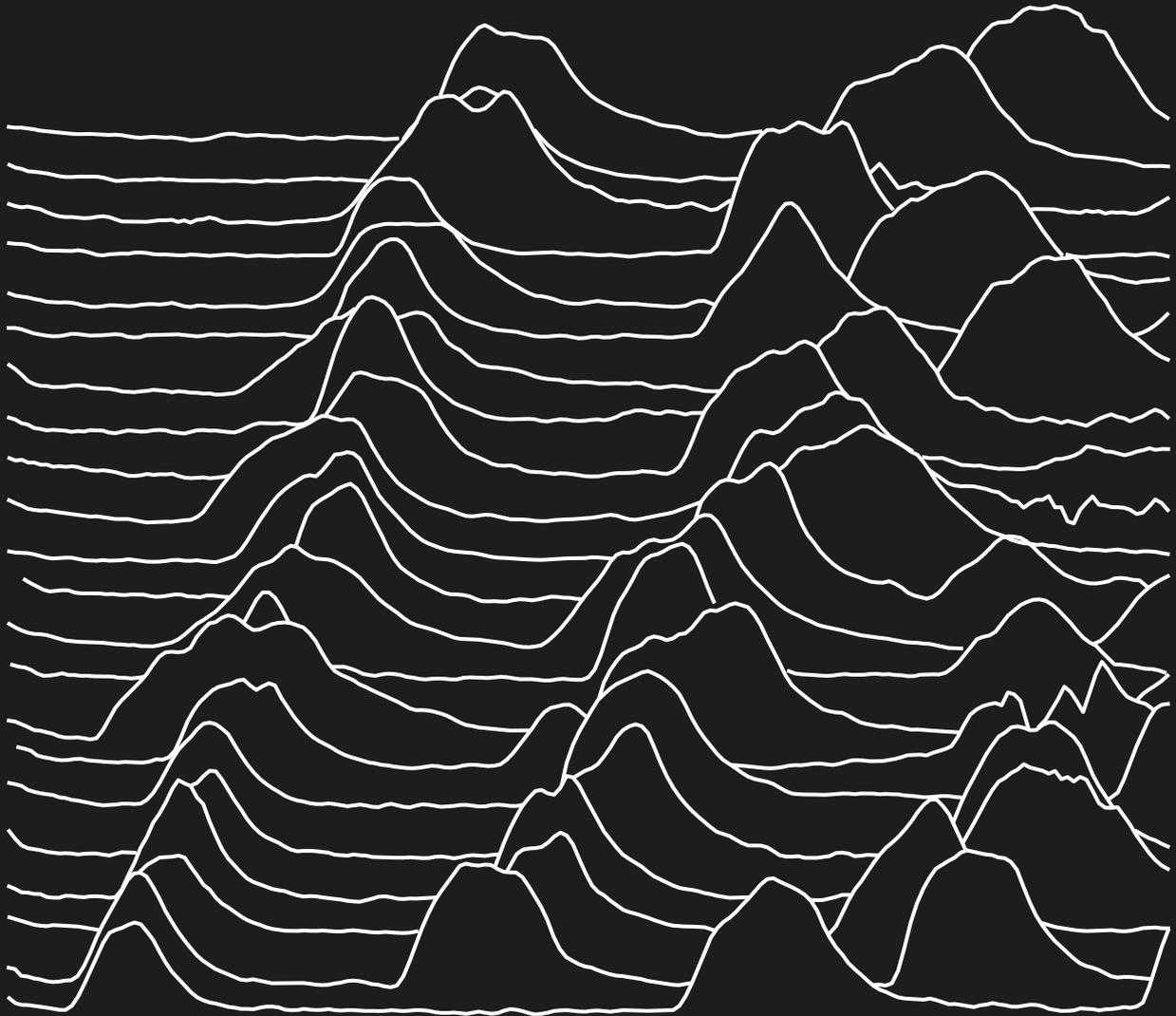


Programming Dynamic Bacterial Behaviour with CRISPRi-Based Genetic Circuits



John Henningsen



Technische Universität München



Technische Universität München

Fakultät für Physik

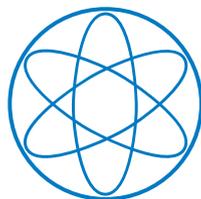
**Programming Dynamic Bacterial Behaviour with
CRISPRi-Based Genetic Circuits**

John Henningsen

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzende: Prof. Dr. Karen Alim
Prüfer der Dissertation: 1. Prof. Dr. Friedrich Simmel
2. Prof. Dr. Ulrich Gerland

Die Dissertation wurde am 23.11.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 07.12.2021 angenommen.



