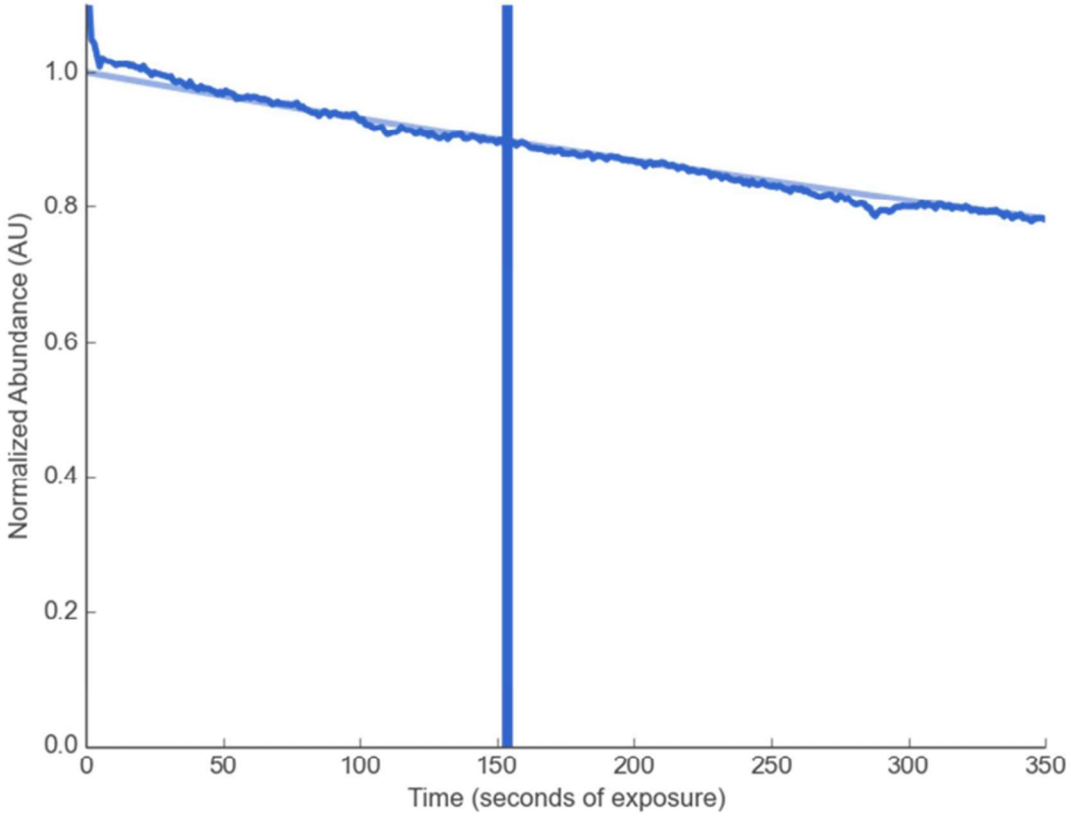


Figure 3.10: Rate of photobleaching. To measure the photobleaching rate in the system we imaged confined reactions in a standard experiment for one hour. We then continuously exposed the reaction to the light source (Nikon Intensilight C-HGFI) for 10 minutes while recording an image every second. The average intensity of 9 chambers was recorded for each time point and plotted against exposure time. Fluorescent intensity was normalized to the initial value and fit with an exponential decay curve. The photobleaching half-life was measured to be 987.47 seconds. The vertical line corresponds to the measured amount of photobleaching during a one-hour experiment, 153 seconds of exposure and a 10% reduction in fluorescence compared to chambers not imaged.

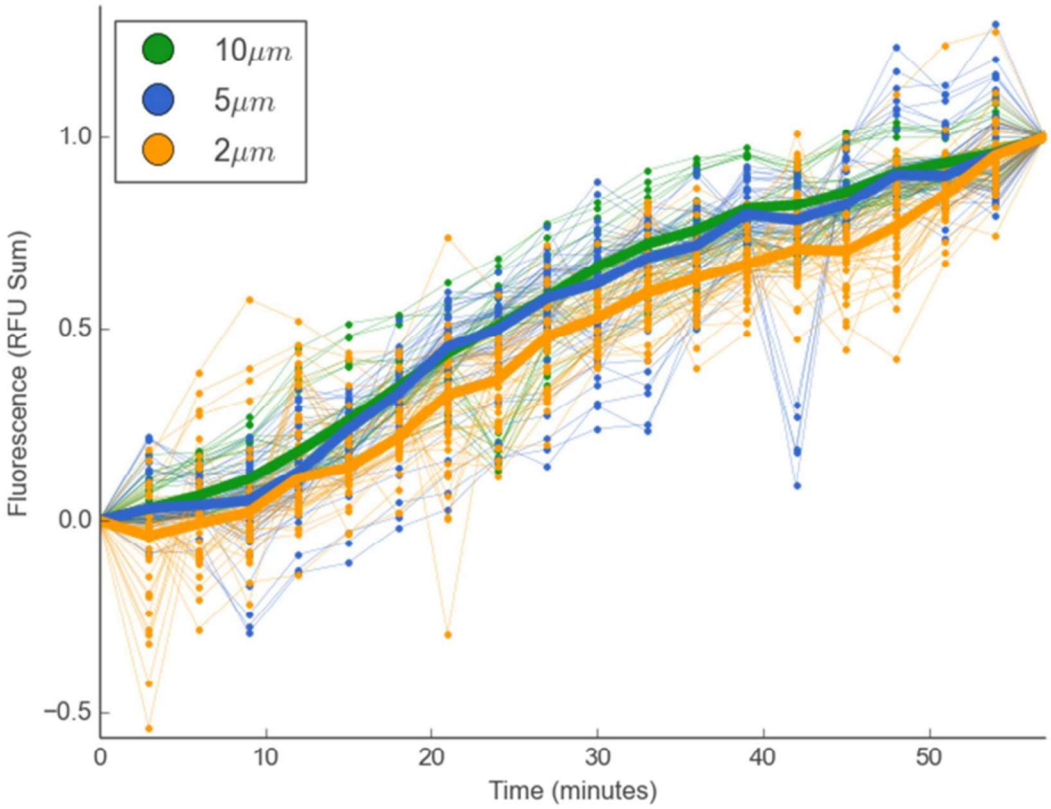


Figure 3.11: Effect of chamber size on fluorescence transient. Normalized fluorescent traces show the transients of different chamber sizes exhibit very similar behavior. Thick lines represent the average behavior of all chambers of a given size, while the thin lines depict individual chamber behavior. Both the average transient behavior and the distributions of the individual behaviors show little variation across the different chamber sizes.

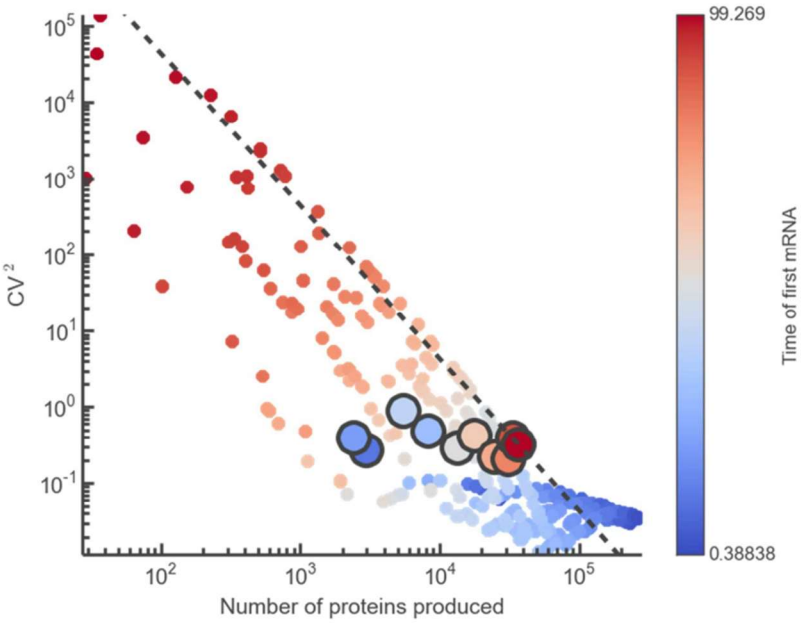


Figure 3.12: Noise plot of simulation data colored by the time the first mRNA was made. Centroids are colored by reaction size (blue 15 genes; red 150 genes). The dashed gray line has the form a/abundance^2. The timing of mRNA production heavily influences both the amount of protein produced and the noise.

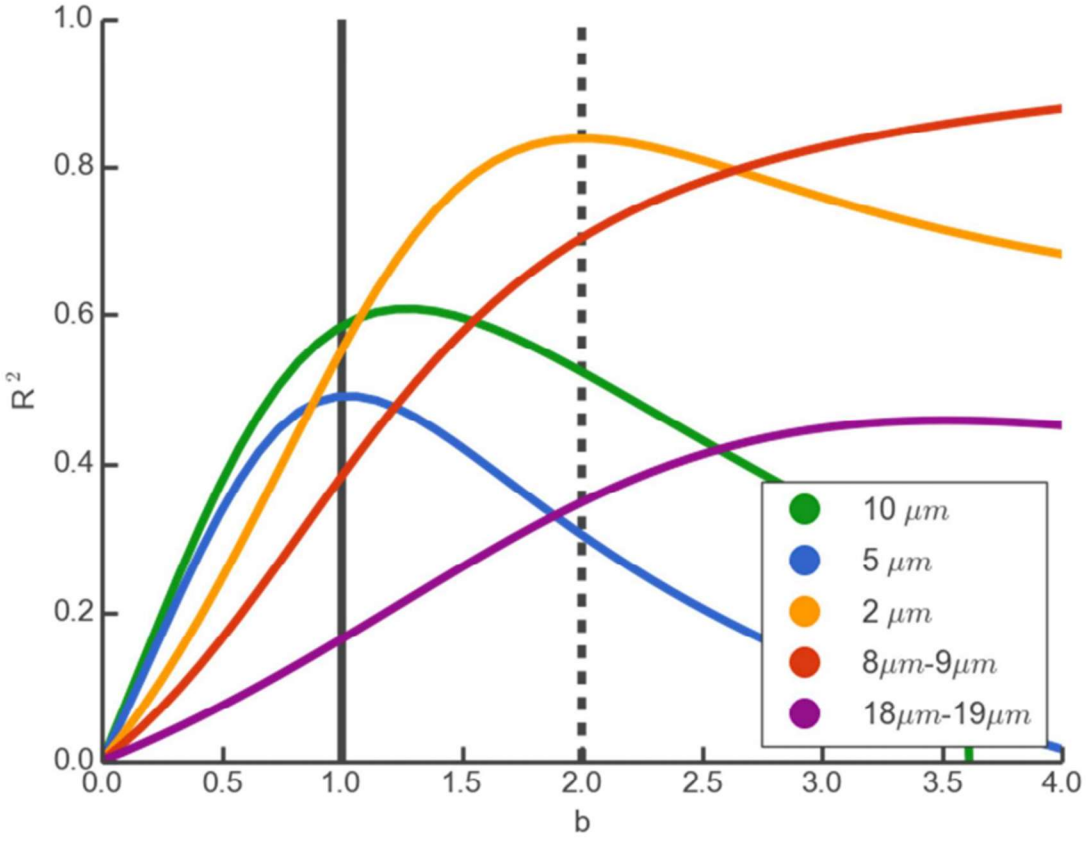


Figure 3.13: R^2 for varying power law fits to experimental data. Chamber and vesicle noise plot data (**Figure 3.3**A and C) were fit with equations of the form a*abundance^{-b}. Values of b ranged between 0.0 and 4.0. For each value of b, the value of a was found by nonlinear least squared fitting. R^2 was measured and plotted for each value of b. Values of b between 1.0 (solid line) and 2.0 (dashed line) show good fits to the experimental data in both chambers and vesicles.

4 MACROMOLECULAR CROWDING INDUCES SPATIAL CORRELATIONS THAT CONTROL GENE EXPRESSION BURSTING PATTERNS

A version of this chapter was originally published by S. Elizabeth Norred, Patrick M. Caveney, Gaurav Chauhan, Lauren K. Collier, C. Patrick Collier, Steven M. Abel, and Michael L. Simpson:

S. Elizabeth Norred, Patrick M. Caveney, Gaurav Chauhan, Lauren K. Collier, C. Patrick Collier, Steven M. Abel, and Michael L. Simpson
ACS Synthetic Biology **2018** *7* (5), 1251-1258 DOI: 10.1021/acssynbio.8b00139

This chapter has been adapted from its published format to accommodate new Figure, Table, and Equation enumeration. The Supplementary Information associated with this work may be found in the Appendix; all references are located at the end of the document. SEN, PMC, CPC, and MLS conceived and planned the experiments. Experiments were performed by SEN, LKC, and PMC. SEN designed and constructed the pRSETb-mCherry-Spinach2 plasmid with the assistance of Drs. Jennifer Morrell-Falvey and Amber Bible. The Spinach2 sequence was acquired, with thanks, from the Jaffery Lab website, accessed in 2014[134,135]. SEN, GC, SMA, and MLS analyzed the results and programmed simulations. SEN, GC, SMA, and MLS wrote the manuscript.

4.1 Abstract

Recent superresolution microscopy studies in *E. coli* demonstrate that the cytoplasm has highly variable local concentrations where macromolecular crowding plays a central role in establishing membrane-less compartmentalization. This spatial inhomogeneity significantly influences molecular transport and association processes central to gene expression. Yet, little is known about how macromolecular crowding influences gene expression bursting – the episodic process where mRNA and proteins are produced in bursts. Here, we simultaneously measured mRNA and protein reporters in cell-free systems, showing that macromolecular crowding decoupled the well-known relationship between fluctuations in the protein population (noise) and mRNA population statistics. Crowded environments led to a 10-fold increase in protein noise even though there were only modest changes in the mRNA population and fluctuations. Instead, cell-like macromolecular crowding created an inhomogeneous spatial distribution of mRNA ("spatial noise") that led to large variability in the protein production burst size. As a result, the mRNA spatial noise created large temporal fluctuations in the protein

population. These results highlight the interplay between macromolecular crowding, spatial inhomogeneities, and the resulting dynamics of gene expression, and provide insights into using these organizational principles in both cell-based and cell-free synthetic biology.

4.2 Introduction

The intracellular environment is highly crowded by proteins and other macromolecules which occupy a significant fraction of the cell (e.g. approximately 30% in *E. coli*)[24,136,137]. In contrast to the dilute environments often used to determine biochemical parameters, macromolecular crowding affects the dynamics of molecular transport and interactions in ways that significantly alter molecular association, ligand binding, and protein folding[24]. Likewise, macromolecular crowding is known to differentially alter average transcription and translation rates[65,138]. Yet, the relationship between macromolecular crowding and gene expression bursting patterns remains largely unexplored.

Gene expression bursting is an episodic process where mRNA and proteins are produced in relatively short bursts separated by periods of no expression[100,112,118]. Expression bursting is often associated with molecular processes such as transcription factor binding kinetics[46,110,111], supercoiling[30], and transcriptional re-initiation[30,37,117], that would seem likely to be influenced by macromolecular crowding. In particular, crowding is likely to modulate burst dynamics (Figure 4.1), which are characterized by burst size (average number of mRNA or protein molecules created per burst) and burst frequency (number of bursts per unit time). Burst size and burst frequency are modulated in response to numerous physical and chemical cues[132,139], and macromolecular crowding could affect expression bursting through inhibited diffusion, physical exclusion from regions of space, and extended co-localization of interacting molecules (e.g. ribosome and mRNA; Figure 4.1).

The effect of exclusion and co-localization on bursting depends on the size of the reacting molecules and the complexity of the reactions. Larger molecules are more readily excluded than smaller molecules[24], while co-localization of only a few large reacting molecules occurs more readily than ones involving several large components.

Accordingly, it seems likely that transcriptional and translational bursting respond

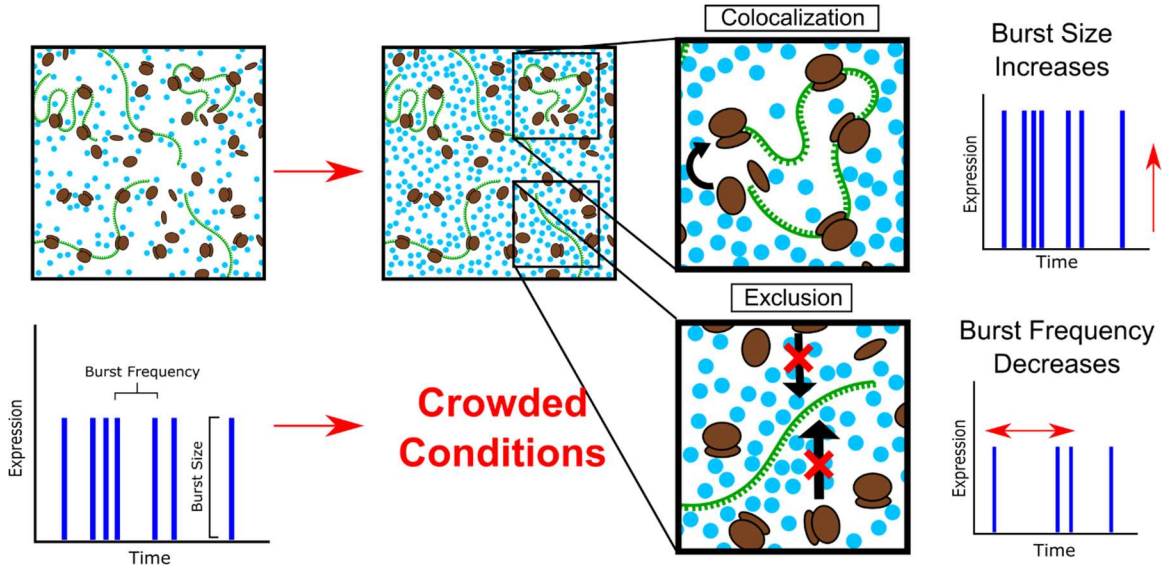


Figure 4.1:: Gene expression happens in an episodic process characterized by a burst size (average number of molecules synthesized during a burst) and a burst frequency (the average rate of burst occurrence). Crowding affects molecular transport by making diffusion to some reaction sites difficult (exclusion) and keeping nearby reactants co-localized. Exclusion seems likely to lower burst frequency, while co-localization seems likely to increase burst size.

differently to changing levels of macromolecular crowding. For example, in cell-free expression T7 transcription requires only the co-localization of two large components (the polymerase and promoter), but translation requires the co-localization of many large structural proteins, as well as the availability of aminoacylated tRNAs in the proper amounts and sequence. Furthermore, ribosome mobility in crowded environments is significantly lower than that of the ~3x smaller polymerase[24,140,141]. These different effects of crowding are intertwined and, in some cases, lead to opposing behavior. As a result, little is known about how the two components of expression bursting (transcription and translation) are differentially modulated in response to macromolecular crowding.

Here we show how macromolecular crowding drives expression bursting patterns in cell-free gene expression, and specifically, how crowding affects the transcriptional and translational components of bursting in very different ways. These experiments feature the simultaneous measurement of the intensity and fluctuations of fluorescent reporters of transcriptional and translational events driven from the same promoter. Total expression (transcription and translation taken together), burst size and burst frequency were modulated by crowding, but in opposite directions – i.e. crowding increased burst

size, but decreased burst frequency. The transcriptional component of bursting was only mildly affected by crowding. In sharp contrast, the translational component of bursting was very sensitive to changes in the crowding level, responding to greater crowding with a large decrease in burst frequency, but a large increase in burst size. These results are consistent with a crowded model consisting of spatially distinct regions with widely varying populations of expression resources (e.g. ribosomes) that control the burst size[34]. Likewise, recent experimental results have demonstrated such subregions in *E. coli* that direct the location and rates of transcription and translation[142], and in eukaryotic cells such subregions may be created through regulated phase transitions[143]. The results here suggest that such spatial features in cells may be responsible for controlling gene expression burst patterns, which may play important roles in the sharing of global gene expression resources[34] and in making translation more variable than transcription[144]. These burst patterns play a central role in setting the low noise limit in gene expression[127], and these noise constraints have consequence for quick adaptation to stress and survival in uncertain environments[145,146].

4.3 Results and Discussion

In order to examine transcription and translation simultaneously, we inserted a sequence coding for a stabilized aptamer, Spinach2[134], into a plasmid coding for a fluorescent red protein, mCherry (Figure 4.2A). Fluorescence readings for both reporters were recorded over 12 hours (Figure 4.2C) at crowding fractions ranging from 0-30% (Figure 4.2D and Methods). These crowding levels were selected to mimic cellular environments where the high total concentration of macromolecules produces crowding between 5-40%[21] .