I like to think (and
the sooner the better!)
of a cybernetic meadow
where mammals and computers
live together in mutually
programming harmony
like pure water
touching clear sky.

I like to think
 (it has to be!)
of a cybernetic ecology
where we are free of our labors
and joined back to nature,
returned to our mammal
brothers and sisters,
and all watched over
by machines of loving grace.

Richard Brautigan (1967)

Contents

1	Introduction	1
2	Background	3
2.1	Engineering of Growing Bacteria	4
2.1.1	Bacterial Growth	4
2.1.2	Control of Growing Bacteria	8
2.1.3	Engineering with Evolution	10
2.2	Systems Biology and Biological Oscillators	13
2.2.1	Systems Biology	13
2.2.2	Oscillator Theory	15
2.3	Information in Biology	21
2.3.1	Introduction of Shannon Entropy	21
2.3.2	Interpretations of Mutual Information	24
2.3.3	Measuring Information	26
3	Results	29
3.1	A Synthetic Bacterial Clock with a Hybrid Feedback Loop Containing dCas9-sgRNA	30
3.1.1	Circuit Design	31
3.1.2	Microfluidics Experiments	33
3.1.3	Single Cell Data Analysis	36
3.1.4	Simulation	38
3.1.5	Synchronization and Phase Stability	39
3.1.6	Temporal Information and Reliability as Clock	41
3.1.7	Interplay with Bacterial Growth	43
3.1.8	Discussion and Conclusion	44
3.2	CRISPRi-Induced Auxotrophy as Selection Principle in Bacterial Cultures	47
3.2.1	Growth in Different Bacterial Culture Systems	47
3.2.2	Amino Acid Auxotrophy Mediated by CRISPRi	56

3.2.3	Outlook	61
3.3	Fabrication of Sub-Micron Scale Microfluidic Chips	62
3.3.1	Fabrication	62
3.3.2	PDMS Chip Integrity	71
3.3.3	Loading Bacteria	72
3.3.4	Timelapse Setup	73
3.3.5	Optimization	74
3.3.6	Outlook	77
3.4	Sampling the Distance Dependence of Quorum Sensing Between Bacterial Colonies	79
4	Conclusion	86
5	Methods	88
5.1	Bacterial Oscillator Experiments	89
5.2	Lithography for Microfluidics	93
5.3	Auxotrophy Experiments	94
5.4	QS experiments	95
6	Appendix	97
6.1	Appendix Figures	98
6.2	Appendix Notes	103
	Bibliography	107
	Acknowledgements	122

1 Introduction

Over the last twenty years, the research field of synthetic biology made significant progress towards the goal to create, control and programme novel and useful biological systems. By drawing inspiration from a broad range of disciplines, synthetic biology takes a rational engineering approach to biology as opposed to unstructured tinkering or pure investigative research.

In spite of early challenges [1], a broad spectrum of success can be attributed to this approach since the foundational work in 2000 [2, 3]. In this work, we focus on synthetic genetic circuits, referring to systems of interacting biological parts in analogy to electrical circuits. Complex behaviour, such as a programmed response based on sensor input and logic or spatial differentiation, has been designed and implemented with genetic circuits in bacteria [4, 5], aided by standardization, characterization and automated design [6, 7]. Applications of the biosynthesis of therapeutics, food or textiles or diagnostics of infectious diseases have reached a commercial stage and are set to support a multibillion dollar industry [8].

The breakthrough discovery of CRISPR and subsequent improvements have enabled bio-engineers to target and manipulate nearly arbitrary locations on a genome by programming a guide RNA sequence. CRISPR interference (CRISPRi) in particular, based on the catalytically inactive dCas9, allows the transcriptional repression of both native genes of the bacterial host and artificially integrated genes [9]. CRISPRi has been utilized to make the construction of gene circuits more modular, predictable and standardized [10, 11].

In spite of these achievements, the scope of the effort to engineer biological systems remains limited with regards to the complexity of the synthetic construct and the respective environment [12]. Individual components have reached a high degree of sophistication, but how a large number of components can work together is still unsolved for the most part. The high number of interactions that need to be considered in more complex engineered systems call for new quantitative methods to aid prediction and design. For communities of diverse cells, ecological topics such as competition, cooperation

and self organization need to be integrated into the traditional engineering approach. This way, cells may be programmed to work together in a modular way for more robust and efficient function [13]. In natural *in vivo* contexts, how to work with and not against evolution is an open challenge. Furthermore, there are still open questions regarding dynamic genetic circuits that display varied behaviour over time in contrast to static circuits that quickly settle into a stable state. How can we programme robust biological clocks to coordinate the timing of other processes, how can a dynamical circuit adapt to a changing environment?

From a broader perspective, the field of synthetic biology is faced with a variety of concerns from society. While issues surrounding the level of popular trust in synthetic biology or corporate control and ownership cannot be dismissed, it is useful to focus on the aspects that are related to the nature of the technology itself and not its relation to general societal issues. Since living systems can reproduce, the uncontrolled spread of an engineered version is possible. For example, engineered pathogens can be considered to be dual-use with both beneficial and harmful applications. It may be prudent not to ignore or silence justified concerns but to engage in active risk management instead [14]. Some risks can be addressed in advance with technological solutions, e.g. synthetic genetic circuits have the potential to prevent the uncontrolled spread of a host system by sensing environmental conditions and reacting if they differ from a safe containment environment [15]. This highlights the need for basic research into suitable tools for the interaction with a complex environment that changes over time.

In this work, we integrate CRISPRi into genetic circuits to explore novel ways to connect them with native genes and the environment of the host. The dynamic behaviour of a bacterial oscillator is studied in detail with regards to its stability and relation to growth. The construction of sub-micron scale microfluidic chips enables the quantitative analysis of such dynamic bacterial circuits. Mathematical tools such as information theory are utilized to quantify the potential of the oscillator as a controller for temporal gene expression programs. The interaction of a CRISPRi based circuit with the growth environment of the bacterial host is achieved by interrupting metabolic pathways. Finally, the communication between spatially separated bacterial colonies is investigated.

2 Background

2.1 Engineering of Growing Bacteria

Growth is a characteristic feature of living systems such as bacterial cells. From a thermodynamic point of view, cells consume energy in order to resist the decay towards unordered equilibrium and maintain an ordered state. In this chapter we will cover fundamental aspects of bacterial growth and implications for the effort to engineer dynamic behaviour with genetic circuits.

2.1.1 Bacterial Growth

Understanding of bacterial growth is key for the engineering and analysis of genetic circuits in bacteria. The basic laws of bacterial growth were first described by Monod (1949) [16], who identified distinct phases of bacterial growth (Figure 2.1). In general, bacteria grow and divide with a variable rate, depending on the environmental conditions and availability of nutrients. Here, we focus on the phase of exponential growth, where bacteria are not limited by nutrient supply and divide with a constant rate. Therefore, they can be considered to be in a steady state. This greatly simplifies modelling, as even if many details of the host organism remain unknown, they can be assumed to be fixed during exponential growth. In natural environments however, bacteria may have to adapt to a volatile environment or spend long periods under challenging conditions without growing, severely limiting the applicability of the exponential model. If more than one species is co-cultured (microbial system), competition for resources or direct interactions may need to be considered. Simple examples of microbial systems are analysed in Murray (2002) [17]. Typically, such population dynamics models consist of a system of ordinary differential equations (ODEs) that describe the rate of change for the concentration of all relevant species. In environments that are not well mixed the spatial context has to be considered as well, in which case partial differential equations (PDEs) may be employed. For a low number of cells, birth and death processes may have to be modelled stochastically.

Continuous Culture Next, we will discuss two important factors for bacterial growth experiments: The cultivation method and the growth medium. The most common bulk cultivation method is batch culture, where a fixed reservoir of growth medium is inoculated with a small dose of bacteria. Those bacteria then grow until the nutrients of the growth medium are used up. Most often, a shaking incubator is used to ensure the optimal temperature and oxygen conditions in the liquid medium and to keep the culture