Expression bursting was investigated by extracting the noise behavior of the two reporters from the expression trajectories (Figure 4.2C) using a previously described method[29,34,43,44]. In agreement with previous reports [65], the protein expression rate was strongly repressed by macromolecular crowding, and was completely repressed at a crowding level of 30% (Figure 4.3A). In contrast, mRNA expression was only mildly responsive to crowding and displayed only small decreases in expression level as the crowding fraction was increased. However, at the highest crowding fraction studied (30%) there was a large decrease in the rate of mRNA synthesis, but synthesis

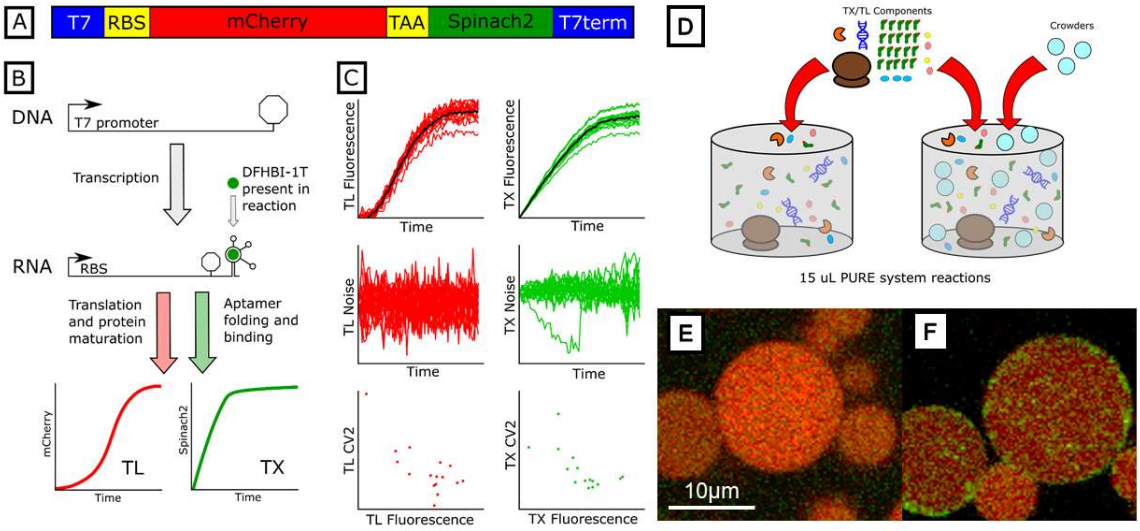


Figure 4.2: Cell-free tracking of mRNA and Protein expression. A) Gene layout for pRSET-b-mCherry-Spinach2 used in cell-free experiments. B) The RNA aptamer, Spinach2, folds and binds to a fluorophore (DFHBI-1T) which fluoresces at 495 nm. The mRNA upstream of the aptamer codes for the mCherry reporter protein that fluoresces at 630 nm. (C) Fluorescence from both reporters was tracked over time for noise analysis. D) Cell-free expression (PURExpress kit) was in microplate wells (15uL) in the presence of DFHBI-1T was measured every 5 minutes for 6 hours. Different crowding levels were achieved by the addition of Ficoll-70 to the cell-free reactions. (E) and (F) The spatial distribution of mRNA and protein was visualized by confocal microscopy imaging of cell-free reaction in liposomes at (E) 0% and (F) 5% crowding fractions. In (E) both protein and mRNA are distributed evenly throughout the volume of the liposome. In (F) protein is distributed throughout the volume of the liposome, but mRNA exhibits a more inhomogeneous distribution.

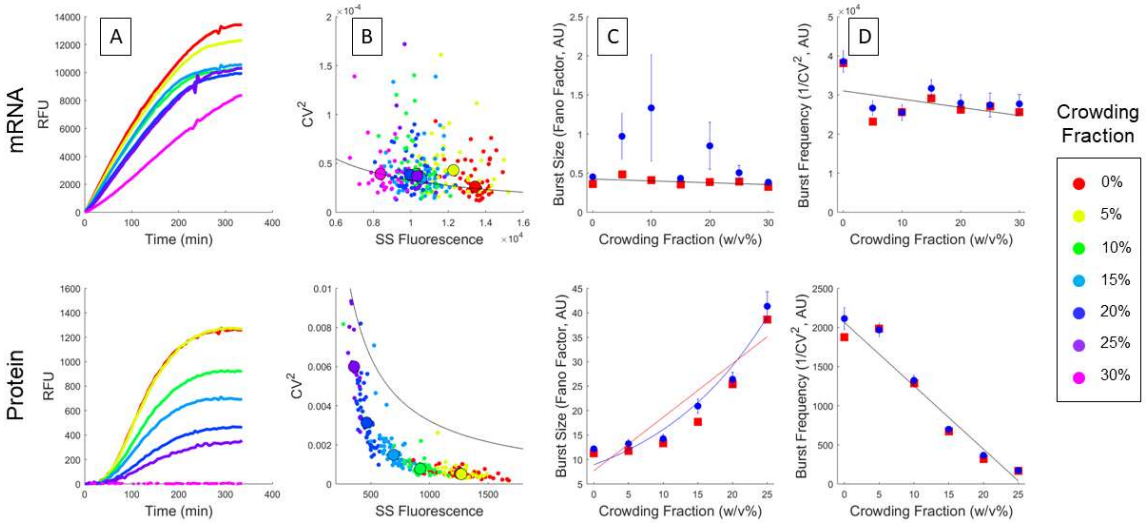


Figure 4.3: Average expression time course (A) and CV^2 vs. steady-state (SS) fluorescence (B) for mRNA and protein at crowding fractions ranging from 0 to 30%. The small data points in (B) represent individual experiments, while the large data points are averages of all experiments at the same crowding fraction. The black lines represent the Poisson scaling trend ($\boldsymbol{CV^2} \propto \frac{\boldsymbol{1}}{\boldsymbol{SS}}$). (C) Mean (blue dots) and median (red square) burst sizes (Fano factor) vs. crowding fraction extracted from mRNA and protein noise measurements. Error bars indicate standard error of the mean. The means of some crowding fractions in the mRNA measurements are pulled up by a small number of high outlier values. The trend line for the mRNA data is a linear fit to the medians. The trend lines shown for the protein graph show assumptions about the relationship between the variance of the burst size and the crowding fraction (CF). The red line assumes a linear relationship ($\boldsymbol{\sigma^2} \propto \boldsymbol{CF}$), and the blue line assumes an exponential relationship ($\boldsymbol{\sigma^2} \propto \boldsymbol{e^{CF}}$). (D) Mean and median burst frequency ($1/CV^2$) for mRNA and protein. The trend lines for both mRNA and protein are linear fits to the medians.

continued to occur for a longer duration (Figure 4.3A). There was a striking difference in the noise behaviors of the mRNA and protein populations in response to crowding-induced changes in expression. Although median mRNA noise magnitude was only mildly sensitive to the size of the mRNA population (number of molecules), protein noise was hypersensitive (i.e. increased faster than Poisson (1/population) scaling) to changes in the protein population (Figure 4.3B). This difference in noise behavior indicated that crowding resulted in very different transcriptional and translational burst behaviors.

Previous analysis and experimentation show that $CV^2 = \frac{B}{\mu}$, where B is the burst size and μ is the mean value of the molecular population[43,99,100,104,125,127,147,148]. Since the Fano factor is equal to μCV^2, it is a direct measure of the burst size. Accordingly, to determine how transcriptional and translational burst sizes varied with crowding, we examined the Fano factors (Methods) of mRNA and protein populations (Figure 4.3C). The transcriptional burst size was insensitive to changes in crowding fraction. There was no statistically significant change in transcriptional burst size across all crowding fractions as determined by one-way ANOVA ($F(6,281) = 1.08$, $p = .38$). In contrast, there was a statistically significant change in protein Fano factor as determined by one-way ANOVA ($F(5,250)=50.04$, $p<.0001$) as the Fano factor increased approximately three-fold as the crowding fraction increased from 0 to 25% (Figure 4.3C). The protein Fano factor could not be measured for the 30% crowding fraction as no protein production was observed for this condition.

To determine how burst frequency varied with crowding, we examined $1/CV^2$ of mRNA and protein populations, which are proportional to the frequency of transcriptional and translational bursts, respectively (Figure 4.3D, also see Appendix). The transcriptional burst frequency had a statistically significant change across all crowding fractions as determined by one-way ANOVA ($F(6,281)=4.1$, $p=0.0006$), but most of this drop was observed between 0% and 5% crowding levels (Figure 4.3D). In contrast, there was a sharp, monotonic decrease in translation burst frequency (Figure 4.3D) and a statistically significant change in translational burst frequency as determined by one-way ANOVA ($F(5, 250)=84.07$, $p<.0001$). The translational burst frequency at a 30% crowding fraction was essentially 0 as no protein synthesis was observed under these conditions.

Although crowding had only a small effect on the average transcriptional behavior, it did lead to a higher propensity for "outlier" transcriptional behavior. At many crowding fractions – and in particular, at the moderate crowding levels – the median and mean Fano Factors were significantly different (Figure 4.3C), indicating that the mean was being pulled up by a small number of traces with very high Fano factors. These outliers were easily identified by inspection of the noise traces, where they often exhibit extended periods of high expression (See Appendix, Figure 4.5). A likely explanation for these transcriptional outliers is repeated binding of RNAP and promoters that have been entrapped together by crowding induced compartments. Although crowding increases the mean time for an initial reaction, it also enhances the possibility for repeated rebinding and may even increase some reaction rates[149] . Such re-initiation events have been shown to increase transcriptional burst size and noise[37]. These outliers were rare events in these experiments, making it impossible to determine the actual likelihood of outlier behavior. As a result, we focused our attention on the "bulk" behavior exhibited by the majority of the experimental runs.

An earlier study[52] of the effects of crowding on expression noise provides a foundation for the results reported here. The important advance reported here came from the simultaneous observation of reporters of both transcriptional and total expression activity under the control of the same promoter and within the same environment. This additional level of detail allowed the unambiguous observation of the individual contributions of transcription and translation to expression burst behavior.

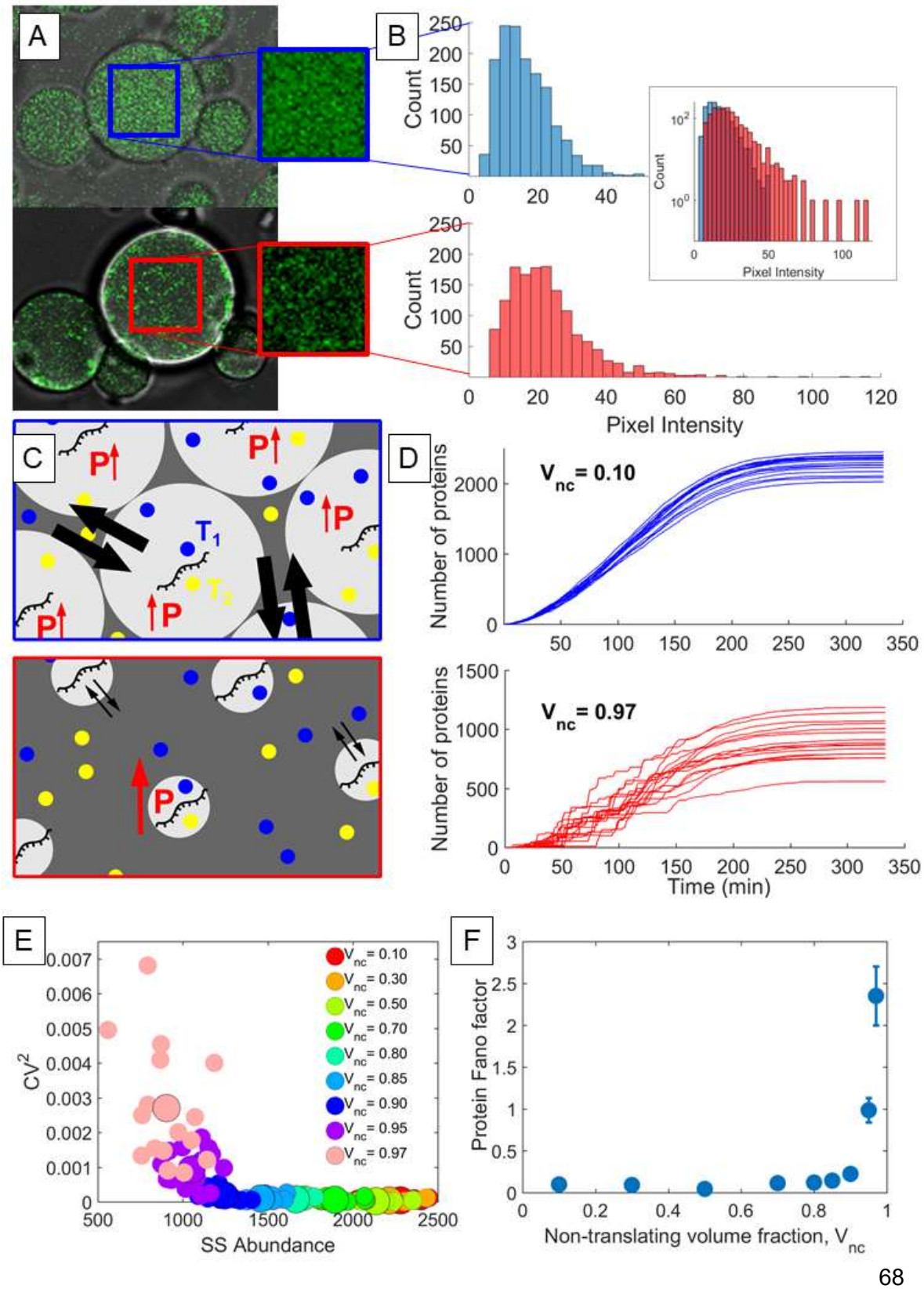
Starting with the mRNA measurements, crowding levels of 25% or lower produced small changes in transcriptional behavior. For all crowding levels, mRNA synthesis occurred for at least 200 minutes and the mRNA population increased at a nearly constant rate before a fairly abrupt cessation of mRNA synthesis (Figure 4.3A and Figure 4.6 in Appendix). The most noticeable effect was a drop of about 35% in the final mRNA population as crowding increased from 0 to 25% (Figure 4.3A). As crowding increased from 25% to 30%, the rate of mRNA production dropped by more than 20%, yet transcription remained active for at least 100 minutes longer (Figure 4.3A). This behavior was observed in conjunction with complete loss of translational activity. Previous studies have shown that the elongation rate of transcription by bacterial polymerases is highly correlated with the rate of translation[150], and our measurements

similarly suggest that the T7 polymerase elongation rate drops in the absence of translation. Furthermore, the absence of translation reduced the energy demand of expression, likely leading to the extended duration of transcription. These observations illustrate the importance of the simultaneous measurement of both transcriptional and translational activity, as transcriptional behavior may change substantially in the absence of translation.

Surprisingly, as crowding increased from 0 to 25%, protein noise changed by more than 10-fold with only a 35% change in the mRNA population and only small changes in mRNA noise (Figure 4.3B). This indicated that mRNA population statistics played very little role in total expression noise. As a possible explanation for the large changes in protein noise behavior, we explored how crowding affected the mRNA spatial distribution. Hansen *et al.* reported evidence of the formation of pockets of mRNA in transcription-only experiments[52]. To examine the spatial distribution of mRNA in a full expression experiment, we performed fluorescence microscopy imaging of cell-free reactions in vesicles (Appendix, Figure 4.7) with and without added crowders (Figure 4.4A-B). The most striking effect of crowding was the emergence of distinct regions with very high mRNA populations (Figure 4.4A-B). These mRNA-rich regions created a high-end tail in the mRNA spatial distribution that is most clearly visible in the log-scale distribution (Figure 4.4B, inset). While this change to the mRNA spatial distribution only modestly affected mRNA population and fluctuation behavior, it led to significant changes in the protein expression behavior. Specifically, increased crowding led to less frequent but more intense translational bursting (Figure 4.4C-D). Since there was no associated drop in the mRNA burst frequency, less frequent translational bursting implies that much of the mRNA was translationally inactive at the higher crowding level. This finding is similar to results we recently reported for gene expression under different levels of spatial confinement[34]. In that study, less confinement (bigger reaction chamber volume) resulted in a sub-population of mRNA that was translationally inactive due to unequal sharing of translational resources.

It is an intriguing result that in our experiments 30% crowding was enough to silence translation completely, yet translation continues at higher crowding fractions in *E. coli* and other cell types. Continued translation at high crowding levels is likely the result of inhomogeneous crowding in cells where there is little translation in very crowded

Figure 4.4 Spatial Inhomogeneity in mRNA and Simulation results (A) Distribution of mRNA in vesicles expressing pRSET-b-mCherry-Spinach2 at 0% (Top) and 5% (Bottom) crowding fractions. The blue and red boxes denote 1440 pixel regions of interest (ROIs) that were analyzed to understand how crowding affected the spatial distribution of mRNA. (B) Histogram of pixel intensities for the two ROIs from (A). The inset shows the two histograms together using a log scale for the y-axis to illustrate the long tail distribution created by crowding. (C) Schematic of physical gene expression model with low crowding (Top) and high crowding (Bottom). The model simulates a crowding-controlled distribution of accessible regions for gene expression. Two factors, T1 and T2, must be present in the region for translation to occur. As crowding increases, the volume of the regions decreases and the diffusion rate (indicated by the thickness of the black arrows) into and out of the region decreases. Translational activity in a region is indicated by red arrow. D) Representative protein expression trajectories for low crowding (Top) and high crowding (Bottom). The crowding level in the model is described by VNC, which represents the fraction of volume outside the expression regions. High crowding level (VNC =0.97, red) trajectories are more variable than low crowding level trajectories (VNC =0.10, blue). E) Noise analysis of trajectories colored by VNC. As VNC increases, protein abundance decreases while CV2 increases. F) The Fano factor of protein expression increases as VNC increases, indicating an increase in burst size.


A
B
Count
Pixel Intensity
C
P
T1
T2
D
Vnc = 0.10
Vnc = 0.97
Number of proteins
Time (min)
E
CV2
SS Abundance
Vnc = 0.10
Vnc = 0.30
Vnc = 0.50
Vnc = 0.70
Vnc = 0.80
Vnc = 0.85
Vnc = 0.90
Vnc = 0.95
Vnc = 0.97
F
Protein Fano factor
Non-translating volume fraction, Vnc

regions (e.g. the nucleoid), but much greater translational activity in less crowded ribosome-rich regions[142].

We hypothesized that crowding leads to unequal sharing of translational resources by parsing the mRNA into spatially distinct, translationally-favored and translationally-disfavored sub-populations. To explore how crowding-induced spatial correlations of mRNA affected expression, we developed a physically-motivated model where transcription and translation occurred in compartmentalized regions of space (Figure 4.4C). In this model, increased crowding reduced the size of the compartmentalized regions and the transport of molecules to and from the compartments[151]. The translational step was modeled by requiring the assembly of two components (referred to generically as T1 and T2) for translation to occur (Figure 4.4C). In this model, higher crowding (smaller compartmentalized regions) led to reduced protein population (Figure 4.4D and E), higher protein CV^2 (lower burst frequency**;** Figure 4.4E), and higher protein Fano factor (higher protein burst size; Figure 4.4F) much like the effects observed in the experiments.

These crowding induced effects in the model were primarily due to wide variability in the size and timing of the translational bursts. Importantly, both increased translational burst size and larger variability in the burst size lead to greater noise in the protein population, with[127]

$$\boldsymbol{FF_p = \bar{B} + \frac{\sigma_B^2}{\bar{B}}}, \qquad 4.1$$

where FF_p is the Fano factor of the protein, and $\bar{B}$ and σ_B^2 are the mean and the variance of the translational burst size, respectively (Appendix, 4.5.1). To study explicitly the effect of translational burst size variability on protein noise, we constructed a simplified two-state expression model (Appendix, 4.5.2) that captured the variability of the translational burst size predicted by the model described above. This two-state model is like the traditional two-state model of expression with one important exception: the average translational burst size remained constant, but the variance of the burst size increased as crowding increased (Appendix Figure 4.8). Gillespie simulations of this model were performed where the variance of the burst size varied over a 25X range, and demonstrate that large variability in the translational burst size – not just changes in the

mean burst size – can generate the crowding-induced changes in protein noise behavior measured here (Appendix Figure 4.8).

The results here demonstrate that spatially inhomogeneous distribution of molecular populations plays an essential role in controlling expression bursting behavior in crowded cell-free systems. Cell-like macromolecular crowding creates an inhomogeneous spatial distribution of mRNA. This spatial "noise" in turn controls expression burst behavior and becomes ingrained in the temporal noise of the protein population. This spatial arrangement in cell-free systems is similar to that found in *E. coli*. Recent superresolution microscopy of *E. coli* shows spatial segregation of the nucleoid and the bulk of the ribosome population[142]. This spatial arrangement in cells affects both transcription and translation, but similar to the results reported here, in differing ways. RNA polymerase can explore the crowded inner nucleoid region, and low levels of transcription can occur throughout the nucleoid[152]. However, ~85% of the ribosome population is excluded from the nucleoid region[142]. The most active region of expression bursting seems to be concentrated at the interface between the nucleoid and ribosome-rich region where dense clusters of transcribing RNAPs congregate[152]. Intriguingly, dynamic reordering of the spatial organization appears to be a mechanism for expression rate regulation in cells[152], and may be a viable strategy in cell-free systems that make use of advanced nanofabricated platforms[153].

It is important to point out differences between the crowded environment in cells and those reported here for the cell-free experiments. For prokaryotic cells such as *E. coli*, water makes up approximately 50-70% of an *E. coli*, while the dry weight is ~55% protein, ~20% RNA, ~10% Lipids, and ~15% of other molecules[154]. Since a large fraction of the protein and RNA is ribosomal[155], ribosomes (~20 nm in diameter in prokaryotes) make a significant contribution to cellular crowding. Most of the other proteins have a globular configuration with radii in the 3-6 nm range[156]. Ficoll 70 (stokes radius of ~5.1 nm) and the cell free expression media used here resemble the distribution of globular protein and ribosome crowders in cells. Conversely, these cell-free experiments do not mimic extended structures like cytoskeletal filaments or elongated proteins, nor larger structures like organelles or the bacterial nucleoid. These extended structures in cells may affect expression bursting in important ways by allowing for facilitated transport along filaments or by creating inhomogeneous crowding (e.g. very crowded nucleoid

region with less crowded ribosome rich areas[142]) that are not well approximated by the experimental platform reported here. While eukaryotic cells have crowded environments like that described for prokaryotic cells, they completely decouple transcription and translation. The coupling of transcription and translation is likely to affect expression burst behavior by modulating the rate of transcription (described above) and by affecting local crowding around the promoter. Accordingly, the results presented here should be used cautiously with regards to eukaryotic transcriptional burst dynamics where promoter structure (e.g. TATA boxes) or nucleosome occupancy patterns play prominent roles in burst behavior[124] not seen in prokaryotic cells.

The vast majority of synthetic biology efforts focus on gene circuit design that relies on specific molecular mechanisms – e.g. transcriptional control with regulatory proteins – and engineering principles of circuit and network design[16]. But even very well characterized gene circuit elements, like the lac promoter, may be more sensitive to location or local concentrations than to specific molecular interactions[157]. Spatial effects on the regulation of gene expression provide both significant challenges and opportunities for cell-free synthetic biology. Chief among the challenges is a limited understanding of how the spatial distribution of molecular populations may be used to design specific functionality. However, cell-free platforms provide the means to vary spatial arrangements in intricate ways[153], and may be especially well-suited for developing a deeper understanding of spatial synthetic biology.