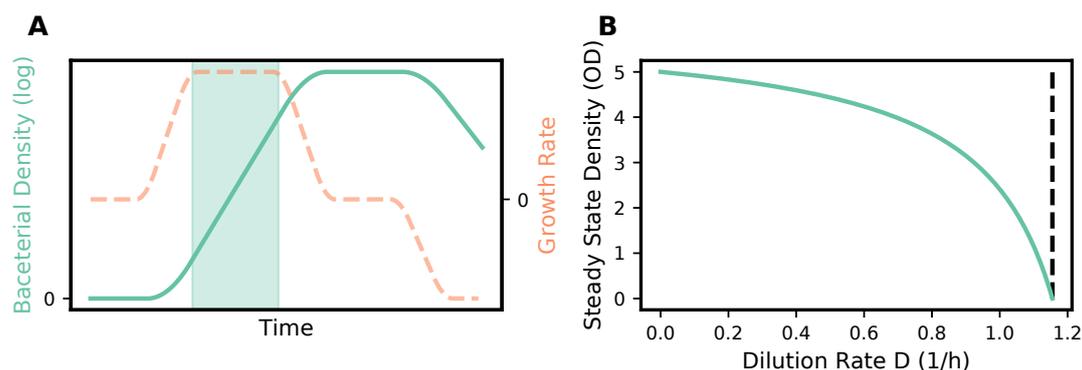


Figure 2.1: (A) Overview of growth phases in batch culture. The shaded area marks the exponential phase, where the growth rate (dotted line) is constant and the bacterial density increases exponentially (straight line with logarithmic axis). During exponential phase, the bacteria can be assumed to be in a steady state. (B) Continuous culture: Variation of the steady state density depending on the dilution rate. Above a critical dilution rate, bacterial growth cannot keep up with dilution and the culture is ‘washed out’.

well mixed. A different approach is continuous culture [18–20]. For a continuous culture, fresh medium is supplied throughout the cultivation. At the same time waste is removed. A typical continuous culture setup can be realized by adding an additional inflow and outflow to a culture vessel. A bioreactor, a device optimized to mix, aerate and heat a culture in an attached vessel, is the most common basis for a continuous culture setup. If the inflow and outflow are set to the same constant rate, this setup is also referred to as chemostat [21]. In a chemostat, bacteria will grow while the culture is being diluted at the same time. With increasing culture density, growth will slow down due to gradual nutrient depletion. A steady state is reached when the growth of the culture is exactly matched by dilution of the culture. Unlike the exponential phase during batch cultivation, this steady state can in principle be maintained indefinitely. This makes a chemostat attractive for the reliable quantitative study of bacteria and, more importantly, allows the long term observation of dynamic systems in bacteria. Mathematically, a chemostat can be described with the following set of ODEs:

$$\dot{N} = \mu N - DN = (\mu - D)N \quad (2.1)$$

$$\dot{S} = D(S_R - S) - \frac{\mu N}{Y}. \quad (2.2)$$

The first equation describes the change in the number of cells N through growth

with rate μ and dilution with rate D . The second equation describes the change of the substrate concentration S (assuming the growth medium contains a single essential substrate). Dilution removes substrate partially consumed by the cells and adds fresh substrate with concentration S_R . The efficiency of the conversion of substrate into new cells is given by the yield constant Y . The growth rate μ is proportional to on the substrate concentration until reaching saturation at a maximal rate of μ_m , with the half maximal growth rate given by the saturation constant K_S . The growth rate can be linked to the doubling time t_d :

$$\mu = \mu_m \frac{S}{K_S + S} \quad (2.3)$$

$$\mu = \frac{\ln_e 2}{t_d}. \quad (2.4)$$

With suitable parameters, equations (2.1) and (2.2) can lead to a steady state. Setting (2.1) equal to zero shows that, for greater than zero N , the substrate concentration needs to be at a level so that the growth rate is exactly equal to the dilution rate. For example, if a culture is diluted with one volume per hour that directly implies a doubling time of 42 min. At the steady state, cell and substrate concentration are as follows:

$$N_{ss} = Y (S_R - S_{ss}) \quad (2.5)$$

$$S_{ss} = K_S \frac{D}{\mu_m - D}. \quad (2.6)$$

In experiments, the steady state concentration is usually influenced by changing the dilution rate and the substrate concentration. Figure 2.1 B shows the non-linear relationship between the steady state culture density and the dilution rate. Above a critical value, the medium can no longer support a growth rate that matches the dilution rate, causing the culture to ‘wash out’. A change in the substrate concentration however has a linear effect on the steady state culture density. Adapting the medium is therefore, if possible, the preferred method to tune the steady state.

Microchemostats Instead of culture vessels on the litre scale, chemostats may also be realized at lower volumes (about 10 ml). Such small and cheap chemostat devices, often constructed using DIY methods, are more conducive to parallelised setups of continuous cultures for increased experimental throughput [22–25].

Using a set of very different techniques and on a much smaller scale, continuous culture

setups are also possible in microfluidic devices for the cultivation of bacteria [26, 27]. Microfluidic devices are an arrangement of small channels and cavities (down to the micrometer scale) that contain very low volumes of fluids [28–30]. Suitable microfluidic devices are fabricated with soft lithography, where a biocompatible polymer such as PDMS (polydimethylsiloxane) is cast from a casting mould, leaving a relief imprint that forms the channels. The casting mould in turn can be accurately fabricated down to the sub micrometer scale using lithography techniques. Different implementations of chemostats in microfluidic devices, so called microchemostats, have been demonstrated in the literature [31–34]. These microchemostats usually consist of a cavity that is suitable for trapping bacterial cells in place while a separate channel supplies a fresh stream of growth medium. In a different approach, the culture is pumped around a loop with intermittent supply of fresh medium to prevent biofilm build up [35]. Biofilm build up is a challenge for all long term cultivation methods because the biofilm layer can physically inhibit the function of the device and interfere with the growth of the suspended culture (e.g. by competing for nutrients) [36, 37]. A translucent microfluidic chip allows the continuous observation of growing bacterial cells *in situ* with optical (fluorescence) microscopy. The resulting time resolved data for single cells enables quantitative studies of cell growth and the dynamic effect of synthetic genetic circuits.

Growth Medium and Auxotrophy Growth media can be broadly categorized according to two features: Minimal or rich and defined or undefined. Minimal medium only includes a minimal number of ingredients essential for growth, such as one type of sugar as energy and carbon source and salts as source for other essential inorganic elements. Strictly according to definition, no single ingredient can be left out from the medium for it to still support bacterial growth. The opposite of a minimal medium is a rich medium, which contains many non-essential ingredients to support fast growth. The second feature for classification is the knowledge about the ingredients of the growth medium: A defined medium consists only of purified ingredients with precisely known concentrations. It is therefore suitable for studies concerned with the relation of growth and specific substances. Minimal medium is also useful to preclude any unwanted side effects on measurements, e.g. background fluorescence. The opposite, undefined medium, contains ingredients such as lysate of other organisms or products of other biotechnological processes, for which the exact individual substance concentrations are not known. Rich undefined media based on yeast extracts are popular because they support fast growth with high convenience and low cost.

Wild type *E. coli* is able to synthesize all molecules required for growth from the nutrients contained in a minimal medium. A strain that cannot synthesize a specific molecule that is essential for its growth is termed *auxotroph*. Auxotrophy implies that a strain can only grow in medium that supplements the specific molecule.

2.1.2 Control of Growing Bacteria

While historically the concept of control theory was first applied to mechanical machines [38], developments such as cybernetics later extended it to other domains such as biology (‘control and communication in the animal and the machine’) [39]. Naturally, ideas and models from control theory are also useful for the engineering of biological systems [40–42]. Control theory seeks to improve the performance, stability and robustness of a system. For this purpose, a controller may be constructed to act on the system depending on its state. For systems implemented in bacteria, a controller can be implemented as a genetic circuit integrated into the cell. A basic example is given by a negative feedback configuration that is employed to reduce the stochastic variation of the expression level of a protein of interest [43, 44]. Such a controller is realized by coupling the expression of the protein of interest to the expression of a transcription factor that represses the expression of both proteins. This way, the desired set point for the output is reached faster and more robustly.

Multicellular Control Moving outside of the context of a single host cell, it may be required to program multicellular control systems to design population level effects or spatial patterning. A key module that can be used for this purpose is quorum sensing [45, 46]. Quorum sensing is based on a small molecule that is able to diffuse through cell walls and the associated synthesis and transcription factor proteins. Through such a diffusible signal, individual cells get access to an ensemble state of the population. To reach a high threshold signal concentration (and e.g. trigger a feedback process), a cell needs to be surrounded by a dense culture of signal producing cells. In the Lux quorum sensing module adapted from the marine bacterium *Vibrio Fischeri*, the AHL molecule (acyl homoserine lactone) is synthesized by a single enzyme. It can bind to a specific transcription factor to activate transcription from the corresponding promoter. Other, related quorum sensing systems have been engineered for orthogonal functionality and allow multiple signal pathways in parallel [47, 48]. The population control system by You et al. consists of a circuit based on Lux quorum sensing, which mediates negative feedback from high population density [49]. The negative feedback is in the form of a

toxic gene, which closes the control loop (the population size is both input and output of the control system). Thus the population density is regulated to a lower steady state level. As suggested by theory, this multi-step negative feedback loop can also produce oscillations in the population density when cultured under specific conditions in a microfluidic chemostat [35].