4.4 Methods

The Spinach2 sequence, which fluoresces after hybridizing with a fluorophore, DFHBI-1T (Lucerna, Inc), was inserted 30 bp downstream of the mCherry stop codon and upstream of the T7 terminator (Figure 4.2A). This placement ensured that all fluorescing gene transcripts contained a complete transcript of the mCherry protein (Figure 4.2B). The plasmid pRSET-b-mCherry-Spinach2 was expressed in a scaled-up PURExpress (NEB) cell-free protein synthesis reaction in the presence of DFHBI-1T. The reaction was divided into 15μL aliquots. For each 15μL reaction, fluorescence readings for Spinach2-DFHBI-1T and mCherry were performed in a microplate reader (Perkin-Elmer, EnSpire Multimode Plate Reader) every 5 minutes for 12 hours to generate expression traces for Spinach2-DFHBI-1T and mCherry fluorescence (Figure

4.2B). All fluorescence traces were zeroed after the first 5 time points in order to exclude initial readings due to autofluorescence. Additionally, all traces were truncated after 72 timepoints, after which all traces reached a maximum level (Appendix Figure 4.9). This maximum level was indicative of the end of expression due to resource limitation, and was not an equilibrium between synthesis and decay[44,47].

To simulate cell-like macromolecular crowding, concentrated Ficoll-70 (Sigma-Aldrich) was added to the cell-free reactions in order to create crowding fractions ranging from 0-30 w/v% (in previous sections and hereafter referred to as % crowding). Concentrated Ficoll-70 (Sigma-Aldrich) was prepared in nuclease free water by adding 375 mg Ficoll-70 to 500µL of nuclease-free water, to create a weight volume percent of approximately 75 w/v%. This was weight/volume percent of the mass of the solute to the volume of the solvent instead of the mass of the solute to the volume of the solution. The crowder solution in nuclease free water was applied to the PURE system reactants in lieu of the water used to dilute all reactants to a 30µL volume. This means all experiments of differing crowding fractions had the same amount of proteins and cell-free reactants.

Ficoll-70 is a highly branched polymer that approximates a sphere with a Stokes radius of ~5.1 nm. This is of a comparable volumetric scale to a T7 polymerase, and much smaller than an intact ribosome (Figure 4.2D). Three experiments were run for each crowding fraction at 0, 5, 10, 15, and 20%. Each experiment was prepared by creating a “master mix” of cell-free reagents and Ficoll-70 at a specific crowding fraction, then aliquoting the master mix into sixteen 15µL reaction volumes. This resulted in 48 reaction traces for each crowding fraction from 0-20%. Additional experiments were conducted for crowding fractions of 25% (16 reaction traces) and 30% (32 reaction traces). Neither the presence of DFHBI-1T nor the Ficoll 70 significantly affected fluorescence measurements (Appendix Figure 4.10).

To visualize the spatial distributions of mRNA and protein, microscopy was performed on cell-free reactions in polydisperse vesicles ranging from approximately 5µm to 30µm in diameter (Figure 4.2E-F). PURExpress reactants with the pRSET-b-mCherry-Spinach2 plasmid and DFHBI-1T were encapsulated in vesicles, following methods described in Nishimura et al, 2014[51]. The vesicle microreactors containing cell-free reactants were allowed to express for 5 hours at room temperature (Appendix,

4.5.3), and imaged using a confocal microscope (Zeiss LSM T-PMT). Vesicles were imaged in uncrowded conditions (0%) and mildly crowded conditions (5%) in two different experiments (Figure 4.2E-F).

Individual Spinach2-DFHBI-1T (mRNA) and mCherry (protein) traces for each uncrowded reaction appeared similar to expression growth traces reported previously[34,44,57], where protein and mRNA populations increased in expression over time before reaching a plateau upon resource limitation of the cell-free reaction (Figure 4.2C)[44,47]. While mRNA-fluorophore fluorescence began to build very quickly after the reaction was initiated, significant protein fluorescence was not seen until ~25 minutes into the experiment (Figure 4.3A), which is consistent with an expected maturation time for the mCherry reporter[158]. Normalized traces of mCherry expression across crowding fractions showed very similar transient behaviors, indicating that protein maturation time was not affected by increased crowding fraction (Appendix Figure 4.6).

The noise in mRNA and protein populations was extracted from each mRNA and protein trace. Briefly, the traces for each reporter within an experiment were averaged to create a general trend for mRNA and protein populations, respectively (Appendix Figure 4.11). This deterministic transient component was subtracted from each trace, leaving noise signals describing the stochastic fluctuations in mRNA and protein populations. The noise magnitude in each molecular population was quantified using the square of the coefficient of variation (where $CV^2 = \frac{\sigma_{fl}^2}{\langle fl \rangle^2}$, where σ_{fl}^2 and $\langle fl \rangle$ were the variance of the fluorescent signal and the final fluorescent level, respectively). For both molecular species, CV^2 was plotted versus its final fluorescence level, which was used throughout the text as a measure of mRNA and protein populations. The Fano factor was calculated as $FF = \frac{\sigma_{fl}^2}{\langle fl \rangle}$.

One-way ANOVA analysis was performed across all crowding fractions (0-30%) for the mRNA experiments (7 groups) and across crowding fractions between 0 and 25% for protein (6 groups). The 30% crowding fraction was not included in the protein analysis since no protein was made for this condition. The significance threshold was set at $p<.05$.

4.5 Appendix

4.5.1 Expression Burst Analysis

Gene expression noise may be characterized using various measures based on the mean value (μ) of a reporter (e.g. fluorescent protein) of a molecular population (e.g. number of mRNA or protein molecules) and the variance (σ^2) of the distribution of these molecular populations. Common measures include coefficient of variation (CV), the square of the coefficient of variation (CV^2), and the Fano factor (FF). These three measures of the relative magnitude of expression variations are related as follows:

$$\boldsymbol{CV^2} = \frac{\boldsymbol{\sigma^2}}{\boldsymbol{\mu^2}} \quad 4.2$$

$$\boldsymbol{CV} = \sqrt{\boldsymbol{CV^2}} \quad 4.3$$

$$\boldsymbol{FF} = \frac{\boldsymbol{\sigma^2}}{\boldsymbol{\mu}} = \boldsymbol{\mu CV^2} \quad 4.4$$

The Fano factor, which has also been known as noise strength[159] or just defined mathematically[147] as μCV^2 , has been used at least since 2001[159] as a direct measure of burst size. Ozbudak et al. experimentally demonstrated[125] that the Fano factor was linearly related to the translational burst size (the average number of proteins produced from an individual mRNA template), and Hasty and Collins[148] commented that these results showed some evidence of transcriptional bursting. So et al. used Fano factor measurements to demonstrate that in *E. coli*, an increased transcriptional rate is often accomplished by an increased transcriptional burst size[100]. Taniguchi et al. showed in *E. coli* that the Fano factor of the protein noise increased approximately linearly with increasing protein population[104], and later analysis showed that this relationship was due to burst size increasing linearly with the protein population[127]. Other studies have used the Fano factor to characterize the transcriptional burst size in expression controlled by the HIV LTR promoter[43,99,147].

Some studies that did not explicitly use the Fano factor as a measure of expression burst size, still used the same underlying Fano factor principle to interpret their results. For example, Hansen et al. noted[52] that in their cell-free expression experiments that increased macromolecular crowding led to a larger CV^2 without a correspondingly large decrease in their measured μ. However, since $FF = \mu CV^2$, their

experimental results demonstrated a crowding induced increase in the Fano factor in a way entirely consistent with the results reported here.

Although the three commonly used measures of noise are closely related, the Fano factor is the measure that is most strongly associated with burst size (B). Conversely, since

$$\boldsymbol{CV^2 = \frac{FF}{\mu} \propto \frac{B}{\mu}} \quad 4.5$$

and

$$\boldsymbol{\mu \propto B f_B}, \quad 4.6$$

then[127]

$$\boldsymbol{f_B \propto \frac{1}{CV^2}}, \quad 4.7$$

where f_B is the burst frequency. Accordingly, throughout this work we use the Fano factor as a measure of burst size and $1/CV^2$ as a measure of burst frequency.

Looking deeper at the relationship between burst size and Fano factor, we can approximate expression bursting as the product of two random processes: Process A (transcriptional initiation) composed of a Poissonian pulse train of impulse functions of weight = 1 and average value $\bar{A}$; and Process B (expression bursting) that is uncorrelated with process A, has a mean value of $\bar{B}$, and a variance of σ_B^2 (note: we have lumped transcriptional and translational bursting into a single process with an average burst size of $\bar{B}$)

$$\boldsymbol{\phi_A(\tau) = \bar{A}\delta(\tau) + \bar{A}^2} \quad 4.8$$

$$\boldsymbol{\phi_B(\tau) = \sigma_B^2\delta(\tau) + \bar{B}^2} \quad 4.9$$

The autocorrelation function of the expression burst is given by the product of the autocorrelation functions of these two functions, or

$$\boldsymbol{\phi_{AB}(\tau) = \phi_A(\tau) * \phi_B(\tau) = \bar{A}\sigma_B^2\delta(\tau) + \bar{A}\bar{B}^2\delta(\tau)} \quad 4.10$$

where we have neglected all the $\bar{A}^2$ terms because $\bar{A} \ll 1$. From this we get

$$\boldsymbol{\sigma_{AB}^2 = \bar{A}\bar{B}^2 + \bar{A}\sigma_B^2} \quad 4.11$$

and the Fano factor (which would be the Fano factor of the protein abundance) is

$$\boldsymbol{FF_{AB} = FF_{<P>} = \frac{\sigma_{AB}^2}{\overline{A}\,\overline{B}} = \overline{B} + \frac{\sigma_B^2}{\overline{B}}} \quad 4.12$$

or

$$\boldsymbol{FF_{<P>} = (\overline{B} + FF_B)}, \quad 4.13$$

where FF_B is the Fano factor of the expression burst size.

Most previous analyses have assumed that expression bursting represents the counting of a Poissonian pulse train[88,159], in which case

$$\boldsymbol{FF_{<P>} = (\overline{B} + 1)} \quad 4.14$$

is the result. Yet, this often-used equation becomes inaccurate when there is greater variability in the expression burst size. As an example, consider a case where the expression burst size distribution is split between a high (B_H) and a low (B_L) level. This would be a simple approximation for an inhomogeneous spatial distribution of resources where some regions are prone to very high translation rates, while other regions produce protein at relatively low rates. In this case

$$\boldsymbol{\overline{B} = P_H B_H + (1 - P_H) B_L \approx P_H B_H} \quad 4.15$$

$$\boldsymbol{\sigma_B^2 \approx P_H (1 - P_H) B_H^2} \quad 4.16$$

$$\boldsymbol{FF_B \approx \frac{P_H (1 - P_H) B_H^2}{P_H B_H} = (1 - P_H) B_H} \quad 4.17$$

$$\boldsymbol{FF_P \approx \overline{B} + (1 - P_H) B_H = P_H B_H + (1 - P_H) B_H = B_H} \quad 4.18$$

where P_H is the probability of having the B_H-sized translational burst. The approximation in these equations assumes that $P_H B_H \gg B_L$. Since $F_{Bapp} \propto \frac{1}{CV^2}$, where F_{Bapp} is the apparent burst frequency,

$$\boldsymbol{F_{Bapp} \propto \frac{\langle P \rangle}{FF_P} \propto \frac{P_H B_H F_B}{B_H} = P_H F_B} \quad 4.19$$

where F_B is the actual burst frequency. So, for the bursting case analyzed above, the apparent burst size is B_H not $\overline{B}$, and the apparent burst frequency is $P_H F_B$ not F_B. In other words, as long as $P_H B_H \gg B_L$, the train of high bursts completely control the noise behavior.

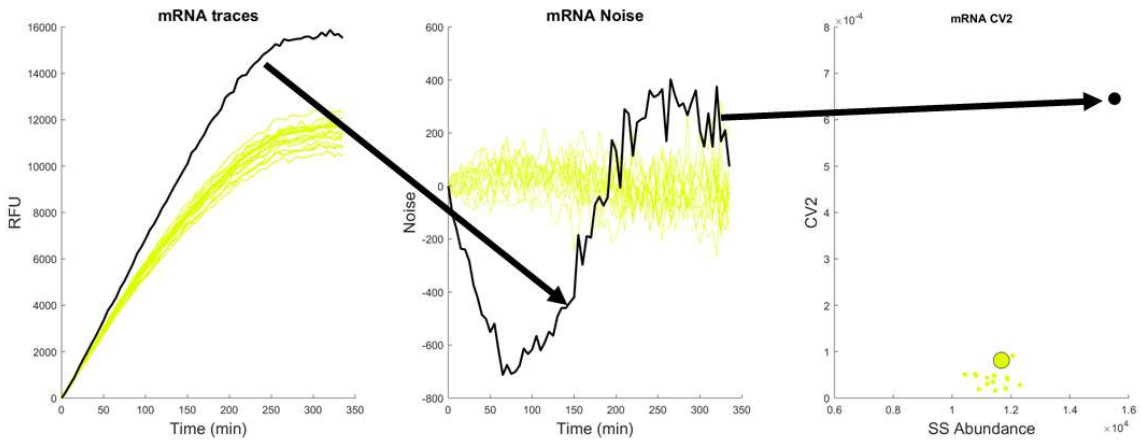


Figure 4.5: "Outlier" behavior in transcription in 5% crowding fraction experiment. A) The individual outlier expression trace is shown here in black. The yellow lines indicate expression traces found in the same experiment. B) Individual noise traces are shown as colored lines. Outlier noise trace shown in black. These outliers typically have high CV^2 values. C) Individual noise values shown as small dots. Mean CV^2 for this experiment shown as a large dot. Outlier noise value shown as black dot.

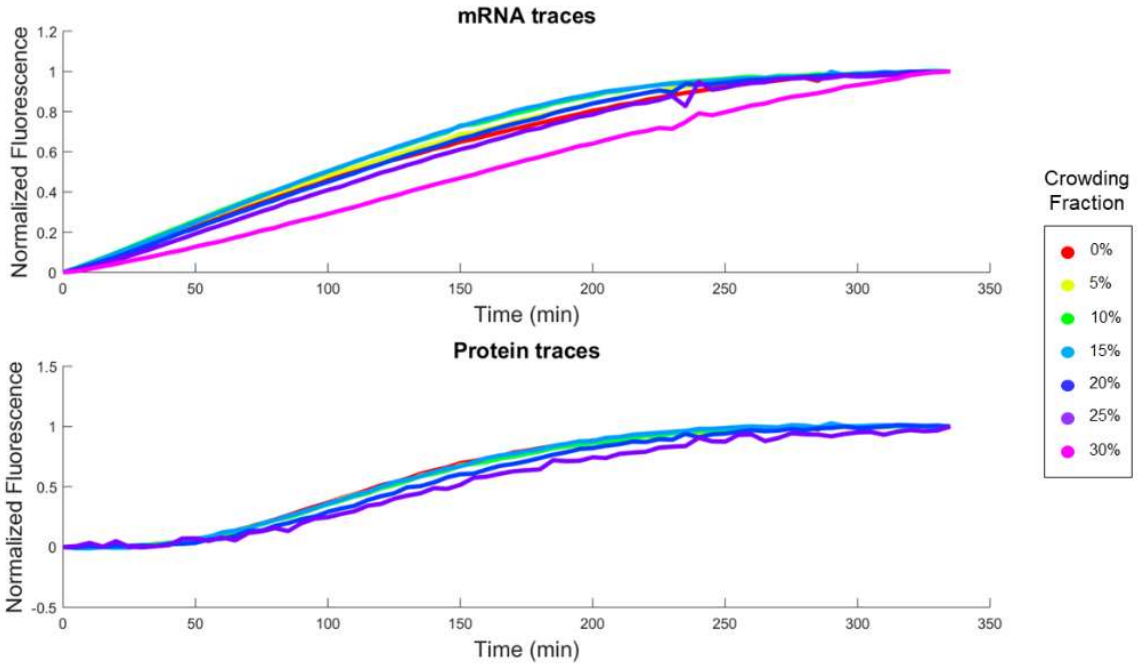


Figure 4.6: Normalized traces of average protein and mRNA expression. Normalized general trends across all crowding fractions up to 25% demonstrate similar shape and behavior over time. Notably, the 30% crowding fraction trace for mRNA appears to deviate from the other trends. No protein expression was detected at the 30% crowding fraction.

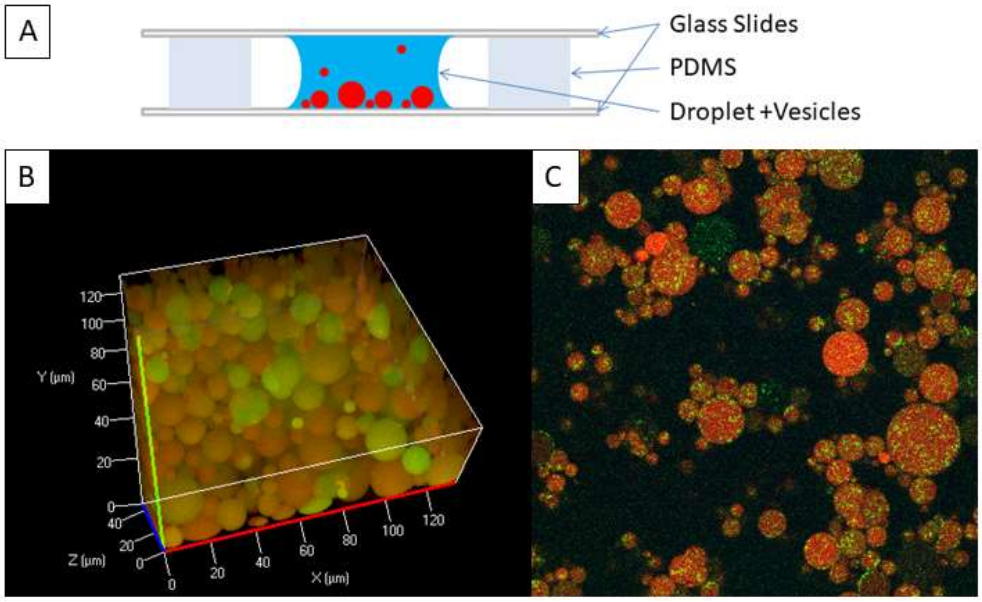


Figure 4.7: Protein synthesis in polydisperse vesicle microreactors. A) An ~10µL droplet containing the vesicles is placed on a glass coverslip, and surrounded by a PDMS spacer. The PDMS spacer is attached to the glass by plasma cleaning both layers, fixing the PDMS to the glass, and heating the layers at 80°C. These steps prevent de-lamination of the PDMS layer once the droplet is applied. Once the droplet has been added to the lower coverslip, another coverslip is placed on top of the PDMS spacer. This assembly allows the vesicles to be viewed over many hours at the interface of the glass, without risk of evaporation. B) Example of reconstructed 3D image of polydisperse vesicles sitting on the glass interface. These vesicles are expressing Green Fluorescent Protein and contain a red fluorescent volume marker. C) Example of 2D image slice of 5% crowded vesicles expressing pRSET-b-mCherry-Spinach2 in presence of DFHBI-1T.

4.5.2 Physical Model of Expression in Crowded Environments

We developed a physically motivated model in which (i) transcription and translation occur in compartmentalized regions of space, (ii) multiple components must assemble for translation to occur, and (iii) crowding influences the size of the compartmentalized regions and the transport of molecules to and from the regions. Increased crowding leads to slower diffusion and, at sufficiently high levels, likely leads to the emergence of isolated regions of transcription and translation that decrease in size as crowding increases.

Consider a system that contains N_c distinct locations at which mRNA is transcribed. Each of these locations is associated with a physical region of space (a "compartment") in which transcription and translation occur. Each compartment has volume V_c and is regarded as well-mixed and independent of the others. As crowding increases, the crowding molecules decrease the volume of each compartment. The total volume of the system is conserved so that $V_{\text{tot}}=V_{\text{nc}}+N_cV_c$, where V_{nc} is the volume outside of the compartments. No reactions occur in this region, but reactants can diffuse from one compartment to another through it. Since the volume of the region outside of the reaction compartments increases with increasing crowding, we use V_{nc} as a proxy to describe the extent of crowding.

Within each compartment, mRNA is produced at rate α and translation is modeled as a series of two bimolecular reactions involving translational components T_1 and T_2:

$$\text{mRNA}+\text{T}_1 \rightarrow C_1$$
$$C_1+\text{T}_2 \rightarrow C_2$$
$$C_2 \rightarrow \text{protein}+\text{mRNA}+\text{T}_1+\text{T}_2$$

The sequence of reactions is motivated by the cooperative binding of multiple chemical species that is needed for translation to occur. We formulate the kinetics in terms of discrete numbers of molecules, so that the binding rates are given by $k_{\text{on},i}=k_{\text{on},i}^{(0)}/V_c$ for $i=1,2$. Here, $k_{\text{on},i}^{(0)}$ is the second-order rate constant that would appear in a well-mixed system with kinetics formulated in terms of concentrations. The expression

for $k_{\mathrm{on},i}$ emphasizes that decreasing V_c leads to higher effective second-order rates within compartments. Hence, small compartment volumes promote binding of mRNA and translational components, as well as rapid rebinding after protein production. The rate of transcription (α) and the rate of protein production (k_p) decay exponentially in time to account for the decay in synthesis observed in cell-free reactions with time constants τ_α and τ_{kp} respectively. Initially, the translational components are populated at random so that the probability of being in a particular region is equal to its fraction of the total volume. Each translational component can transition from the non-compartmental region to a compartment and vice versa. The entry rate into a compartment is Γ and the exit rate is γ. These values are constrained such that the net flux of translational components across compartmental boundaries is zero in the absence of mRNA: $\Gamma/\gamma = V_c/V_{\mathrm{nc}}$. Since crowding impedes diffusion, the rates of entry and exit decrease with increasing V_{nc}. Bulk diffusion rates for a wide variety of molecular species decrease by about a factor of 30 over the range of crowding that we investigated experimentally [151]. To match this magnitude of change in diffusion rates, we assumed the linear relationships $\Gamma = (1 - V_{\mathrm{nc}})\Gamma_0$ and $\gamma = (1 - V_{\mathrm{nc}})\gamma_0$ and allowed V_{nc} to vary between 0.1 and 0.97. Due to their larger size, we assume that the complexes C_1 and C_2 remained confined within a compartment once formed.

To analyze the model, we generated stochastic simulation trajectories using the Gillespie algorithm and tracked the number of mRNA and proteins over time. Noise analysis on the resulting trajectories was performed using the same procedure as for experimental results.