Multicellular control systems can also be extended to multiple interacting bacterial species that are cultured together (microbial systems) [50]. Engineered circuits with orthogonal signaling modules coupled to outputs that change the relative fitness of a species can control the population dynamics of the microbial system. In an engineered predator-prey system, Balagadde et al. used two orthogonal quorum sensing modules to engineer the interactions between two *E. coli* species [51]. By placing a toxic gene under a promoter sensitive to the quorum sensing signal produced by the other species, the growth and survival of one species becomes dependent on the population density of the other species. The presence/absence of a negation of the signal marks the difference between predator/prey species. A similar approach based on two quorum sensing modules was used by Scott et al. to stabilize the co-culture of two species [52]. An alternative to quorum sensing signals for the control of multi species population dynamics may be based on metabolic engineering, where the interactions stem from competition for shared resources or the release of metabolites that are either essential or harmful to growth. Importantly, multicellular coordination also enables the programming of spatial behaviour, which will not be covered here.

In Silico Control Another possibility for the control of bacterial systems is to make use of a computer as control unit, called *in silico* control. With a computer the programming of a controller is very easy, but the implementation challenge transfers to the now required interface between computer and bacteria. The measurement of the state of bacteria is usually based on fluorescent reporters, which can be observed under a spectrometer, microscope or flow cytometer. Another potential point of measurement is the density of a culture, which works based on absorbance at 600 nm wavelength, changes in light scattering or changes in impedance [53]. Based on the measurement, the control computer infers the state of the bacterial system and calculates the input required to reach the desired state. This input can be actuated in multiple ways: Light exposure, change in osmotic pressure, redox potential, temperature or inducer concentration [40]. The use of a computer in the control loop allows for more sophisticated control schemes than can be realized *in vivo*. Initial filtering steps may be employed to robustly process noisy

measurement data. If a predictive model of the system is available, advanced methods like model predictive control (MPC) can be used [54, 55]. With MPC, the simulated result of a possible input to the measured state can be compared against the desired set point. This presents an optimization problem that can be solved for the optimal input for every time step.

Retroactivity From the viewpoint of control theory, unintended effects in a genetic network can be predicted. *Retroactivity* to the output means the load effect that the connection of a downstream node (e.g. a reporter) has on the main system. It may be quantified by considering retroactivity effects as additional signal from the downstream node to the upstream signal that binds away product of the upstream signal. Based on this model, retroactivity effects are mostly sensitive to the affinity of the output signal to the downstream binding sites and the abundance of signal compared to the number of downstream binding sites [56]. Influence of such effects may be limited with mitigation strategies like an insulation device with high gain negative feedback, or separation of timescales between upstream and downstream systems [41, 56].

2.1.3 Engineering with Evolution

While it may be useful to implement concepts from conventional control engineering, a number of characteristic aspects of synthetic biological systems cannot be compared to other disciplines such as mechanical or electrical engineering. A growing, self reproducing system with genetic heritability undergoes evolution¹.

Here we will discuss some examples of how (continuous in vivo) evolution is used in a lab towards a specific goal (see Castle et. al. (2021) [58] for a recent review). In the long-term evolution experiment by Lenski, an *E. coli* culture has been diluted daily for over 30 years. The continued bulk growth of the culture generates diversity and selects for faster growth under the given conditions. As a result, a non-linear (decelerating) increase in relative fitness has been observed while the number of mutation rose linearly [59]. Using the same principle, the expansion speed of a motile *E. coli* strain growing on an agar plate can be enhanced with a selection procedure if a new round of growth is inoculated by picking from the leading edge of the previous colony [60, 61]. This enhancement comes at the cost of a reduced growth rate. The resulting library of strains with different expansion speeds may be used to generate patterns from the competitive co-culture of multiple strains. A high throughput automated culture platform has been

¹‘Nothing in biology makes sense except in the light of evolution.’ - Theodosius Dobzhansky [57]

used to screen how populations evolve to adapt to different environments by precisely applying various selection pressures [24], underlining the law that states ‘you get what you screen for’ [62]. Recently, a ‘design, mutate, screen’ approach was utilized to modify a bacterial oscillator circuit based on synchronized lysis. Libraries were generated with directed mutagenesis and screened with multiplexed microfluidics to tune the dynamics of the oscillator circuit [63]. For a similar purpose, single cells were isolated selectively with optical tweezers after time-lapse imaging in a mother machine microfluidic chip [64]. This technologically involved approach allowed the selection of mutants on a single cell level and was utilized to further improve the repressilator circuit [3, 65], achieving a theoretical phase stability of 200 h for a drift of half a period.