Main text Figure 4.4D shows representative simulation trajectories for two values of V_{nc} that represent low ($V_{\mathrm{nc}} = 0.10$) and high ($V_{\mathrm{nc}} = 0.97$) levels of crowding. Trajectories associated with the higher crowding level are more highly variable and, by inspection, appear to have episodic periods of rapid protein production. Main text Figure 4.4**E** displays results of the noise analysis for simulations ranging from low to high crowding levels. The number of proteins decreases with increasing crowding, with a

corresponding increase in CV^2 at large values of V_{nc}. The Fano factor increases markedly in this regime as well.

Physically, as the crowding increases, fewer translational components are found in reaction compartments. Additionally, their mobility decreases, leading to a slower sampling of the entire system volume. As a consequence, the waiting time to assemble a full translational complex (C_2) increases and the average number of proteins decreases. This is associated with a decrease in burst frequency. However, once all of the translational components are colocalized within a compartment, they are likely to rapidly rebind when the compartment volume is small. This has the effect of increasing burst size at large values of V_{nc}, which is reflected by the results in main text Figure 4.4F.

To study explicitly the effect of translational burst size variability on protein noise, we constructed a simplified two-state expression model (Figure 4.8) that captured the variability of the translational burst size predicted by the model described above. In this model, the average translational burst size remained constant, but the variance of the burst size increased with increasing crowding. In agreement with the measurements, the burst frequency (~k_{on} in the model) was reduced as the crowding increased. Although the distribution of the translational burst size was not measured explicitly, the spatial organization of the mRNA population suggested a distribution with a high degree of skew and kurtosis **(**Figure 4.4B). Accordingly, we chose a simple bimodal distribution with one high and one low translational burst size, where increased crowding led to greater separation between the low and high states without affecting the average translational burst size (Figure 4.8B). Gillespie simulations of this model were performed where the variance of the burst size varied between 1x and 25x the mean burst size, and demonstrate that large variability in the translational burst size – not just changes in the mean burst size – can generate the crowding-induced changes in protein noise behavior measured here (Figure 4.8D).

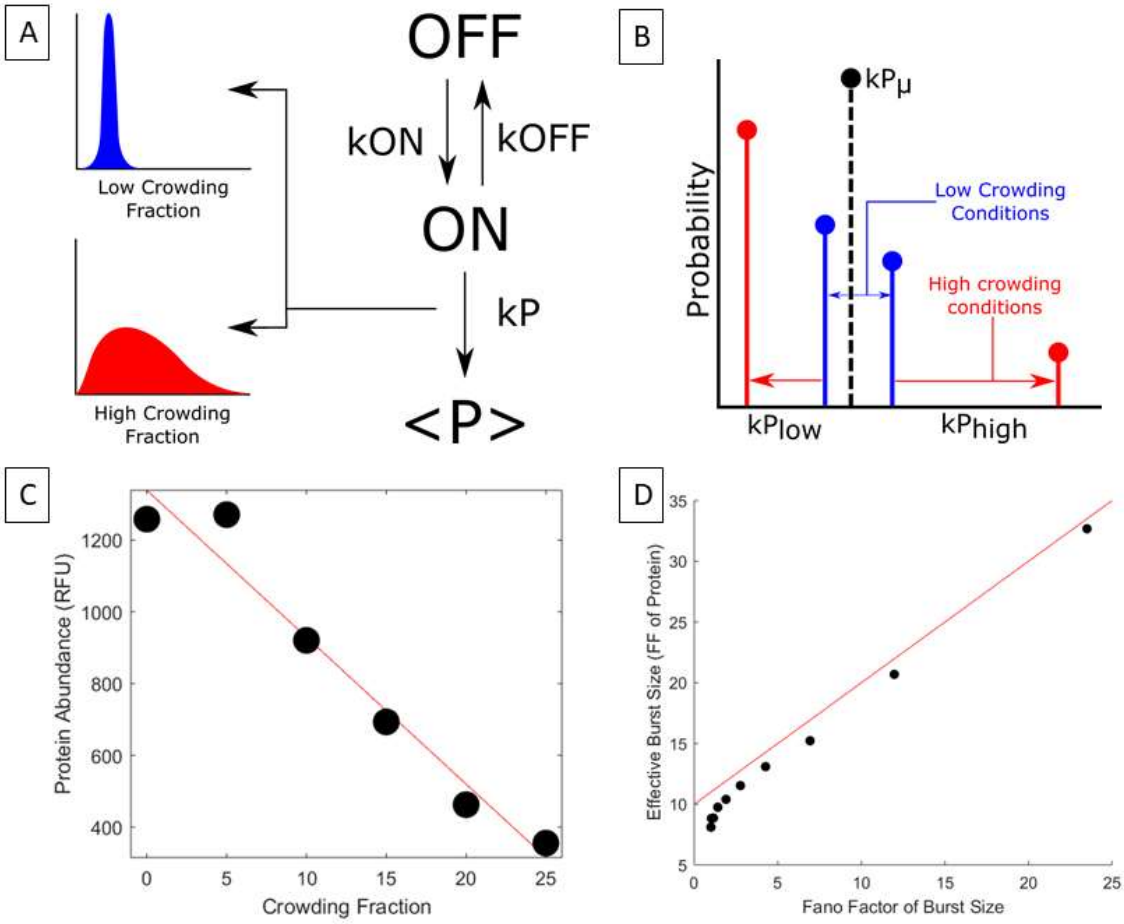


Figure 4.8: Simplified two-state expression model. A) A simplified two-state model of expression which allows modelling of a crowding-controlled distribution of translational burst sizes. B) Simple bimodal distribution of translation rate constant (k_p in (A)) used in the simulations. At low crowding levels k_{phigh} and k_{plow} were nearly equal. At higher crowding levels k_{phigh} and k_{plow} were further apart and the probability of k_{phigh} was reduced. (C) The variation of k_{ON} with crowding fraction for the simulations was found from a linear fit (y= -40.96x+1339; red line) to the protein abundance vs crowding fraction measurements. (D) Gillespie simulation of the two-state model in (A) showing the relationship between the measured burst size (Fano factor of the protein) and Fano factor of the burst size. These simulations used the translation rate distributions in (B) and assumed an average burst size of 10. The red line uses the relationship in main text Eq. (1) with $\overline{\boldsymbol{B}} = \mathbf{10}$.

Table 4.1: List of Model Parameters Used

Model Parameter	Value
N_c	10
Total number, T1	10
Total number, T2	10
α	0.01 s^{-1}
$k_{on,1}^{(0)}$	0.0001 s^{-1}
$k_{on,2}^{(0)}$	0.03 s^{-1}
k_p	20.0 s^{-1}
γ^0	0.7 s^{-1}
I^0	γ^0 V_c/V_{nc}
τ_α	55.55 min
τ_{kp}	27.78 min

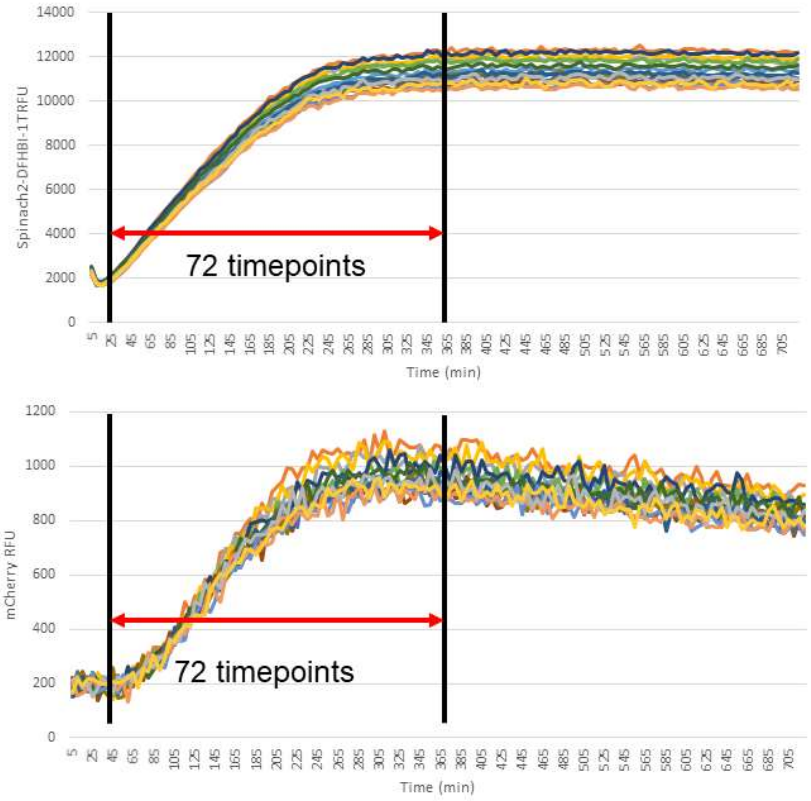


Figure 4.9: Protein and mRNA expression in PURExpress protein synthesis kit. A timescale of Spinach2-DFHBI-1T and mCherry fluorescence was read at 5 minute timesteps over 720 minutes. Each of the 16 time traces represents a single 15μL reaction; each individual color in both graphs corresponds to mRNA or protein expression from one reaction. All time traces used in this work were truncated at 72 time points (denoted by the red arrows), after all traces in all experiments reached maximum expression in both mRNA and protein. Steady state fluorescence shown afterward does not correspond to equilibrium between expression and decay, but rather is governed by the elimination of resources and buildup of toxic side-reactions and by-products[47,49]. The first 5 time points were excluded as the Spinach2-DFHBI-1T readings likely correspond to initial autofluorescence, and not mRNA expression.

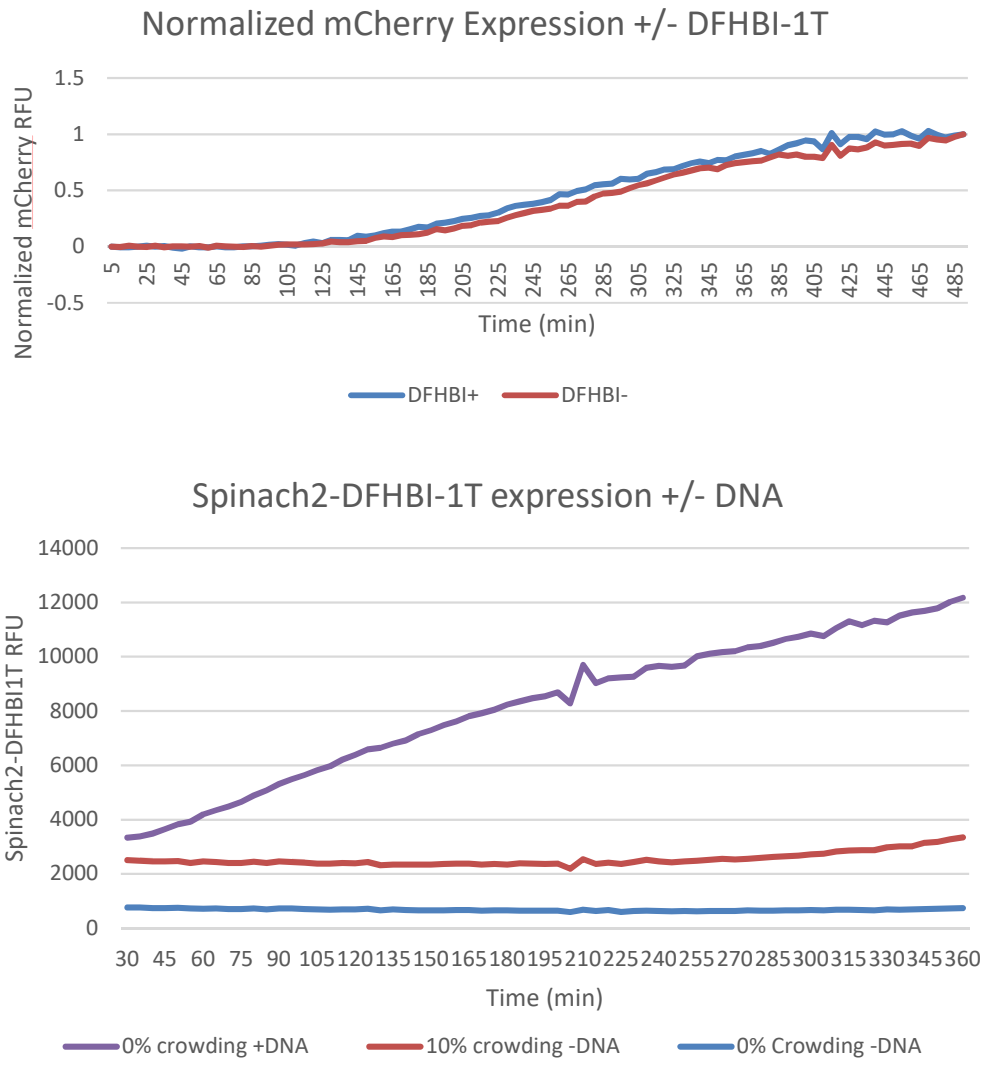


Figure 4.10: Control Experiments for Cell-free reactions. Top) Normalized expression of mCherry in the Promega S30 T7 High-Yield Protein Expression System. Traces are averages of triplicate reactions. The presence of DFHBI-1T did not significantly affect the shape of the average protein traces, indicating that the presence of the fluorophore does not have a significant effect on the translational timescale. Bottom) Spinach2-DFHBI-1T fluorescence in cell-free reactions in presence and absence of DNA. All reactions initially have some autofluorescence which appears to photobleach to a uniform background level after the first few time points. The presence of Ficoll-70 increases this baseline level of fluorescence, but cell-free experiments in the absence of plasmid do not increase significantly in fluorescence over time. This indicates that autofluorescence from Ficoll-70 does not substantially affect the zeroed data.

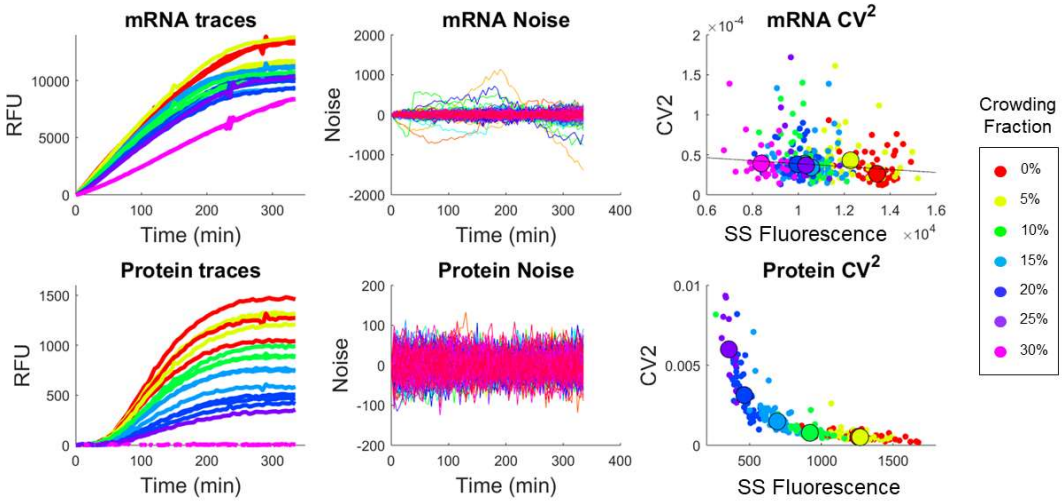


Figure 4.11: mRNA and Protein Expression in PURE System. A single experiment consisted of 16 traces from 15μL microplate well reactions. Fluorescent reads were performed at 495nm and 630nm. The method for performing noise analysis has been detailed in previous work[29,34,43,45,49]. The first panel shows the average timescale trace for each 16-reaction experiment, colored by crowding fraction. The average trace is corrected for the transient and subtracted from each individual reaction trace to create noise traces (shown as individually colored traces in the second panel). These individual traces are plotted as individual points in "noise space", or CV^2 vs steady-state fluorescence, colored by crowding fraction. Large dots represent the average for each crowding fraction group.

4.5.3 Experimental Methods Supplement

Gene Structure

The pRSETb-mCherry-Spinach2 plasmid was constructed using an in-house pRSET-b backbone with an mCherry insertion. The restriction enzymes EcoRI and HindIII were used to insert the Spinach2 oligomer sequence, which was constructed by IDT DNA. The publicly available sequence for this aptamer and framing tRNA scaffolds was taken from the Jaffery Lab website[134]. The aptamer sequence was inserted downstream of the protein sequence to ensure that a full transcript of the mCherry sequence was produced.

Gene Sequence (From T7 Promoter to T7 terminator)

TAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACA
TATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCTGT
ACGACGATGACGATAAGGATCCCGCCACCATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTT
CATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCG
CCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCT

GTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCT
TCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCT
CCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCA
GAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAA
GCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGT
GCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAA
CAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAAGAATTCGAGCTCGAG
ATCTGCAGCTGGTACCATGGCCCGGATAGCTCAGTCGGTAGAGCAGCGGCCGGATGTAACTGAATGAAATGGTGAA
GGACGGGTCCAGTAGGCTGCTTCGGCAGCCTACTTGTTGAGTAGAGTGTGAGCTCCGTAACTAGTTACATCCGGCC
GCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCAAAGCTTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

Gene Preparation

A plasmid midiprep kit (Quantum Prep Plasmid Midiprep Kit, BioRad) was used to isolate pRSETb-mCherry-Spinach2 from Top 10 *E.coli* following manufacturer's instructions. The plasmid was purified by Isopropanol precipitation and the pellet washed with 70% Ethanol. The plasmid was resuspended in nuclease free water at a concentration of 1000 ng/μL.

CFPS Experiment Formulation

Cell-free experiments used the PURExpress cell-free protein expression kit (NEB) diluted with nuclease-free water to the maximum manufacturer-suggested reaction volume of 30μL per reaction. A final plasmid concentration of 8.33 ng/μL was used for all reactions (250ng plasmid/reaction). DFHBI-1T (Lucerna, Inc) was diluted in DMSO and used in the reaction at a final concentration of ~13 μM. From these ratios, experiments were scaled up to a total volume of 300μL. For each experiment, a 300μL master mix was created and divided into 15μL microplate well reactions.

The assembled reactions were applied to a 384-well microplate (Corning #3540, black, clear-bottom) in 15μL aliquots. A 12-hour kinetic read was performed in a microplate reader (Perkin-Elmer EnSpire Multimode Plate Reader), with fluorometric reads at 495nm and 630nm every 5 minutes. Reactions were incubated at 30°C with 2 minutes shaking. The microplate was covered with a qPCR film to prevent evaporation.

Cell-Free Protein Synthesis (CFPS) previous work and control experiments

Similar gene structures to the one constructed for this paper have been examined in cell-free conditions[57,160]. Of particular relevance was Van Nies et al. 2013 which used a

yellow fluorescent protein and an earlier version of the Spinach aptamer and DFHBI. The final concentration of DFHBI-1T used in this work (13μM) approximates the DFHBI concentration used in the previous work (20μM).

The CFPS experiments were initially tested for the effects of Spinach2-DFHBI-1T activity in the Promega S30 T7 High-Yield Protein Expression System. In order to determine how protein expression timescales were affected by the presence or absence of DFHBI-1T, reactions were run with and without the fluorophore, finding that difference in normalized timescales for both conditions are not significant (Figure 4.10, top). In the PURExpress system, reactions were performed in the presence and absence of DNA and crowders in order to determine background fluorescence levels in the Spinach2-DFHBI-1T range (Figure 4.10, bottom).

Vesicle Preparation Methods

Vesicle preparation was adapted from Nishimura et al. 2012[51]; the experiment was modified to observe mRNA and protein expression simultaneously. Briefly, vesicles are prepared by preparing the PURE System as described previously ("Inner solution"), with the addition of sucrose in order to aid visualization of vesicles in brightfield images. The inner solution is placed into a paraffin oil mixture containing phospholipids (POPC, Avanti Polar Lipids) and then vortexed to create a disperse population of vesicles. This "oil phase" vesicle emulsion is layered onto an aqueous "Outer Solution" mixture balanced with the aqueous Inner solution. The layered solutions are then centrifuged for 20 minutes at high speed (~14k g) at 4C. Vesicles are collected by pipetting from the bottom layer. The majority of vesicle diameters range from approximately 5-30 μm. Vesicles were imaged in a method similar to that described in Caveney et al., 2016[34]. Vesicles in the outer solution mixture were pipetted onto a glass coverslip. The droplet containing the vesicles was surrounded by a ~2mm PDMS spacer, and another coverslip was applied on top of the spacer to create an airtight chamber. This setup prevents evaporation and global drift in the imaged vesicle solution. ImageJ was used to perform the analytical steps of the intensity image taken from fluorescence values of the mRNA and Protein. For the Figure 4.4 analytical images of mRNA in vesicles, a square 1444 pixel region of interest was selected in the interior of two representative vesicles, one in an uncrowded (0%) reaction and the other in a mildly crowded 5% reaction. The intensity

Table 4.2: Vesicle Reactants

Inner Solution	Outer Solution
10 µL PURE Solution A, 7.5 µL PURE Solution B, 250 ng pRSETb-mCherry-Spinach2 plasmid 0.5 µL DFHBI-1T (1.56 mM) 5 µL Sucrose (1 M), and filled to 30 µL with Nuclease-free water.	3.6 mL Amino Acid mix (50 mM), 4.9 mL ATP (460 mM), 3.0 mL GTP (500 mM), 1.5 mL CTP (500 mM), 1.5 mL UTP (500 mM), 3.6 mL Spermidine (250 mM), 7.5 mL Creatine Phosphate (1 M), 9 mL DTT (100 mM), 1.5 mL Folinic Acid (4 mg/mL), 168 mL Potassium Glutamate (1 M), 22.6 mL Magnesium Acetate (0.5 M), 60 mL HEPES (1 M), 120 mL Glucose (1 M), and filled to 600 mL with nuclease-free water.

values of the ROIs were extracted and analyzed in MATLAB. 3D representations of the vesicles may be constructed from z-stack imaging, simplifying the estimation of vesicle diameter. The vesicles rest on the glass coverslip; they are largely spherical. Larger vesicles (d>20 microns) tend to be easier to image because they settle on the glass quickly and do not move significantly for several hours.