Since the expression of a synthetic genetic circuit carries a fitness burden for the host, loss of function frequently occurs due to evolutionary pressure. The stability of a circuit in this context may be improved with suitable design strategies like avoiding homologous sequences and keeping expression levels as low as possible [66]. It has been shown that loss of function may be reversed when the function becomes beneficial to the host after a change in environmental conditions [67].

Cellular Resource Allocation For a synthetic genetic circuit inside a bacterial host, the interactions between circuit and host are another complicating factor. The availability of cellular resources strongly depends on the current state of the cell. Resources for protein synthesis may be allocated to produce more ribosomes or towards nutrient uptake and metabolism machinery. This allocation is proportional to the growth rate of the cell, which determines the priority of either fraction. Without knowledge of the details of this complex regulation phenomenon, the resource allocation may be predicted with simple growth laws that express its dependence on the growth rate [68–70]. Overexpressing endogenous proteins reduces the resources available for cell growth and therefore has a negative effect on the growth rate. Conversely, the growth rate determines the ‘decay’ of protein concentrations through dilution and thus influences both steady state levels of a continuously produced output or the characteristic time scale of a dynamic genetic circuit. By considering growth rate as an input to the system, these non-trivial interactions with a dynamic circuit have been used to sense variations of microbiota and inflammation in a mammalian gut [71].

To stabilize the function of a synthetic genetic circuit in this dynamic environment, a number of different strategies have been proposed. By identifying promoters that are activated under stress, a negative feedback control based on dCas9 repression of a synthetic

circuit in response to a burden was implemented [72]. Modelling the competing resource demands and resource sensitivity of synthetic circuit modules, indirect interactions between modules can be understood and tuned by varying the resource demand (e.g. with different DNA copy number). [73, 74] Another potential mitigation strategy for resource competition effects is to introduce orthogonal resources decoupled from host resources like T7 transcription machinery [75].

Stochastic Nature Biological systems are complex and stochastic in nature, which can cause significant variation in the expression of a genetic circuit between individual cells or experiments with slightly different conditions. In the context of synthetic circuits, a number of concrete effects that contribute to this stochasticity have been identified, such as low DNA/molecule copy numbers or translational bursting [76, 77]. Overall, protein expression often follows a log-normal distribution. This can be explained by the fact that protein expression is the end result of a number of independent random processes (i.e. reactions with associated rate) that are connected by multiplication. The logarithm of the product is a sum (of independent random variables) and therefore normally distributed (central limit theorem) [78, 79].

Stochasticity in addition to high context sensitivity [80] (and general complexity under natural conditions [81]) call for suitable experimental strategies to cope with these limitations. A reduction in experimental variability may alleviate lack of mechanistic understanding. By using the same host strain under carefully fixed growth conditions (details matter), a consistent output may be achieved. These fixed conditions can be used to characterize the system and tune a phenomenological model to make useful predictions. In contrast to individual bulk measurements, high throughput experiments and single cell measurements can resolve population distributions, which should better represent the stochastic nature of the underlying system.

2.2 Systems Biology and Biological Oscillators

The goal of this chapter is to provide a background for the quantitative analysis and design of (synthetic) genetic networks, with particular focus on genetic oscillators.

2.2.1 Systems Biology

The field of systems biology offers qualitative and quantitative insight into the function of organisms and their constitutive parts that can be readily transferred to engineered biological applications. Systems biology is a quantitative, interdisciplinary approach to the study of biology with emphasis on mathematical models. A top-down viewpoint allows us for a moment to disregard the structure of proteins, the base sequences of nucleic acids or soft matter physics, which in principle govern the molecules that make up life. Instead, suitable layers of abstraction give rise to principles for biological systems found in nature that can readily be applied to engineer similar behaviour. This section is based on the seminal book by Alon (2019) [82].

The model presented here focuses on transcription factors (TFs), proteins that regulate the transcription rate of specific genes and can change with signal input. Viewed mathematically, transcription factors mediate how an output (the transcription of a gene under a specific promoter) is affected by an input signal (e.g. a small molecule from the environment). Disregarding a mechanistic description, a transcription factor can be characterized by an input function. Mostly, this function is monotonic and bounded, which implies a sigmoidal shape (s-shape). For experimental data, the production rate of a gene p depending on the concentration of transcription factor R is usually well described by a Hill function of the form

$$p([R]) = \alpha \frac{[K]^n}{K^n + [R]^n},$$

with the hill coefficient n , the maximal production rate α and the concentration K leading to half maximum production rate (often referred to as K_D because of similarity with a dissociation constant) [83]. Written as above, it describes a repressor, a transcription factor that decreases the promoter activity with increasing concentration. A reversed sigmoidal would describe an activator. The Hill coefficient n is supposed to represent the cooperative action of n molecules in the regulation [84]. For high values of n , the Hill function starts to resemble a step function with the threshold K . In fact, the Heaviside step function is an even simpler model that can be useful to find approximate analytical

solutions of the dynamics of a component. Square brackets denoting the concentration of a species are omitted for brevity from here on.