5 SYNERGISTIC INTERACTIONS BETWEEN CONFINEMENT AND MACROMOLECULAR CROWDING SPATIALLY ORDER TRANSCRIPTION AND TRANSLATION IN CELL-FREE EXPRESSION

A version of this chapter was originally published by S. Elizabeth Norred, Rosemary M. Dabbs, Gaurav Chauhan, Patrick M. Caveney, C. Patrick Collier, Steven M. Abel, and Michael L Simpson:

S. Elizabeth Norred, Rosemary M. Dabbs, Gaurav Chauhan, Patrick M. Caveney, C. Patrick Collier, Steven M. Abel, and Michael L. Simpson "Synergistic Interactions Between Confinement and Macromolecular Crowding Spatially Order Transcription and Translation in Cell-Free Expression" Pre-print available at: *bioRxiv*: 445544

This chapter has been adapted from its published format to accommodate new Figure, Table, and Equation enumeration. The Supplementary Information associated with this work may be found in the Appendix; all references are located at the end of the document. SEN, PMC, CPC, and MLS conceived and planned the experiments. Experiments were performed by SEN and RMD. Lauren K. Collier performed preliminary imaging experiments that led to the three-dimensional rendering seen in Figure 5.4B. SEN designed and constructed the pRSETb-mCherry-Spinach2 plasmid with the assistance of Drs. Jennifer Morrell-Falvey and Amber Bible. The Spinach2 sequence was replicated, with thanks, from the Jaffery Lab website, accessed in 2014[134,135]. SEN, RMD, and PMC performed image analysis. SEN wrote MATLAB scripts for image segmentation, data extraction, and analysis. GC and SMA developed the three-dimensional simulations. All authors participated in data analysis. SEN, GC, SMA, and MLS wrote the manuscript.

5.1 Abstract

Synergistic interactions between macromolecular crowding and confinement spatially organize transcription and translation in cells. Yet, reproducing such spatial ordering in cell-free expression platforms has proven to be elusive. Here we report crowding- and confinement-driven spatial self-organization of cell-free expression that mimics expression behavior within and around the nucleoid of prokaryotes. These experiments use Ficoll-70 to approximate cellular macromolecular crowding conditions within cell-size lipid vesicles. Intriguingly, there was an abrupt change in transcriptional dynamics when crowding reached physiologically relevant levels. Imaging experiments revealed that this change in transcriptional dynamics was coincident with localization of

plasmid DNA and mRNA at the vesicle wall. Computer simulations demonstrated that crowding leads to an entropically induced attraction between plasmid DNA and the wall, causing localization of DNA near the wall at sufficiently high crowding levels. The experiments demonstrate cell-like spatial organization of translation, where translational activity is controlled by chromosomally-templated positioning of mRNA. This cell-free system provides a flexible experimental platform to probe the underlying mechanisms of self-organization of membrane-less structures in cells and the spatial control of gene expression.

5.2 Introduction

Cellular volumes are confined in a range from roughly one femtoliter[18] to several picoliters[17], and much of this volume (e.g. approximately 30% in *E. coli*) is occupied by proteins and other macromolecules[24,136,137]. The physical consequences of macromolecular crowding and cell-relevant confinement has dramatic effects on complex molecular processes, especially ones with diverse molecular components and reaction requirements like gene expression. Cell-free gene expression studies have provided a detailed (expression levels, noise, burst parameters) picture of how confinement alone [34,63] or crowding alone[52,65,67] affect gene expression bursting (Figure 5.1). Unfortunately, little is known about how synergistic interactions between confinement and crowding[69] affect expression. It is an intriguing possibility that crowding and confinement together may have surprising effects on the complex and multi-component diffusion, binding, and re-initiation events of gene expression.

Numerous studies in confined and crowded cellular environment demonstrate how the gene expression process self-organizes into spatial subregions[142,161,162]. Superresolution microscopy in *E.coli* shows that transcriptional and translational components localize preferentially in different microenvironments[142], and that transcripts often remain localized near their origin[162]. In eukaryotes, regulated phase transitions can drive spatial organization of super-enhancers that control transcriptional behavior [143,163]. This self-organized, membrane-less structure in cells creates heterogeneous environments of crowding and confinement that control the sharing of gene expression resources and tune the patterns of expression bursting[142,152,162,164,165]. Cell-free systems can mimic some physical features of cells[34,60,62,67,161], and crowding studies lacking cell-

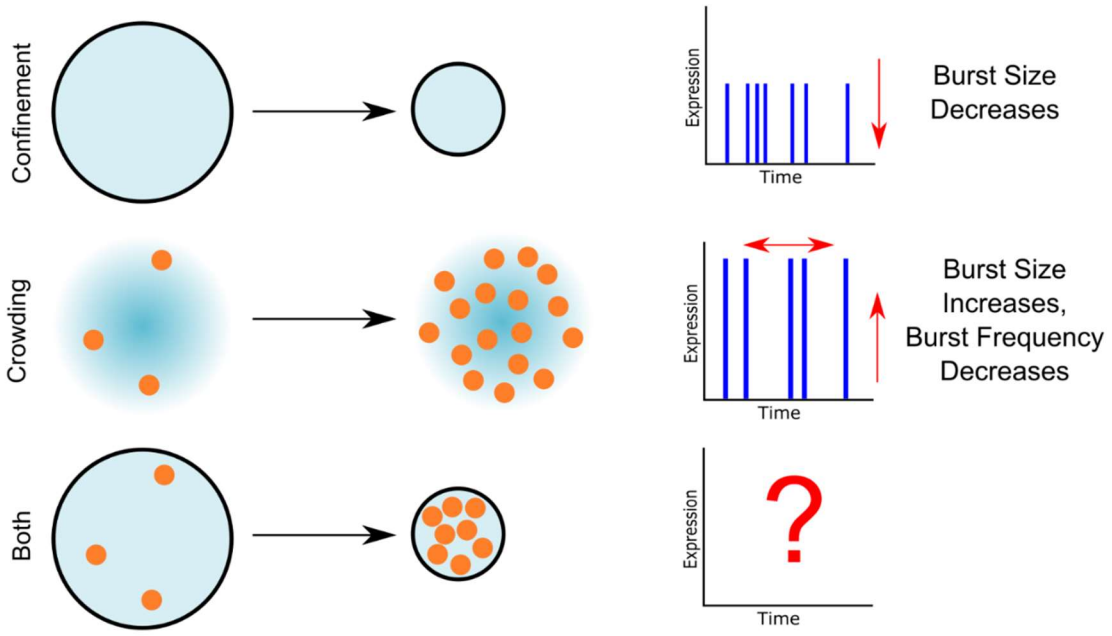


Figure 5.1: Confinement and Crowding affect gene expression bursting parameters. Gene expression bursting is sensitive to confinement (Top) and crowding (middle), but little is known about synergistic effects between crowding and cell-like confinement (bottom).

relevant confinement show some spatial organization of transcription[52,67]. Yet, more fully mimicking cell-like self-organization has been elusive.

Here we report synergistic interactions between macromolecular crowding and confinement of cell-free expression in vesicles that mimics aspects of spatial self-organization observed in prokaryotic cells. Ficoll-70 was used to approximate cellular macromolecular crowders, and crowding levels were varied from 0 to 90 mg/mL. Intriguingly, there was an abrupt change in transcriptional dynamics as crowding reached physiologically relevant levels (>40 mg/mL). Imaging experiments showed that localization of plasmid DNA and mRNA near the vesicle wall generated the change in transcriptional behavior. Computer simulations demonstrated that crowding leads to an entropically induced attraction between plasmid DNA and the wall, causing localization of DNA near the wall at sufficiently high crowding levels. At these higher crowding levels, the mRNA remained localized in the dense DNA region at the vesicle walls and was largely inaccessible for translation. These results demonstrate the spatial organization of

transcription and translation in a cell-free platform that mimics the behavior within and around the nucleoid of prokaryotes, where translational activity is controlled by chromosomally-templated positioning of mRNA[162]. This work demonstrates a flexible experimental platform to understand the underlying mechanisms of self-organization of membrane-less structures in cells and the spatial control of gene expression.

5.3 Results

To understand how the combination of crowding and confinement affects gene expression, we performed cell-free protein synthesis (CFPS) reactions in vesicles crowded with Ficoll-70. Transcription and translation were tracked simultaneously using a coupled mRNA/protein reporter technique described in previous work[57,67,166,167]. Briefly, Spinach2[134], an RNA aptamer which fluoresces in the green range upon hybridization with the fluorophore DFHBI-1T, was inserted downstream of a gene coding for a red fluorescent protein, mCherry[158] (Figure 5.2A). The Spinach2 fluorescence intensity was indicative of the mRNA population and transcriptional dynamics, while the mCherry fluorescence intensity was indicative of the protein population and total (transcriptional and translational) expression dynamics. Ficoll-70 at concentrations from 0-90 mg/mL was added to the Cell-free Protein Synthesis (CFPS) reactions. The concentrations of Ficoll used here mimics lower levels of physiological macromolecular crowding, which can range from 50 to 400 mg/mL[21]. Polydisperse vesicles containing the CFPS reactions were fabricated using a shearing method adapted from Nishimura *et al.* 2012[51] (Figure 5.2C). Vesicles between 14-16 μm in diameter were observed using confocal microscopy over 6 hours (Figure 5.2D). Spinach2 and mCherry fluorescence were measured for individual vesicles over time (Figure 5.2B, 2E). Each experiment was performed in duplicate on separate days for Ficoll-70 concentrations of 0, 10, 40, 60 and 90 mg/mL. Between 93 and 191 vesicles were analyzed per crowding condition, for a total of 694 vesicles. Transcriptional and total expression transients were extracted from individual vesicles using custom MATLAB code for image processing. The expression noise was extracted from these transients using a protocol described in previous work[34,44,67,161] (Figure 5.2E; Methods).

In contrast to either confinement[34] or crowding[67] alone, the shape and timing of the transcriptional transient response varied significantly as crowding was increased in a

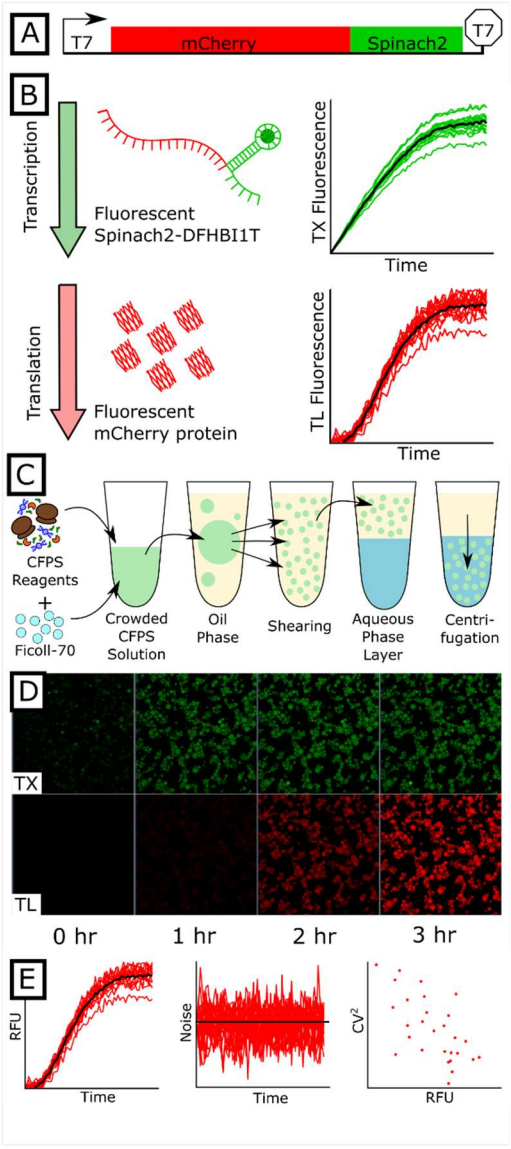


Figure 5.2: Observation of transcription and translation in cell-free reactions. A) The plasmid used for these experiments included a T7 promoter, a gene coding for mCherry, and a sequence encoding an untranslated RNA aptamer, Spinach2. B) Transcription was tracked over time by measuring the fluorescence from Spinach2-DFHBI-1T, the fluorescent hybrid of the Spinach2 aptamer and DFHBI-1T. Total expression was tracked over time by measuring the fluorescence from mCherry. C) Fabrication steps for forming vesicle microreactors. Cell-free reagents and Ficoll-70 were placed in an oil phase solution containing phospholipids, sheared into polydisperse vesicles by vortexing, layered onto a balanced aqueous phase solution, and centrifuged into the solution. D) Confocal images over time of both mRNA and protein expression in a field of polydisperse vesicles. E) Protein and mRNA expression (fluorescence of mCherry and Spinach2 in Relative Fluorescence Units (RFU)) and noise were tracked over time in individual vesicles.

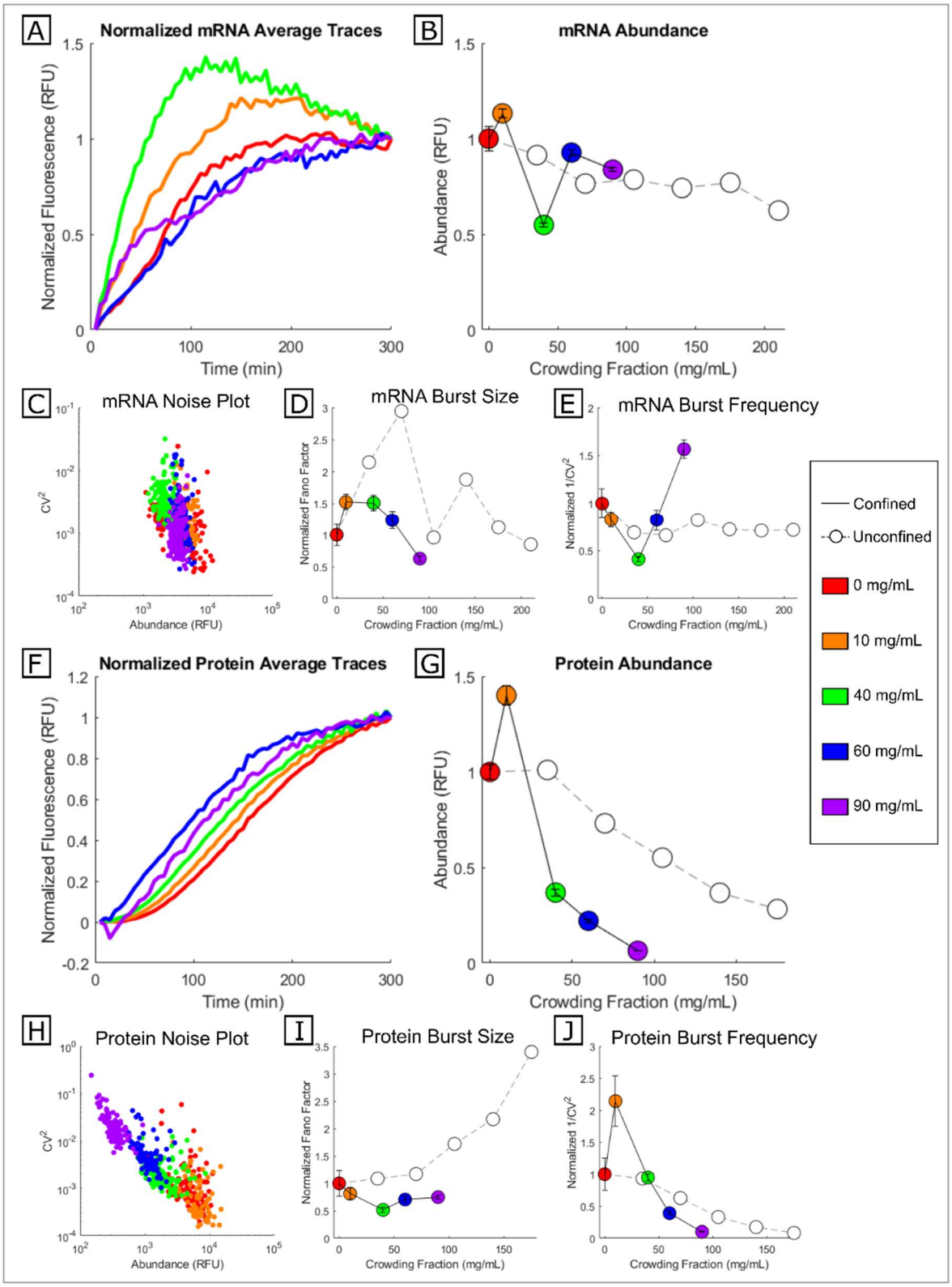
confined environment (Figure 5.3A). With confined crowding, transcription started without delay and persisted over a 100-200 minute (crowding level dependent) duration, at which point the Spinach2-DFHBI1T fluorescence reached its peak value (Figure 5.3A). After the cessation of transcription, the Spinach2-DFHBI1T fluorescence decayed due to photobleaching (Figure 5.3A). At the lower crowding levels (0-40 mg/mL), increased crowding decreased the transcriptional transient risetime to its peak value (Figure 5.3A), with the 40 mg/mL trace reaching its peak value ~125 minutes sooner than the 0 mg/mL trace. Surprisingly, increasing the crowding level beyond 40 mg/mL reversed this trend, with the 60 and 90 mg/mL traces having risetimes similar to the 0 mg/mL transient.

A one-way ANOVA showed that confined crowding resulted in statistically significant differences in mRNA concentrations across the different crowding conditions ($F(689,4)=61.47$, $p<0.001$). In contrast to the unconfined condition[67], the mean mRNA population was quite sensitive to crowding with confinement (Figure 5.3B). Even relatively high levels of unconfined crowding (175 mg/mL) only reduced the mRNA population by about 20%, while a low level of confined crowding (40 mg/mL) reduced the mRNA population by nearly 2-fold (Figure 5.3B). Surprisingly, the mRNA population did not decrease monotonically with increasing crowding fraction. Instead, a crowding level of 40 mg/mL produced the lowest mRNA population (Figure 5.3B) even though this condition produced the quickest risetime.

The protein transients exhibited a delayed start in fluorescence – indicative of the maturation time of mCherry[67,158]– followed by a smooth ~250 minute rise to a peak value. In contrast to unconfined crowding[67], mCherry maturation was significantly altered by confined crowding. The highest levels of crowding decreased maturation time by ~40 minutes (Figure 5.3F) but did not otherwise significantly alter the shape of the mCherry transient (See Appendix, Figure 5.5).

A one-way ANOVA showed that confined crowding produced statistically significant differences in protein concentration across crowding levels ($F(689,4)=526.86$, $p<0.001$). Increased crowding reduced protein synthesis in both unconfined and confined conditions, but cell-like confinement significantly amplified the effects of crowding (Figure 5.3G). Compared to no crowding, confined crowding of 90 mg/mL reduced protein production by more than an order of magnitude. In contrast, a similar

Figure 5.3: Transcription and translation in confined reaction chambers. A) Normalized average mRNA expression transient (Relative Fluorescence Units (RFU) versus crowding fraction). B) Normalized peak mRNA abundance in confined (colored dots) and unconfined (white dots) reactions. Error bars indicate standard error of the mean. C) mRNA expression noise vs. peak mRNA abundance. D) Normalized transcriptional burst size (i.e. mRNA Fano Factor) for confined and unconfined reactions. E) Normalized transcriptional burst frequency ($1/CV^2$) for confined and unconfined reactions. F) Normalized average protein expression transient versus crowding fraction over time. G) Normalized peak abundance of protein in confined (colored dots) and unconfined (white dots) reactions. H) Protein expression noise (CV^2) vs. peak protein abundance. I) Normalized total expression burst size (protein Fano Factor) for confined and unconfined reactions. J) Normalized total expression burst frequency ($1/CV^2$) for confined and unconfined reactions. All unconfined batch reactions performed in a microplate reader (data taken from Norred *et al.* 2018)[67].


A
Normalized mRNA Average Traces
Normalized Fluorescence (RFU)
Time (min)
B
mRNA Abundance
Abundance (RFU)
Crowding Fraction (mg/mL)
C
mRNA Noise Plot
CV²
Abundance (RFU)
D
mRNA Burst Size
Normalized Fano Factor
Crowding Fraction (mg/mL)
E
mRNA Burst Frequency
Normalized 1/CV²
Crowding Fraction (mg/mL)
Confined
Unconfined
0 mg/mL
10 mg/mL
40 mg/mL
60 mg/mL
90 mg/mL
F
Normalized Protein Average Traces
Normalized Fluorescence (RFU)
Time (min)
G
Protein Abundance
Abundance (RFU)
Crowding Fraction (mg/mL)
H
Protein Noise Plot
CV²
Abundance (RFU)
I
Protein Burst Size
Normalized Fano Factor
Crowding Fraction (mg/mL)
J
Protein Burst Frequency
Normalized 1/CV²
Crowding Fraction (mg/mL)

decrease in protein population required an unconfined crowding level exceeding 175 mg/mL (Figure 5.3G). Consistent with other reports[65], there was a statistically significant 1.4 fold increase in protein synthesis with a low level (10mg/mL) of confined crowding (Figure 5.3G).

The details of expression behavior were investigated by examining the noise behavior of the two reporters using the relationships:

$$\boldsymbol{\mu = B f_B} \quad 5.1$$

$$\boldsymbol{CV^2 = \frac{1}{f_B} = \frac{B}{\mu}}, \quad 5.2$$

where μ is the mean population of the reporter (i.e. mCherry or Spinach2-DFHBI1T); CV^2 is the square of the coefficient of variation (variance of reporter population/ μ^2); and B and f_B are parameters that describe the expression pattern. In the 2-state model of expression bursting from an individual gene, B is the burst size (average number of molecules created per burst) and f_B is the burst frequency (number of bursts per unit time)[43,67,99,100,104,125,127,147,148]. With multiple copies of plasmids in each vesicle, the burst frequency may be thought of as the number of statistically independent expression centers, and the burst size as the intensity of expression within the centers[161]. There is evidence of expression patterns indicative of these distinct expression centers even without crowding[161], but with crowding these centers (at least at the transcriptional level) are visible using optical microscopy[52,67]. The transcriptional and total expression burst sizes ($B = \mu CV^2$ (also known as the Fano factor); Figure 5.3D and 3I) and frequencies ($f_B = \frac{1}{CV^2}$; Figure 5.3E and 3J) were calculated using the measured values of μ and CV^2 for each reporter in every vesicle (Figure 5.3C and H).

There was a drastic shift in transcriptional burst behavior between the lower and higher crowding levels. The change in transcriptional burst size (B_{TX}) across all crowding levels with confinement was statistically significant as determined by one-way ANOVA ($F(689,4)=11.48$, $p<0.001$). B_{TX} increased by ~1.5 fold in response to low levels of confined crowding (10, 40 mg/mL; Figure 5.3D), but decreased at higher ($\geq$ 60 mg/mL) crowding levels (Figure 5.3D). The abundance of transcriptional expression centers (f_{BTX}) had a statistically significant change across all crowding levels as determined by one-way ANOVA ($F(689,4)=50.38$, $p<0.001$). As the crowding level increased from 0 to 40

mg/mL, f_{BTX} decreased by ~5 fold. Yet at higher levels of crowding ($\geq$ 60 mg/mL), f_{BTX} exceeded the value measured with no crowding (Figure 5.3E).

Similar to unconfined crowding, there was little change in the total expression burst size (B_T; protein Fano factor) for the crowding levels measured here (Figure 5.3I). Although this change in the protein Fano factor was marginally statistically significant as determined by one-way ANOVA ($F(689,4)=2.77$, $p=0.0263$), only the two groups that produced the highest and lowest protein Fano factor (0 mg/mL and 40 mg/mL) were significantly different from each other. This is similar to unconfined reactions, where protein Fano factors remained relatively unchanged for crowding <140 mg/mL, although higher levels of unconfined crowding did result in large increases (~4-fold) in the protein Fano factor[67]. In contrast, the abundance of total translational expression centers (f_{BT}) was quite sensitive to the crowding level (Figure 5.3J) and varied by more than 30-fold from its peak value at low crowding (10 mg/mL) to its lowest value at high crowding (90 mg/mL). This change was statistically significant as determined by one-way ANOVA ($F(689,4)=154.38$, $p<0.001$). As found for protein concentration, confinement amplified the crowding induced decrease in f_{BT} (Figure 5.3J).

Intriguingly, the results here show a decoupling between transcriptional and translational expression centers. While all translational expression centers must be initiated by a transcriptional center, many transcriptional centers were not translationally active. For example, the spike in protein population with 10 mg/mL confined crowding (Figure 5.3G) was much larger than the associated mRNA concentration increase (Figure 5.3B) because this low level of crowding increased the number of transcriptional expression centers (as indicated by f_{BTX}) that were translationally active (as indicated by f_{BT}; Figure 5.3J). Yet, as the crowding level was increased, transcriptional expression centers became more elusive for the translational machinery. The abundance of transcriptional expression centers reached a peak with 90 mg/mL of confined crowding, but nearly none of these centers were translationally active (Figure 5.3J).

Previous reports demonstrated that crowding without cell-relevant confinement affected translational activity by creating an inhomogeneous spatial distribution of mRNA[52,67], and with increased crowding, much of the mRNA became inaccessible for translation. An especially intriguing feature of expression with confined crowding was the abrupt change in transcriptional behavior as crowding increased from 40 to 60 mg/mL

(Figure 5.3B and 3E), implying a pronounced shift in the mRNA spatial organization. To examine the evolution of mRNA spatial organization with increased crowding, representative vesicles of the same approximate size were compared visually using confocal microscopy. Figure 5.4A shows fluorescence cross-sections of vesicles of ~15μm diameter at the endpoint of a 300-minute reaction. With no crowding, mRNA was visible in distinct spots of relatively uniform intensity and spatial distribution (Figure 5.4A). Low levels of crowding (10 mg/mL) had little discernable effect on the spatial distribution of the mRNA, but as we previously reported[67] did lead to the emergence of a few hot spots of higher local mRNA concentration (Figure 5.4A). At a crowding level of 40 mg/mL these hot spots preferentially appeared near the walls of the vesicles, and were almost exclusively found at the walls with crowding of 60 mg/mL (Figure 5.4A). This localization of mRNA at the vesicle walls was not seen in larger (~90 μm diameter) cell-free reaction chambers[52], indicating that the synergistic effects of confinement and crowding[69] emerge at cell-relevant confinement volumes.

There was no evidence of mRNA diffusion to the wall after synthesis at other locations. Instead, time-lapse images (Figure 5.6) indicate that with crowded confinement, transcription occurred at the wall, and the mRNA remained localized near the site of transcription throughout the experiment. A 3D reconstruction of a crowded vesicle, created from z-stack confocal imaging, showed that mRNA were synthesized in hot spots around the periphery of the vesicles (Figure 5.4B). Since imaging experiments in prokaryotes show that mRNA often exhibit limited dispersion from their site of transcription[162], we hypothesized that confined crowding led to localization of the plasmid DNA near the vesicle wall. To test this hypothesis, we prepared the vesicles with a DNA dye, Pico488, and examined fluorescence cross-sections of ~15 μm diameter vesicles using a confocal microscope. These measurements showed DNA was spatially organized in the same patterns as the mRNA. Without crowders, there was a sparse distribution of DNA throughout the interior of the vesicle. As the crowding level increased, the DNA appeared in localized hot spots distributed throughout the interior of the vesicle. At a crowding level of 40 mg/mL and higher, DNA localized near the vesicle wall (Figure 5.4A, 4C).

We further explored this phenomenon using Brownian Dynamics computer simulations. We utilized a coarse-grained model of a DNA plasmid in a crowded and

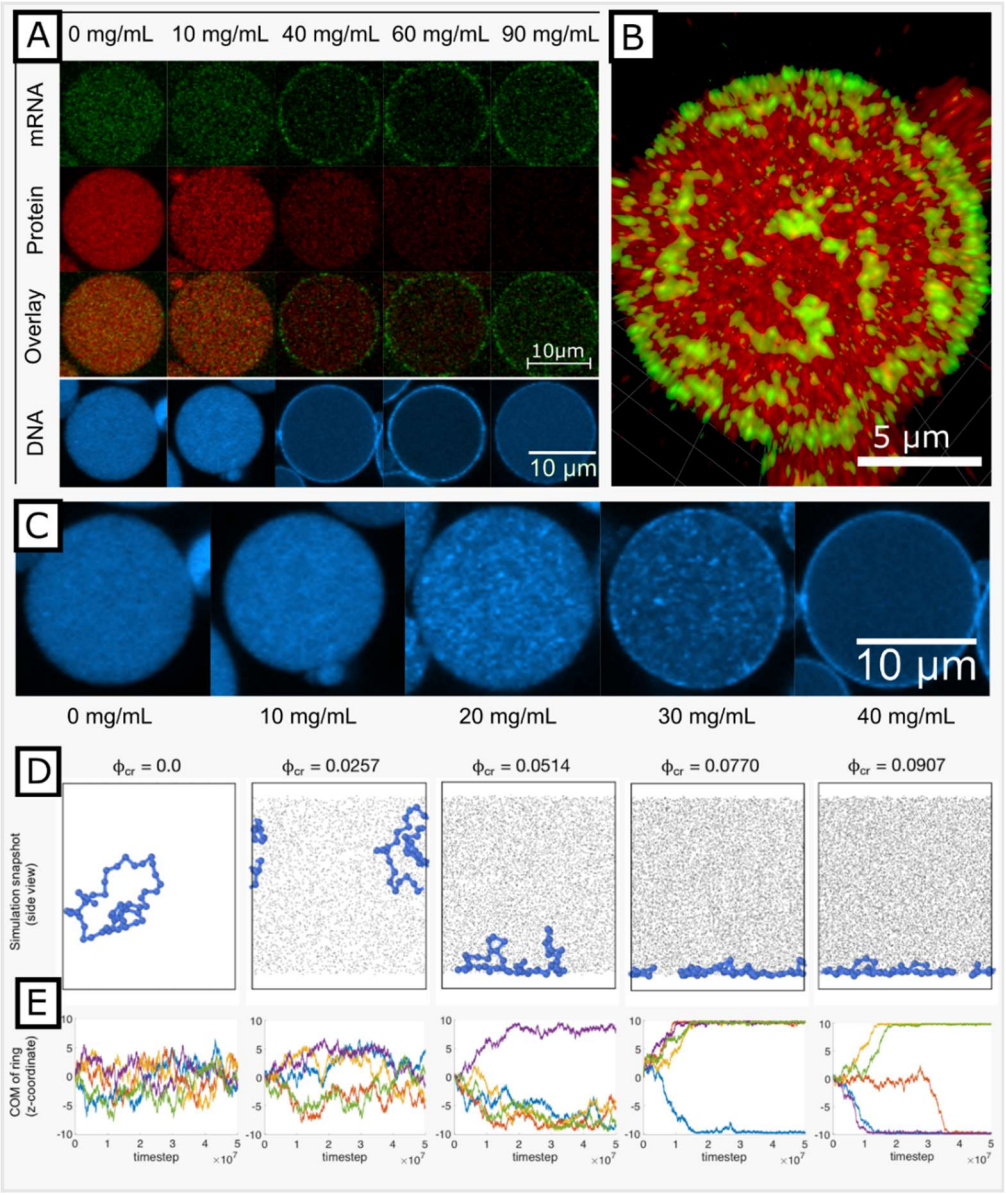


Figure 5.4: Spatial distribution of protein, mRNA, and DNA in vesicles A) Representative vesicles demonstrating spatial distribution of mRNA and protein. Overlay shows combination of mRNA and protein signals in the same vesicle. Bottom row shows representative vesicles from separate experiments where DNA was fluorescently labelled with Pico488 dye (false colored to cyan). B) Three dimensional z-stack reconstruction of an individual vesicle. C) Distribution of DNA demonstrated by fluorescent labelling with Pico488 (false colored to cyan). D) Representative snapshots from equilibrated simulation trajectories at five different crowding fractions (DNA in blue, crowders in grey). The system is viewed from the side and is confined at the top and bottom. E) The z-component of the center of mass of the polymer, with walls confining the system at z = +10 and z = -10 in reduced units. Results from five independent simulations are shown for each crowding fraction. The polymer starts from z = 0 in each simulation.

confined environment: The DNA plasmid was modeled as a flexible ring polymer, the volume fraction of crowders (ϕ_{cr}) was varied by changing the number of crowding particles, and all components interacted via short-ranged repulsive interactions. The system was confined by repulsive walls in one dimension (z) and had periodic boundary conditions in the other dimensions. A full description of the model is provided in the SI. Figure 5.4D shows the behavior of the polymer, initially located in the interior of the system, as a function of the volume fraction of crowders. At low crowding fractions, the ring polymer adopts coiled configurations in the bulk of the system. However, at higher crowding fractions, it localizes to the wall. At ϕ_{cr} = 0.077 and higher, the polymer nearly completely flattens against the wall even though there are no specific attractive interactions between the two. The effective attraction is a consequence of depletion interactions resulting from the presence of crowding particles[168,169]. The crowding-induced localization observed in simulations is consistent with the experimental observations (Figure 5.4C) in which DNA plasmids are found near vesicle walls at high crowding levels.

The cell-free results reported here are strikingly similar to expression behavior in prokaryotic cells where mRNA localization determines translational efficiency[162]. Superresolution microscopy of *E. coli* shows that high-rate transcription preferentially occurs at the periphery of the nucleoid[152], and this mRNA population is efficiently translated as it resides at the boundary with ribosome-rich regions of the cell[142]. Conversely, lower rate transcription occurs throughout the ribosome-poor[142] nucleoid, and the resulting relatively immobile mRNA populations are inefficiently translated[162]. At low levels of confined crowding in cell-free reactions, the mRNA is expressed in distinct translationally-active regions. At higher levels of crowding, the DNA is localized and compacted near the vesicle wall, and the resulting mRNA remains localized in this dense DNA region that appears to be largely inaccessible for translation. By generating a spatial organization of transcription and translation that mimics key aspects of prokaryotic cell membrane-less structure, these cell-free experiments provide a flexible experimental platform to probe the underlying mechanisms of cellular self-organization.

While Ficoll 70 provides a reasonable approximation of cellular crowding, there are important differences to note. An *E. coli* cell is approximately 50-70% water, and the dry weight is ~55% protein, ~20% RNA, ~10% Lipids, and ~15% of other molecules[154].

As much of the protein and RNA is ribosomal[155], ribosomes (~20 nm in diameter in prokaryotes) are a significant contributor to cellular macromolecular crowding. Most of the non-ribosomal protein has radii in the 3-6 nm range with a globular configuration[156]. Ficoll 70 (stokes radius of ~5 nm) and the PURE expression media used here accurately approximate the distribution of globular protein and ribosome crowders in cells. However, extended structures like cytoskeletal filaments or elongated proteins, and structures like the bacterial nucleoid or eukaryotic organelles are not well approximated in these experiments. In contrast to these cell-free experiments, in cells these extended structures may affect expression bursting by allowing for facilitated transport or by creating inhomogeneous crowding[142]. Finally, the results presented here should be applied carefully concerning eukaryotic expression. First, eukaryotic transcriptional burst dynamics are highly sensitive to promoter structure (e.g. TATA boxes) or nucleosome occupancy patterns [124] not present in these cell-free experiments. Furthermore, eukaryotic cells completely decouple transcription and translation, and includes addition steps (e.g. mRNA export from the nucleus) that may affect expression noise[170].

The history of cell-based synthetic biology is one of gene circuit design using specific molecular mechanisms (e.g. promoter/transcription factor interactions) and principles borrowed from electronic circuit design[16]. Much of cell-free synthetic biology has followed a similar path[47,49,171]. However, there is a growing realization that the manipulation of the expression environment – from the composition of the expression reaction media[172] to the physical (confinement and crowding) arrangements – provide another dimension to cell-free synthetic biology. One big advantage of cell-free platforms is that they provide the ability to intricately vary spatial arrangements[153] and are especially well-suited for spatial synthetic biology as a strategy to achieve specific functionality. However, the results here – which show the cell-free spatial organization of expression much like that seen in prokaryotes – suggest that the most immediate application of these experimental systems is to understand the underlying mechanisms of self-organization of collective behavior in cells.

5.4 Methods

In order to simultaneously track transcription and translation outputs, a plasmid vector coding for mCherry and a downstream fluorescent mRNA aptamer, Spinach2,

was expressed[134,166]. The plasmid pRSET-b-mCherry-Spinach2 transcribes from a T7 polymerase promoter to create a transcript with a translated region coding for mCherry, followed by an untranslated aptamer tag which fluoresces after folding and binding with the fluorophore DFHBI-1T[135] ((Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4 H)-one, Lucerna, Inc). A commercial cell-free protein synthesis kit (PURExpress, NEB) was used to express the plasmid in the presence of DFHBI-1T and a crowding molecule, Ficoll-70 (Sigma-Aldrich).

Vesicles containing the cell-free expression system and added components (the "Inner Solution" were prepared by a shearing method adapted from Nishimura *et al.* 2012. In summary, vesicles are prepared by assembling the cell-free reaction mixture, plasmid, DFHBI-1T, sucrose (to aid with visualizing vesicles), and Ficoll-70. Concentrated Ficoll-70 was added at a final concentration from 0-90 mg/mL. The Inner Solution is vortexed in a paraffin oil solution containing phospholipids (POPC, Avanti Polar Lipids) to create a polydisperse population of water-in-oil droplets. This paraffin oil mixture with droplets is layered on to an aqueous "Outer Solution" and then centrifuged for 20 minutes at 4C at ~14k g. The Outer Solution is balanced with the inner solution, containing small molecules found in the PURE system reactions[64,130] (See Appendix for list of reactants). Vesicles are collected by pipetting and are prepared for microscopy by placing ~10 µL of vesicles in Outer Solution between two glass coverslips separated by a ~2 mm PDMS spacer. Most vesicle diameters range from 5-30 µm.

Vesicles were observed while resting on a coverslip, using a (Zeiss LSM 710 Axio Observer) confocal microscope to image every 5 minutes for 6 hours. A 20x objective (Zeiss Plan Apochromat 20x/0.8 M27) was used for the timescale data, followed by a 63x objective (Zeiss Plan Apochromat 63x/1.40 Oil DIC M27) for a more detailed image of fluorescence distribution at the end of the experiment. The Spinach2-DFHBI-1T signal was measured using a 561 nm laser from 488/536 nm Ex/Em. The mCherry was measured using a 561 nm laser from 561/637 nm Ex/Em. Brightfield images were also acquired contemporaneously. For each timepoint, the vesicles were imaged using a z-stack capture, using ~14 images per slice at 2 µm increments. The images were analyzed using ImageJ and custom MATLAB code to detect vesicle size and location and to acquire intensity values for Spinach2-DFHBI-1T and mCherry from individual vesicles over time.

To determine spatial DNA distribution in the vesicles, vesicles were prepared as normal with the 0.25 µL of the 200x fluorophore Pico488 (Lumiprobe) in the Inner Solution, instead of DFHBI-1T. These vesicles were imaged using confocal microscopy, using a 63x objective and 561nm laser at 488/536 Ex/Em. Ficoll-70 was added at a final concentration at 0, 10, 20, 30, 40, 60, and 90 mg/mL. A control reaction containing no DNA was also performed. Z-stack renderings and cross-sections in the middle of vesicles were used to characterize DNA distribution within vesicles.

Two experiments were performed for each Ficoll-70 crowder concentration of 0, 10, 40, 60, and 90 mg/mL. For an individual vesicle in each experiment, a 6-hour mRNA and protein expression trace was extracted for noise analysis, as described in previous work[34,67,127,161]. Built-in functions in ImageJ and custom MATLAB code were used to identify the boundaries of vesicles in a brightfield view, select regions of interest around each vesicle, and extract fluorescence information from each ROI. Fluorescence for individual vesicles was tracked over time for detected vesicles between a diameter of 14-16µm.

Briefly, for each population of zeroed expression traces in a single experiment, an average trace, or “general trend” was calculated for all vesicles, and then was subtracted from each individual vesicle’s expression trace. This was done for both reporters, revealing the “noise signals”, or the stochastic variation in mRNA or protein reporter at each timepoint. The coefficient of variation squared (CV^2) was used to quantify the noise magnitude in the molecular populations of mRNA or protein. The coefficient of variation squared is defined in Eqn. 5.2. The steady state fluorescence level was defined as the maximum fluorescence level attained per fluorescence trace, instead of the endpoint of the trace. This was due to trace decay caused by photobleaching, causing the final fluorescence value not to be descriptive of the total molecular populations produced. Since the mRNA traces reached their final value prior to the protein traces, noise traces were only calculated based on the first 150 minutes of the mRNA reactions. However, protein noise traces, which are derived from traces that generally express continuously over the entire experiment, were calculated over 300 minutes. CV^2 is plotted against these maximum values, which is useful for describing changes in the bursting patterns between experimental conditions.

5.5 Appendix

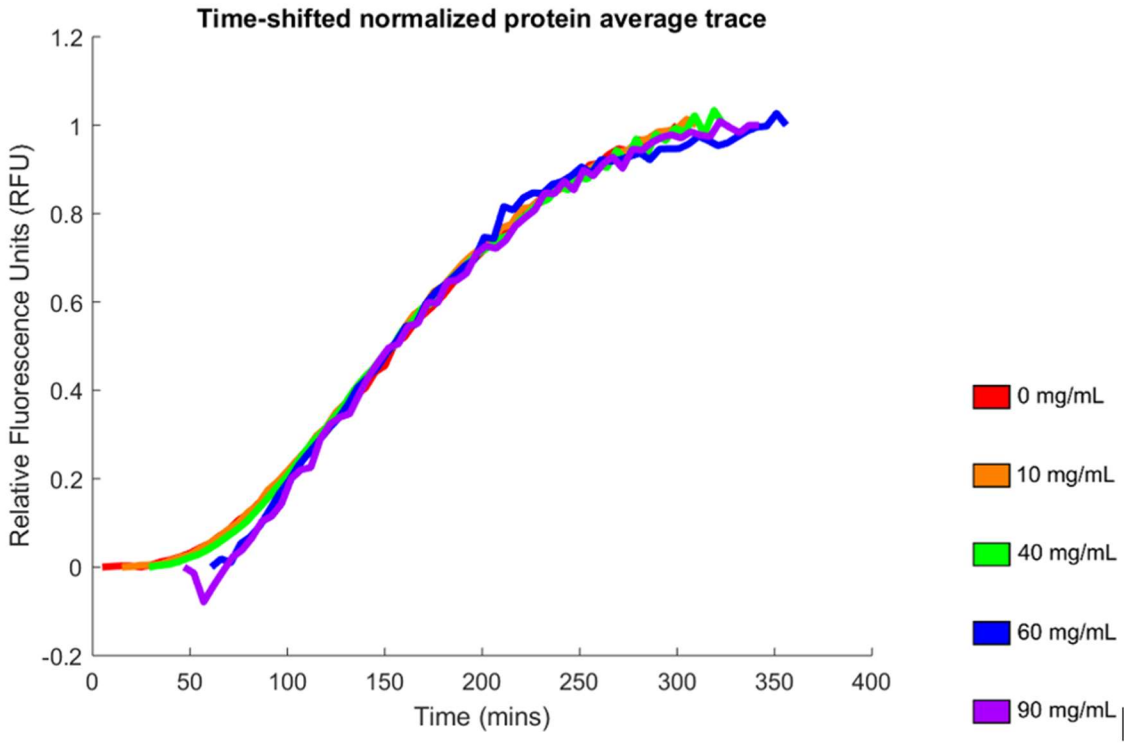


Figure 5.5: Time-shifted normalized average protein expression transients Protein expression transients (Relative Fluorescence Units versus Crowding Fraction) shifted in time to demonstrate similarity in transient profile. The crowding fractions were shifted forward in time by 10 mins for 10 mg/mL, 24 mins for 40 mg/mL, 56 mins for 60 mg/mL, and 42 mins for 90 mg/mL.

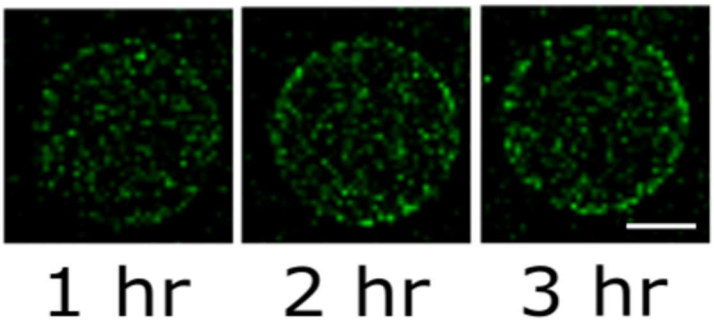


Figure 5.6: Timescale of mRNA expression in a single vesicle Expression of mRNA in a single vesicle at a 40 mg/mL crowding. Scale bar indicates 10 microns. The mRNA remains localized at the periphery of the vesicle for the duration of the experiment.

5.5.1 Description of Computational Model

We employed Brownian Dynamics computer simulations to explore the physical consequences of crowding on a DNA plasmid in a confined system. We utilized a coarse-grained description of the system, with the DNA plasmid modeled as a flexible ring polymer and the level of crowding controlled by changing the number of crowding particles. All components interacted via short-ranged repulsive interactions. The system was confined by repulsive walls in one dimension (z) and had periodic boundary conditions in the other dimensions. The ring polymer consisted of 50 monomer units, with adjacent monomers connected via a finitely extensible nonlinear elastic (FENE) bond[173] potential, given by

$$U_{FENE} = -0.5KR_0{}^2\, ln\left[1 - \left(\frac{r}{R_0}\right)^2\right]$$

where r is the center-to-distance between two adjacent monomers. The maximum distance between two monomers connected via a FENE bond was $R_0 = 2.0\sigma$ with spring constant $K = 15.0\varepsilon/\sigma^2$.

The DNA plasmid used in experiments consisted of 3772 bp. In the coarse-grained model, each monomer unit corresponds to approximately 75 bp, which is approximately 25.5nm in length. All components interacted via the short-ranged and purely repulsive Weeks-Chandler-Andersen (WCA) potential[174],

$$U_{ij} = \begin{cases} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \epsilon_{ij} & r_{ij} < 2^{1/6}\sigma_{ij} \\ 0 & r_{ij} \geq 2^{1/6}\sigma_{ij} \end{cases}$$

where r_{ij} is the center-to center distance between particles i and j. For crowder-crowder, monomer-monomer and monomer-crowder interactions, $\varepsilon_{ij} = \varepsilon$ and $\sigma_{ij} = R_i + R_j$, where R_i and R_j denote the size (radius) of particles i and j respectively. The size of monomer particles was $R_m = 0.75\sigma$ and the size of crowder particles was $R_c = 0.25\sigma$. The simulation box ranged from -10σ to 10σ in the x- and y-directions, with periodic boundary conditions used. The simulation box ranged from -12σ to 12σ in z-direction, with confinement introduced by positioning an immobile array of wall particles at $z = -12\sigma$ and $z = 12\sigma$. The wall particles interacted with crowders and monomers via the WCA potential with $\varepsilon_{wm}=\varepsilon_{wc}=100\varepsilon$ and $\sigma_{wm}=\sigma_{wc}=2.0\sigma$. This gives the size of the box confined in z-direction to be $20\sigma \times 20\sigma \times 20\sigma$.

The number of crowding particles (N_c) was varied, with the volume occupied defined as $V_c = N_c\, 4\pi R_c^3/3$.

We used the LAMMPS[175] simulation package to conduct the Brownian Dynamics simulations. The temperature used was $T=\varepsilon/k_B$. The equations of motion were integrated in time using the velocity-Verlet algorithm with a timestep of 0.01τ, where $\tau=\sigma\sqrt{m/\varepsilon}$. The mass of a monomer and the mass of a crowder particle were 14.02m. The friction coefficient was $140.2\tau^{-1}$. The resulting trajectories were visualized using OVITO[176].

5.5.2 Gene Sequence

(From T7 Promoter to T7 terminator)

TAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACA
TATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCTGT
ACGACGATGACGATAAGGATCCCGCCACCATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTT
CATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCG
CCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCT
GTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCT
TCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCT
CCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCA
GAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAA
GCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGT
GCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAA

CAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAAGAATTCGAGCTCGAG
ATCTGCAGCTGGTACCATGGCCCGGATAGCTCAGTCGGTAGAGCAGCGGCCGGATGTAACTGAATGAAATGGTGAA
GGACGGGTCCAGTAGGCTGCTTCGGCAGCCTACTTGTTGAGTAGAGTGTGAGCTCCGTAACTAGTTACATCCGGCC
GCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCAAAGCTTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

5.5.3 Gene Preparation

A commercial DNA preparation kit (Quantum Prep Plasmid Midiprep Kit, BioRad) was used to extract pRSETb-mCherry-Spinach2 from Top 10 *E. coli.* Plasmid DNA was purified by Isopropanol precipitation and washed with 70% Ethanol. Plasmid DNA was resuspended at a concentration of 1000 ng/μL in nuclease-free water.

5.5.4 Vesicle Inner Solution Formulation

Cell-free experiments were performed using the PURExpress cell-free protein expression kit (NEB) diluted to a volume of 30μL per reaction. A final plasmid concentration of 33 ng/μL was used for all reactions (1000 ng plasmid/reaction). The fluorophore DFHBI-1T (Lucerna, Inc) was diluted in DMSO to concentration of 1.56 mM. A stock solution of Ficoll-70 (Sigma-Aldrich) was prepared in nuclease free water at 50 w/v%. The crowder solution replaced the nuclease-free water component of each reaction in order to alter the crowding fraction of the solution. Therefore, all experiments across all crowding fractions had the same amount of proteins and other cell-free reactants.

Table 5.1: Vesicle Reactants

Inner Solution	Outer Solution
10 µL PURE Solution A, 7.5 µL PURE Solution B, 250 ng pRSETb-mCherry-Spinach2 plasmid 5 µL Sucrose (1 M) 1000 ng plasmid DNA 0.5 µL DFHBI-1T (1.56 mM) (mRNA analysis) OR 0.25 µL Pico488 (200x) (DNA analysis) filled to 30 µL with Nuclease-free water.	3.6 mL Amino Acid mix (50 mM), 4.9 mL ATP (460 mM), 3.0 mL GTP (500 mM), 1.5 mL CTP (500 mM), 1.5 mL UTP (500 mM), 3.6 mL Spermidine (250 mM), 7.5 mL Creatine Phosphate (1 M), 9 mL DTT (100 mM), 1.5 mL Folinic Acid (4 mg/mL), 168 mL Potassium Glutamate (1 M), 22.6 mL Magnesium Acetate (0.5 M), 60 mL HEPES (1 M), 120 mL Glucose (1 M), and filled to 600 mL with nuclease-free water.

6 CONCLUSION

This thesis focused on understanding how cell-like confinement and macromolecular crowding influenced gene expression behavior. Enabled by the bottom-up capabilities of cell-free reaction platforms, both confinement and crowding were examined individually, without the confounding influence of all the complexities (e.g. global gene expression, growth, cell division) seen in cell. Of particular note, a gene construct allowing simultaneous tracking of transcription and translation was developed. Gene expression behavior was extracted from a detailed analysis of gene expression noise, which is intimately linked to expression bursting. It was found that physiologically-relevant levels of confinement or crowding alone had dramatic effects on gene expression bursting patterns. After characterizing each of these physical conditions individually, both confinement and crowding were examined together to develop an understanding of the synergistic effects of crowding and confinement on gene expression bursting. The results presented here provided new insights about the allocation of limited expression resources and spatial orientation of gene expression in cell-free systems, and how both of these factors affect temporal noise in translation.

This work was published in a series of four papers, presented as chapters, that explored expression behavior with confinement, with macromolecular crowding, and finally with both confinement and macromolecular crowding. Each chapter reflected an important component in understanding how spatial properties of the molecular environment affected cell-free reactions. Chapter 2 described a technique for encapsulating cell-free reactants in microfluidic chambers of varying confinement and revealed that this confinement method could be used to analyze gene expression noise traces similarly to methods used to analyze noise in cells. Using the platform described in Chapter 2, Chapter 3 demonstrated the results of varying confinement in cell-free gene expression, revealing that resource sharing and gene expression bursting are closely linked. Chapter 4 examined macromolecular crowding in the absence of confinement and examined both mRNA and protein expression simultaneously. This chapter revealed that increased crowding decoupled mRNA and protein expression statistics and that the spatial distribution of mRNA became more heterogeneous with increased crowding, resulting in large temporal variation (noise) in both mRNA and protein populations. Chapter 5 explored the effects of macromolecular crowding and

confinement together on gene expression noise, revealing more information about the crowding-dependent distribution of mRNA in confined spaces and the effect of this spatial organization on transcriptional and translation behavior.

The first paper in the series[45] (Chapter 2) demonstrated an experimental platform that confined expression reactions into cell-like volumes. A microfluidic reaction encapsulation platform was developed for acquiring expression noise data in cell-free systems. The key feature of the design was the use of arrays of micron-scale reaction chambers of varying sizes with a flexible “control valve” that initiated the capture of the cell-free reactants at a well-defined start time of expression (“time zero”). The static, fabricated design of these reaction chambers was useful for time-lapse microscopy, allowing easy imaging of reaction kinetics involved in protein synthesis. This platform allowed the evaluation of expression noise of genetic circuits in a cell-free environment, analogous to techniques used to evaluate gene expression noise in cells.

The second paper in the series[34] (Chapter 3) used the cell-free platform described above to study how variations in the level of confinement affected expression burst behavior. In particular, this study examined the intimate link between resource sharing and gene expression bursting. Examining the expression noise of protein made in these reaction chambers revealed numerous insights about how resources are shared within a confined volume. Protein statistics from small, confined chambers were summed together and compared to statistics from large, less-confined chambers of equivalent volume—i.e. resources were allocated into small discrete pools and large shared pools of equivalent resource concentration and volume. Though the abundance of protein made in both cases was roughly equivalent, the sums of small discrete pools produced much lower expression noise. This result indicated that expression in the larger, shared pool exhibits larger expression bursts than in smaller, divided pools. This finding suggested that genes in the large pool were more likely to infrequently acquire or consume a large fraction of the total resources instead of more frequently creating new, smaller bursting centers. So in environments with many shared resources, resources are more likely to be allocated to existing expression locations instead of nucleating new bursting centers, suggesting that expression bursts are self-reinforcing.

The third paper in the series[67] (Chapter 4) explored how physiological levels of macromolecular crowding affect expression behavior. While Chapter 3 described the

influence of reaction confinement, the reactions were performed in a comparatively dilute cell-free expression system that did not reflect physiological levels of crowding. In this chapter, an artificial crowder, Ficoll-70, was added to the reactions in order to mimic physiological levels of macromolecular crowding. Ficoll-70 is a high molecular weight, highly branched polysaccharide with an average Stokes radius of ~5 nm[22,151]. Insights in Chapter 3 about transcriptional activity were also limited as the experimental methods tracked only protein fluorescence, requiring information about transcriptional activity to be inferred. Chapter 4 introduced a new gene expression tool which allowed the simultaneous tracking of protein and mRNA expression, allowing transcriptional activity to be measured directly. This method also provided the insight of coupled transcriptional and translational signals, as each fluorescent reporter (a fluorescent mRNA aptamer and a fluorescent protein) was derived from the same DNA promoter. In brief, as the crowding fraction of Ficoll 70 increased, final abundance in mRNA decreased modestly, but the final abundance in protein decreased dramatically. At the highest crowding fraction tested, no translation was detectable for the entire course of the experiment, and the mRNA timescale trace demonstrated a decrease in rate and duration not comparable to the previous, less-crowded conditions. Corresponding to previous studies showing that bacterial transcription elongation is correlated with the rate of translation, these results suggest that the elongation rate in this cell-free context decreases in the absence of translation. Low translational activity also reduced the energy demand of the overall gene expression reaction, allowing an extended duration of transcription reactions. The disparity in the overall abundances of mRNA and protein was also reflected in the noise behavior—as crowding increased, mRNA noise changed very little, whereas protein noise increased substantially with higher crowding. This indicates that the mRNA abundance statistics do not have a strong controlling role in overall expression noise.

The final paper in the series[68] (Chapter 5) looked at how the combination of both crowding and confinement affected expression behavior. This investigation further explored the relationship between crowding and confinement by examining bursting behavior in crowded lipid vesicle reaction chambers. It was shown that bursting parameters for confined crowding were different than those for unconfined crowding, and that the expression bursting behaviors shifted collectively near physiological crowding

and confinement levels. Confocal imaging of these vesicles showed that the spatial organization of nucleic acids shifted dramatically with increased crowding, with both DNA and RNA appearing at the periphery of the reaction chambers as crowding increased. The results from this work reinforced the findings from the previous paper describing a decoupling between mRNA and protein statistics, adding a spatial dimension to the noise analysis that revealed how transcription and translation centers may not share the same levels of activity. At crowding fractions above 40 mg/mL, statistics for transcriptional and overall bursting also shifted dramatically with the spatial reorganization of mRNA and DNA at the vesicle wall, further underlining how spatial noise in transcription can affect temporal noise in translation.

The results here demonstrate that gene expression bursting is controlled by several spatial mechanisms, including the local concentrations of resources and distribution of molecular populations. Cell-relevant levels of crowding and confinement create an inhomogeneous distribution of nucleic acids, controlling the expression burst behavior and temporal noise of the protein population. Above a certain crowding threshold, the confined cell-free positioning of transcriptional activity reflects expression behavior and positioning of transcripts in prokaryotes. Superresolution microscopy studies in *E.coli* demonstrate that the nucleoid and ribosome population are spatially segregated, where transcription occurs at high rates around the peri-nucleoid space and mRNA is translated efficiently at the boundary between the nucleoid and ribosome-populated space[142,152,162]. Since RNA polymerases more easily permeate the nucleoid[152], transcription also occurs within the nucleoid interior, but because of the lack of ribosomes in this region, this mRNA is inefficiently translated. This “crowded nucleoid” model in cells is mimicked in the crowded and confined cell-free expression platform; at higher levels of crowding approaching physiological crowding levels, DNA and the resulting mRNA are localized at the vesicle periphery. The mRNA remains persistently localized in this dense region at the periphery of the vesicle, and it appears to be largely inaccessible to translation. The spatial organization of transcription and translation reactions within this cell-free structure provides a useful method for examining mechanisms of self-organization in gene expression present in cells.

Though the cell-free platform organization mimics the gene expression organization found in cells, there are some important differences between the crowded

cellular environment and the environment reported here. At least half of an *E. coli* cell weight is water, and the dry weight is composed of protein, RNA, lipids, and other molecules[154]. Ribosomes themselves make up a large portion of the cell weight and volume occupancy[155], and themselves contribute significantly to crowding. In the cell-free environment, ribosomes are the largest protein structures, but the cell contains numerous larger structures, like cytoskeletal filaments and other protein assemblies. The cell-free environment used here also does not mimic the dense central bacterial nucleoid, and it is far-removed from eukaryotic models where transcription and translation are segregated by a nucleus and other mechanisms.

Many factors in the spatial self-organization of gene expression components presented here have yet to be explored. Of particular interest is the positioning of ribosomes and other translation machinery in the crowded and confined cell-free context. In the crowded-nucleoid model of prokaryotic gene expression, ribosomes are excluded from the crowded nucleoid center and much of translation does not occur co-transcriptionally[142]. Intact ribosomes are also the largest molecular component in these reactions and themselves may act as active crowding molecules in these reactions. Understanding how ribosomes are positioned in this space may reveal important information about the spatial influence of translation machinery on gene expression as well as deepen an understanding of cooperativity between transcription and translation.

Other interesting factors that could be explored would involve manipulating the vesicle expression platform to investigate different membrane chemistries, rigidity, or other chemical factors that may alter how nucleic acids adsorb to the interior of the vesicle. While simulations from Chapter 5 indicate that nucleic acid localization at the vesicle wall is dominated by entropic effects, the strength of this effect may possibly be modulated by manipulating the charge profile of the membrane, for example. Changing the size or type of artificial crowder to a material like Ficoll-400 or PEG-8000 may also alter expression and bursting patterns in interesting ways. Both of these crowders have been shown to decrease protein abundance in batch reactions at lower w/v% than Ficoll-70[65]. This suggests that these crowders may halt translation at lower crowding fractions than Ficoll-70 in batch reactions, and may be an interesting method for investigating rate dependencies between transcription and translation in both batch and confined reactions.

In general, the crowded vesicle platform provides an interesting molecular arrangement for investigating how gene circuits could express in the crowded peri-nucleoid space and the less-crowded translation-heavy region. Introducing new genes with feedback or regulatory mechanisms could be an immediate pathway for investigating how crowding and confinement affect the bursting profiles of more complicated circuits.

Synthetic biology gene design has traditionally relied on the availability of diverse gene circuit control elements in order to engineer functions and network behaviors. However, even well-characterized gene circuit elements (e.g. transcriptional control using the lac promoter) may be strongly influenced by local molecular conditions surrounding the gene. Manipulation of the expression environment, from placing circuits in different cellular contexts to fine-tuning physical molecular arrangements with cell-free reactions, gives researchers a new pathway for altering gene circuit functionality. Cell-free platforms in particular are useful for spatially arranging gene expression while allowing direct access to the molecular components of a reaction. The results shown here indicate that cell-free systems may be used to spatially arrange gene expression in an orientation reflective of the crowded-nucleoid prokaryote model. Platforms like these are well-suited for understanding the mechanisms controlling self-organization of gene expression in cells, and for developing new spatial controls in the future of gene design.

REFERENCES

1 Administration, U. S. E. I. International Energy Outlook 2016. Report No. DOE/EIA-0484(2016), (Department of Energy, 2016).
2 Vesborg, P. C. & Jaramillo, T. F. Addressing the terawatt challenge: scalability in the supply of chemical elements for renewable energy. *RSC Advances* **2**, 7933-7947 (2012).
3 BP. BP Statistical Review of World Energy 2015. (2015).
4 Administration, U. S. E. I. Monthly Energy Review. Report No. DOE/EIA-0035(2016/4), (2016).
5 Langholtz, M., Stokes, B. & Eaton, L. 2016 Billion-ton report: Advancing domestic resources for a thriving bioeconomy, Volume 1: Economic availability of feedstock. (2016).
6 Fortman, J. *et al.* Biofuel alternatives to ethanol: pumping the microbial well. *Trends in biotechnology* **26**, 375-381 (2008).
7 Savage, D. F., Way, J. & Silver, P. A. Defossiling fuel: how synthetic biology can transform biofuel production. *ACS Chemical Biology* **3**, 13-16 (2008).
8 Purnick, P. E. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nature reviews Molecular cell biology* **10**, 410-422 (2009).
9 Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: thinking outside the cell. *Metabolic engineering* **14**, 261-269 (2012).
10 Kalscheuer, R., Stölting, T. & Steinbüchel, A. Microdiesel: Escherichia coli engineered for fuel production. *Microbiology* **152**, 2529-2536 (2006).
11 Wang, Y., Chen, S. & Yu, O. Metabolic engineering of flavonoids in plants and microorganisms. *Applied Microbiology and Biotechnology* **91**, 949-956 (2011).
12 Rice, M. K. & Ruder, W. C. Creating biological nanomaterials using synthetic biology. *Science and technology of advanced materials* **15**, 014401 (2013).
13 Westfall, P. J. *et al.* Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proceedings of the National Academy of Sciences* **109**, E111-E118, doi:10.1073/pnas.1110740109 (2012).
14 Brophy, J. A. N. & Voigt, C. A. Principles of Genetic Circuit Design. *Nature methods* **11**, 508-520, doi:10.1038/nmeth.2926 (2014).
15 Noireaux, V., Bar-Ziv, R. & Libchaber, A. Principles of cell-free genetic circuit assembly. *Proceedings of the National Academy of Sciences* **100**, 12672-12677 (2003).
16 Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nature Reviews Genetics* **11**, 367-379 (2010).
17 Puck, T. T., Marcus, P. I. & Cieciura, S. J. Clonal growth of mammalian cells in vitro: growth characteristics of colonies from single HeLa cells with and without a" feeder" layer. *Journal of Experimental Medicine* **103**, 273-284 (1956).

18 Kubitschek, H. E. Growth During the Bacterial Cell Cycle: Analysis of Cell Size Distribution. *Biophysical Journal* **9**, 792-809 (1969).
19 Yeates, T. O., Crowley, C. S. & Tanaka, S. Bacterial microcompartment organelles: protein shell structure and evolution. *Annual review of biophysics* **39**, 185 (2010).
20 Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C. & Shively, J. M. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat Rev Microbiol* **6**, 681-691 (2008).
21 Ellis, R. J. & Minton, A. P. Cell biology: join the crowd. *Nature* **425**, 27-28 (2003).
22 Zimmerman, S. B. & Minton, A. P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual review of biophysics and biomolecular structure* **22**, 27-65 (1993).
23 Minton, A. P. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *Journal of biological chemistry* **276**, 10577-10580 (2001).
24 Zhou, H. X., Rivas, G. N. & Minton, A. P. in *Annual Review of Biophysics* Vol. 37 *Annual Review of Biophysics* 375-397 (Annual Reviews, 2008).
25 Tabaka, M., Kalwarczyk, T., Szymanski, J., Hou, S. & Holyst, R. The effect of macromolecular crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics* **2**, doi:10.3389/fphy.2014.00054 (2014).
26 Morelli, M. J., Allen, R. J. & Ten Wolde, P. R. Effects of macromolecular crowding on genetic networks. *Biophysical journal* **101**, 2882-2891 (2011).
27 Süel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J. & Elowitz, M. B. Tunability and Noise Dependence in Differentiation Dynamics. *Science* **315**, 1716-1719, doi:10.1126/science.1137455 (2007).
28 Eldar, A. & Elowitz, M. B. Functional roles for noise in genetic circuits. *Nature* **467**, 167-173 (2010).
29 Weinberger, L. S., Dar, R. D. & Simpson, M. L. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat. Genet.* **40**, 466 (2008).
30 Chong, S., Chen, C., Ge, H. & Xie, X. S. Mechanism of transcriptional bursting in bacteria. *Cell* **158**, 314-326, doi:10.1016/j.cell.2014.05.038 (2014).
31 Carey, L. B., Van Dijk, D., Sloot, P. M. A., Kaandorp, J. A. & Segal, E. Promoter sequence determines the relationship between expression level and noise. *PLoS Biol.* **11**, e1001528 (2013).
32 Sanchez, A., Garcia, H. G., Jones, D., Phillips, R. & Kondev, J. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS Comput. Biol.* **7**, e1001100 (2011).
33 Lee, W. *et al.* A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39**, 1235-1244,

doi:http://www.nature.com/ng/journal/v39/n10/suppinfo/ng2117_S1.html (2007).
34 Caveney, P. M. *et al.* Resource Sharing Controls Gene Expression Bursting. *Acs Synth Biol*, doi:10.1021/acssynbio.6b00189 (2016).
35 Austin, D. W. *et al.* Gene network shaping of inherent noise spectra. *Nature* **439**, 608-611, doi:10.1038/nature04194 (2006).
36 Bar-Even, A. *et al.* Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**, 636 (2006).
37 Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633 (2003).
38 Newman, J. R. *et al.* Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840 (2006).
39 Soltani, M., Vargas-Garcia, C. A., Antunes, D. & Singh, A. Intercellular Variability in Protein Levels from Stochastic Expression and Noisy Cell Cycle Processes. *PLoS Comput Biol* **12**, e1004972, doi:10.1371/journal.pcbi.1004972 (2016).
40 Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6**, 451 (2005).
41 Sanchez, A. & Golding, I. Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188 (2013).
42 Munsky, B., Neuert, G. & van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183 (2012).
43 Dar, R. D. *et al.* Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17454 (2012).
44 Karig, D. K., Jung, S. Y., Srijanto, B., Collier, C. P. & Simpson, M. L. Probing cell-free gene expression noise in femtoliter volumes. *ACS Synth. Biol.* **2**, 497 (2013).
45 Norred, S. E. *et al.* Sealable Femtoliter Chamber Arrays for Cell-free Biology. *J. Visualized Exp.*, e52616, doi:10.3791/52616 (2015).
46 Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V. & Murray, R. M. Gene circuit performance characterization and resource usage in a cell-free “breadboard”. *Acs Synth Biol* **3**, 416-425 (2014).
47 Sun, Z. Z. *et al.* Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *J. Visualized Exp.*, e50762 (2013).
48 Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **4**, 220 (2008).
49 Karig, D. K., Iyer, S., Simpson, M. L. & Doktycz, M. J. Expression optimization and synthetic gene networks in cell-free systems. *Nucleic acids research* **40**, 3763-3774 (2012).

50 Iyer, S., Karig, D. K., Norred, S. E., Simpson, M. L. & Doktycz, M. J. Multi-input regulation and logic with T7 promoters in cells and cell-free systems. *PloS one* **8**, e78442 (2013).
51 Nishimura, K., Tsuru, S., Suzuki, H. & Yomo, T. Stochasticity in gene expression in a cell-sized compartment. *Acs Synth Biol* **4**, 566-576 (2014).
52 Hansen, M. M. *et al.* Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets. *Nature nanotechnology* **11**, 191-197 (2016).
53 Okano, T., Matsuura, T., Kazuta, Y., Suzuki, H. & Yomo, T. Cell-free protein synthesis from a single copy of DNA in a glass microchamber. *Lab Chip* **12**, 2704-2711, doi:10.1039/c2lc40098g (2012).
54 Siuti, P., Retterer, S. T. & Doktycz, M. J. Continuous protein production in nanoporous, picolitre volume containers. *Lab Chip* **11**, 3523 (2011).
55 Nomura, S. M. *et al.* Gene expression within cell-sized lipid vesicles. *ChemBioChem* **4**, 1172 (2003).
56 Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17669-17674, doi:10.1073/pnas.0408236101 (2004).
57 van Nies, P. *et al.* Unbiased Tracking of the Progression of mRNA and Protein Synthesis in Bulk and in Liposome-Confined Reactions. *ChemBioChem* **14**, 1963-1966 (2013).
58 Elani, Y., Law, R. V. & Ces, O. Protein synthesis in artificial cells: using compartmentalisation for spatial organisation in vesicle bioreactors. *Physical Chemistry Chemical Physics* **17**, 15534-15537, doi:10.1039/C4CP05933F (2015).
59 Saito, H. *et al.* Time-resolved tracking of a minimum gene expression system reconstituted in giant liposomes. *ChemBioChem* **10**, 1640-1643 (2009).
60 Sokolova, E. *et al.* Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11692-11697, doi:10.1073/pnas.1222321110 (2013).
61 Hansen, M. M. *et al.* Cell-Like Nanostructured Environments Alter Diffusion and Reaction Kinetics in Cell-Free Gene Expression. *Chembiochem* **17**, 228-232, doi:10.1002/cbic.201500560 (2016).
62 Thiele, J. *et al.* DNA-functionalized hydrogels for confined membrane-free in vitro transcription/translation. *Lab on a Chip* **14**, 2651-2656, doi:10.1039/C3LC51427G (2014).
63 Kato, A., Yanagisawa, M., Sato, Y. T., Fujiwara, K. & Yoshikawa, K. Cell-sized confinement in microspheres accelerates the reaction of gene expression. *Sci. Rep.* **2**, 283 (2012).

64 Shimizu, Y., Kanamori, T. & Ueda, T. Protein synthesis by pure translation systems. *Methods* **36**, 299-304, doi:http://dx.doi.org/10.1016/j.ymeth.2005.04.006 (2005).
65 Ge, X., Luo, D. & Xu, J. Cell-free protein expression under macromolecular crowding conditions. *PLoS One* **6**, e28707, doi:10.1371/journal.pone.0028707 (2011).
66 Tan, C., Saurabh, S., Bruchez, M., Schwartz, R. & LeDuc, P. Molecular crowding shapes gene expression in synthetic cellular nanosystems. *Nature nanotechnology* **8**, 602-608, doi:10.1038/nnano.2013.132 (2013).
67 Norred, S. E. *et al.* Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns. *Acs Synth Biol* **7**, 1251-1258 (2018).
68 Norred, S. E. *et al.* Synergistic interactions between confinement and macromolecular crowding spatially order transcription and translation in cell-free expression. *bioRxiv*, 445544 (2018).
69 Fowlkes, J. D. & Collier, C. P. Single-molecule mobility in confined and crowded femtolitre chambers. *Lab Chip* **13**, 877 (2013).
70 Klammt, C. *et al.* High level cell-free expression and specific labeling of integral membrane proteins. *European Journal of Biochemistry* **271**, 568-580 (2004).
71 Wong, R. W. & Blobel, G. Cohesin subunit SMC1 associates with mitotic microtubules at the spindle pole. *Proceedings of the National Academy of Sciences* **105**, 15441-15445 (2008).
72 Niederholtmeyer, H., Stepanova, V. & Maerkl, S. J. Implementation of cell-free biological networks at steady state. *Proceedings of the National Academy of Sciences* **110**, 15985-15990 (2013).
73 Shin, J., Jardine, P. & Noireaux, V. Genome replication, synthesis, and assembly of the bacteriophage T7 in a single cell-free reaction. *Acs Synth Biol* **1**, 408-413 (2012).
74 Algire, M. A. *et al.* Development and characterization of a reconstituted yeast translation initiation system. *RNA* **8**, 382-397 (2002).
75 Iizuka, N., Najita, L., Franzusoff, A. & Sarnow, P. Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from Saccharomyces cerevisiae. *Molecular and Cellular Biology* **14**, 7322-7330 (1994).
76 Shin, J. & Noireaux, V. An E. coli Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *Acs Synth Biol* **1**, 29-41, doi:10.1021/sb200016s (2012).
77 Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. H. Programmable on-chip DNA compartments as artificial cells. *Science* **345**, 829-832, doi:10.1126/science.1255550 (2014).
78 Zawada, J. F. *et al.* Microscale to Manufacturing Scale-up of Cell-Free Cytokine Production—A New Approach for Shortening Protein Production

Development Timelines. *Biotechnology and Bioengineering* **108**, 1570-1578, doi:10.1002/bit.23103 (2011).
79 Simpson, M. L., Sayler, G. S., Fleming, J. T. & Applegate, B. Whole-cell biocomputing. *Trends in Biotechnology* **19**, 317-323, doi:10.1016/S0167-7799(01)01691-2 (2001).
80 Korobkova, E., Emonet, T., Vilar, J. M. G., Shimizu, T. S. & Cluzel, P. From molecular noise to behavioural variability in a single bacterium. *Nature* **428**, 574, doi:10.1038/nature02404 https://www.nature.com/articles/nature02404#supplementary-information (2004).
81 Weinberger, L. S., Burnett, J. C., Toettcher, J. E., Arkin, A. P. & Schaffer, D. V. Stochastic Gene Expression in a Lentiviral Positive-Feedback Loop: HIV-1 Tat Fluctuations Drive Phenotypic Diversity. *Cell* **122**, 169-182, doi:10.1016/j.cell.2005.06.006 (2005).
82 Arkin, A., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics* **149**, 1633-1648 (1998).
83 Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene Expression in a Single Cell. *Science* **297**, 1183-1186, doi:10.1126/science.1070919 (2002).
84 McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences* **94**, 814-819 (1997).
85 Rao, C. V., Wolf, D. M. & Arkin, A. P. Control, exploitation and tolerance of intracellular noise. *Nature* **420**, 231-237 (2002).
86 Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-226 (2008).
87 Pedraza, J. M. & van Oudenaarden, A. Noise Propagation in Gene Networks. *Science* **307**, 1965-1969, doi:10.1126/science.1109090 (2005).
88 Simpson, M. L., Cox, C. D. & Sayler, G. S. Frequency domain analysis of noise in autoregulated gene circuits. *Proceedings of the National Academy of Sciences* **100**, 4551-4556 (2003).
89 Nourian, Z. & Danelon, C. Linking genotype and phenotype in protein synthesizing liposomes with external supply of resources. *Acs Synth Biol* **2**, 186-193 (2013).
90 Pereira de Souza, T., Stano, P. & Luisi, P. L. The Minimal Size of Liposome-Based Model Cells Brings about a Remarkably Enhanced Entrapment and Protein Synthesis. *ChemBioChem* **10**, 1056-1063, doi:doi:10.1002/cbic.200800810 (2009).
91 Park, N., Um, S. H., Funabashi, H., Xu, J. & Luo, D. A cell-free protein-producing gel. *Nature Materials* **8**, 432, doi:10.1038/nmat2419 https://www.nature.com/articles/nmat2419#supplementary-information (2009).
92 van Swaay, D. & deMello, A. Microfluidic methods for forming liposomes. *Lab on a Chip* **13**, 752-767, doi:10.1039/C2LC41121K (2013).
93 Ito, T. & Okazaki, S. Pushing the limits of lithography. *Nature* **406**, 1027, doi:10.1038/35023233 (2000).

94 Herold, K. & Rasooly, A. Lab-on-a-Chip Technology: Biomolecular Separation and Analysis, vol. 2. *The Publisher* (2009).
95 Unger, M. A., Chou, H.-P., Thorsen, T., Scherer, A. & Quake, S. R. Monolithic Microfabricated Valves and Pumps by Multilayer Soft Lithography. *Science* **288**, 113-116, doi:10.1126/science.288.5463.113 (2000).
96 Cox, C. D., McCollum, J. M., Allen, M. S., Dar, R. D. & Simpson, M. L. Using noise to probe and characterize gene circuits. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10809-10814, doi:10.1073/pnas.0804829105 (2008).
97 Jewett, M. C. & Swartz, J. R. Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnology and bioengineering* **87**, 465-471 (2004).
98 Simpson, M. L. *et al.* Noise in biological circuits. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **1**, 214-225, doi:doi:10.1002/wnan.22 (2009).
99 Dar, R. D. *et al.* Transcriptional bursting explains the noise–versus–mean relationship in mRNA and protein levels. *PloS one* **11**, e0158298 (2016).
100 So, L. *et al.* General properties of transcriptional time series in Escherichia coli. *Nat. Genet.* **43**, 554 (2011).
101 Retterer, S. T., Siuti, P., Choi, C. K., Thomas, D. K. & Doktycz, M. J. Development and fabrication of nanoporous silicon-based bioreactors within a microfluidic chip. *Lab Chip* **10**, 1174 (2010).
102 Hensel, Z. *et al.* Stochastic expression dynamics of a transcription factor revealed by single-molecule noise analysis. *Nat. Struct. Mol. Biol.* **19**, 797 (2012).
103 Levine, J. H., Lin, Y. & Elowitz, M. B. Functional roles of pulsing in genetic circuits. *Science* **342**, 1193 (2013).
104 Taniguchi, Y. *et al.* Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533 (2010).
105 Skupsky, R., Burnett, J. C., Foley, J. E., Schaffer, D. V. & Arkin, A. P. HIV promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLoS Comput. Biol.* **6**, e1000952 (2010).
106 Ceroni, F., Algar, R., Stan, G. B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415 (2015).
107 Gyorgy, A. & Del Vecchio, D. *Limitations and trade-offs in gene expression due to competition for shared cellular resources*. (2014).
108 Guantes, R., Díaz-Colunga, J. & Iborra, F. J. Mitochondria and the non-genetic origins of cell-to-cell variability: More is different. *BioEssays* **38**, 64 (2016).
109 De Vos, D., Bruggeman, F. J., Westerhoff, H. V. & Bakker, B. M. How molecular competition influences fluxes in gene expression networks. *PLoS One* **6**, e28494 (2011).

110 Mather, W. H., Hasty, J., Tsimring, L. S. & Williams, R. J. Translational cross talk in gene networks. *Biophys. J.* **104**, 2564 (2013).
111 Kepler, T. B. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81**, 3116 (2001).
112 Simpson, M. L., Cox, C. D. & Sayler, G. S. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.* **229**, 383 (2004).
113 To, T. L. & Maheshri, N. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science* **327**, 1142 (2010).
114 Blake, W. J. *et al.* Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell* **24**, 853 (2006).
115 Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
116 Suter, D. M. *et al.* Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472 (2011).
117 Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811 (2004).
118 Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025 (2005).
119 Zenklusen, D., Larson, D. R. & Singer, R. H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **15**, 1263 (2008).
120 Chubb, J. R., Trcek, T., Shenoy, S. M. & Singer, R. H. Transcriptional pulsing of a developmental gene. *Curr. Biol.* **16**, 1018 (2006).
121 Yu, J., Xiao, J., Ren, X., Lao, K. & Xie, X. S. Probing gene expression in live cells, one protein molecule at a time. *Science* **311**, 1600 (2006).
122 Siuti, P., Retterer, S. T., Choi, C. K. & Doktycz, M. J. Enzyme reactions in nanoporous, picoliter volume containers. *Anal. Chem.* **84**, 1092 (2012).
123 Jewett, M. C. & Swartz, J. R. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **86**, 19 (2004).
124 Dar, R. D., Karig, D. K., Cooke, J. F., Cox, C. D. & Simpson, M. L. Distribution and regulation of stochasticity and plasticity in Saccharomyces cerevisiae. *Chaos* **20**, 037106 (2010).
125 Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nat. Genet.* **31**, 69 (2002).
126 Dey, S. S., Foley, J. E., Limsirichai, P., Schaffer, D. V. & Arkin, A. P. Orthogonal control of expression mean and variance by epigenetic features at different genomic loci. *Mol. Syst. Biol.* **11**, 806 (2015).

127 Dar, R. D., Razooky, B. S., Weinberger, L. S., Cox, C. D. & Simpson, M. L. The Low Noise Limit in Gene Expression. *PLoS One* **10**, e0140969 (2015).
128 Peccoud, J. & Ycart, B. Markovian modeling of gene-product synthesis. *Theor. Popul. Biol.* **48**, 222 (1995).
129 Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415 (2004).
130 Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751 (2001).
131 Chandra, K. Statistical Time Division Multiplexing. *Handbook of Computer Networks* **1**, 579 (2011).
132 Padovan-Merhar, O. *et al.* Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Mol. Cell* **58**, 339 (2015).
133 Tinevez, J.-Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80-90, doi:https://doi.org/10.1016/j.ymeth.2016.09.016 (2017).
134 Strack, R. L., Disney, M. D. & Jaffrey, S. R. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nature methods* **10**, 1219-1224 (2013).
135 Song, W., Strack, R. L., Svensen, N. & Jaffrey, S. R. Plug-and-play fluorophores extend the spectral properties of Spinach. *Journal of the American Chemical Society* **136**, 1198-1201 (2014).
136 Ellis, R. J. Macromolecular crowding: obvious but underappreciated. *Trends in biochemical sciences* **26**, 597-604 (2001).
137 Zimmerman, S. B. & Trach, S. O. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of molecular biology* **222**, 599-620 (1991).
138 Tan, C., Saurabh, S., Bruchez, M. P., Schwartz, R. & LeDuc, P. Molecular crowding shapes gene expression in synthetic cellular nanosystems. *Nat Nano* **8**, 602-608, doi:10.1038/nnano.2013.132 http://www.nature.com/nnano/journal/v8/n8/abs/nnano.2013.132.html#supplementary-information (2013).
139 Dar, R. D., Hosmane, N. N., Arkin, M. R., Siliciano, R. F. & Weinberger, L. S. Screening for noise in gene expression identifies drug synergies. *Science* **344**, 1392-1396 (2014).
140 Minton, A. P. Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers* **99**, 239-244 (2013).
141 Rivas, G. & Minton, A. P. Macromolecular Crowding In Vitro, In Vivo, and In Between. *Trends in Biochemical Sciences* (2016).
142 Bakshi, S., Siryaporn, A., Goulian, M. & Weisshaar, J. C. Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. *Mol Microbiol* **85**, 21-38, doi:10.1111/j.1365-2958.2012.08081.x (2012).

143 Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340, doi:10.1038/nature10879 (2012).
144 Chizzolini, F. *et al.* Cell-free translation is more variable than transcription. *Acs Synth Biol* **6**, 638-647 (2017).
145 Balázsi, G., van Oudenaarden, A. & Collins, J. J. Cellular decision making and biological noise: from microbes to mammals. *Cell* **144**, 910-925 (2011).
146 González, C. *et al.* Stress-response balance drives the evolution of a network module and its host genome. *Molecular systems biology* **11**, 827 (2015).
147 Singh, A., Razooky, B., Cox, C. D., Simpson, M. L. & Weinberger, L. S. Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical journal* **98**, L32-L34 (2010).
148 Hasty, J. & Collins, J. J. Translating the noise. *nature genetics* **31**, 13 (2002).
149 Kim, J. S. & Yethiraj, A. Effect of macromolecular crowding on reaction rates: a computational and theoretical study. *Biophysical journal* **96**, 1333-1340 (2009).
150 Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504-508 (2010).
151 Dauty, E. & Verkman, A. Molecular crowding reduces to a similar extent the diffusion of small solutes and macromolecules: measurement by fluorescence correlation spectroscopy. *Journal of molecular recognition* **17**, 441-447 (2004).
152 Stracy, M. *et al.* Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proceedings of the National Academy of Sciences* **112**, E4390-E4399 (2015).
153 Doktycz, M. J. & Simpson, M. L. Nano-enabled synthetic biology. *Molecular systems biology* **3**, 125 (2007).
154 Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the bacterial cell: a molecular approach*. Vol. 20 (Sinauer Associates Sunderland, MA, 1990).
155 Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia coli and Salmonella: cellular and molecular biology* **2**, 1553-1569 (1996).
156 Erickson, H. P. Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online* **11**, 32-51, doi:10.1007/s12575-009-9008-x (2009).
157 Bryant, J. A., Sellars, L. E., Busby, S. J. & Lee, D. J. Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic acids research* **42**, 11383-11392 (2014).

158 Hebisch, E., Knebel, J., Landsberg, J., Frey, E. & Leisner, M. High Variation of Fluorescence Protein Maturation Times in Closely Related Escherichia coli Strains. *PLoS ONE* **8**, e75991, doi:10.1371/journal.pone.0075991 (2013).
159 Thattai, M. & Van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences* **98**, 8614-8619 (2001).
160 Chizzolini, F., Forlin, M. & Mansy, S. S. Cell-free translation is more variable than transcription. *bioRxiv*, doi:10.1101/073841 (2016).
161 Caveney, P. *et al.* Self-organization controls expression more than abundance of molecular components of transcription and translation in confined cell-free gene expression. *bioRxiv*, doi:10.1101/401794 (2018).
162 Llopis, P. M. *et al.* Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77-81, doi:10.1038/nature09152 (2010).
163 Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13-23 (2017).
164 Golding, I. & Cox, E. C. Physical nature of bacterial cytoplasm. *Physical review letters* **96**, 098102 (2006).
165 Plochowietz, A., Farrell, I., Smilansky, Z., Cooperman, B. S. & Kapanidis, A. N. In vivo single-RNA tracking shows that most tRNA diffuses freely in live bacteria. *Nucleic acids research* **45**, 926-937 (2016).
166 Pothoulakis, G., Ceroni, F., Reeve, B. & Ellis, T. The spinach RNA aptamer as a characterization tool for synthetic biology. *Acs Synth Biol* **3**, 182-187 (2013).
167 Pothoulakis, G. & Ellis, T. in *Methods in enzymology* Vol. 550 173-185 (Elsevier, 2015).
168 Kim, J., Jeon, C., Jeong, H., Jung, Y. & Ha, B.-Y. A polymer in a crowded and confined space: effects of crowder size and poly-dispersity. *Soft Matter* **11**, 1877-1888 (2015).
169 Marenduzzo, D., Finan, K. & Cook, P. R. The depletion attraction: an underappreciated force driving cellular organization. *J Cell Biol* **175**, 681-686 (2006).
170 Hansen, M. M., Desai, R. V., Simpson, M. L. & Weinberger, L. S. Cytoplasmic amplification of transcriptional noise generates substantial cell-to-cell variability. *Cell Systems* (2018).
171 Pardee, K. *et al.* Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* **165**, 1255-1266 (2016).
172 Garcia, D. C. *et al.* Elucidating the potential of crude cell extracts for producing pyruvate from glucose. *Synthetic Biology* **3**, ysy006 (2018).
173 Grest, G. S. & Kremer, K. Molecular dynamics simulation for polymers in the presence of a heat bath. *Physical Review A* **33**, 3628 (1986).
174 Weeks, J. D., Chandler, D. & Andersen, H. C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *The Journal of chemical physics* **54**, 5237-5247 (1971).

175 Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* **117**, 1-19 (1995).
176 Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Modelling and Simulation in Materials Science and Engineering* **18**, 015012 (2009).

VITA

S. Elizabeth Norred was born in Chattanooga, TN to Robert and Jennifer Norred. After attending Baylor School in Chattanooga, she received her Bachelor of Science in Biomedical Engineering from Georgia Institute of Technology in 2011. She received her Master of Engineering in 2012 from the University of Tennessee, Knoxville, where she specialized in cell and tissue engineering. After interning at Oak Ridge National Laboratory in the Center for Nanophase Materials Science, she joined the Bredesen Center for Interdisciplinary Research and Education. She performed her research at ORNL in the Nanofabrication Research Laboratory, in collaboration with the Biological and Nanoscale Systems Group. She graduated with a PhD in Energy Science and Engineering in December 2018.