The dynamics of the output protein concentration Y can be described by a differential equation². In simple cases we can find an analytical solution. Often, however, only numerical solutions are available. Let's consider a very simple example: Production of a protein Y regulated by a step function. We must also consider a degradation term with constant μ , which describes the reduction in concentration by dilution or protein degradation. The absence of a degradation term would lead to unbounded, non-physical solutions.

$$\dot{Y} = \alpha\Theta(X - K) - \mu Y$$

If the activator X is above the threshold K , the protein Y is produced until it reaches a steady state level, where degradation balances out production.

$$Y_{ss} = \frac{\alpha}{\mu}$$

Finally, the time course of the protein concentration $Y(t)$ can be obtained by solving the differential equation.

$$Y(t) = Y_{ss} \left(1 - e^{-\mu t}\right)$$

On a higher abstraction level, a network of transcription factor interactions can be analysed as a graph. Such an abstraction is useful to find commonly occurring modules and general principles. Given a certain number of nodes (transcription factors, signals and genes), the exponential nature of combinatorics means those nodes could be connected in a huge number of ways. A detailed graph of the *E. coli* transcription network has been established with bottom up molecular biology techniques. A comparison between the *E. coli* transcription network and a randomly generated network with the same number of nodes and edges shows that only a small number of subgraphs are commonly discovered and chosen by evolution. These commonly found subgraphs are referred to as network motifs [85, 86]. Because of their repeated independent occurrence across many domains they can be viewed as an example of convergent evolution, and thus a robust solution for tasks where noisy signals are processed with noisy components to initiate appropriate actions.

A typical motif is the feed forward loop (FFL). It consists of 3 transcription factors: The first TF X regulates the second, the second TF Y regulates the third; additionally

²For low molecule counts, a stochastic description is better suited.

the first TF also regulates the third TF Z . Here we will consider the coherent feed forward loop type 1 (C1-FFL), where all transcription factors are activators. Building on the simple example, we can understand how the C1-FFL fulfills a precise *dynamical* function. Assuming the two activators X and Y combine with AND logic for the input function of Z and using a simple step function model, we can write the differential equations:

$$\begin{aligned}\dot{Y} &= \alpha_Y \Theta(X - K_{XY}) - \mu_Y Y \\ \dot{Z} &= \alpha_Z \Theta(X - K_{XZ}) \Theta(Y - K_{YZ}) - \mu_Z Z,\end{aligned}$$

with subscripts denoting the affiliation of the constants. Assuming the input X crosses the threshold K_{XY} at $t = 0$, we know that the concentration of Y will be

$$Y(t) = \frac{\alpha_Y}{\mu_Y} \left(1 - e^{-\mu_Y t}\right)$$

In turn, Z will become active after the concentration of Y also passes the threshold K_{YZ} . Using the above formula, the time delay T_d can thus be written as:

$$T_d = -\frac{1}{\mu_Y} \log \left(1 - \frac{K_{YZ} \mu_Y}{\alpha_Y}\right)$$

This simplified model thus shows how a basic network motif, the C1-FFL with AND logic, functions as delay element. The delay of the expression of output Z may be tuned through evolutionary adaptation primarily by changing the ratio between the activation threshold of the promoter for Z and the steady state concentration of activator Y . Conversely the AND logic of the Z promoter means that a removal of input X leads to an immediate halt in the expression of Z . Taken together, this motif filters out input pulses shorter than T_d and only longer pulses lead to the activation of Z . It has been shown that *E. coli* makes use of this C1-FFL motif in the regulation of its sugar metabolism, where the investment into the machinery for processing the less valuable sugar arabinose is only made in a *persistent* shortage of the preferred sugar glucose [87, 88].

2.2.2 Oscillator Theory

The systems biology approach is useful to understand rhythmic, periodic behaviour in living systems. The principles behind such biological oscillators can be understood by studying the interactions of its components. Furthermore, collective behaviour such as synchronization depends on the interactions of individual oscillator systems. A detailed

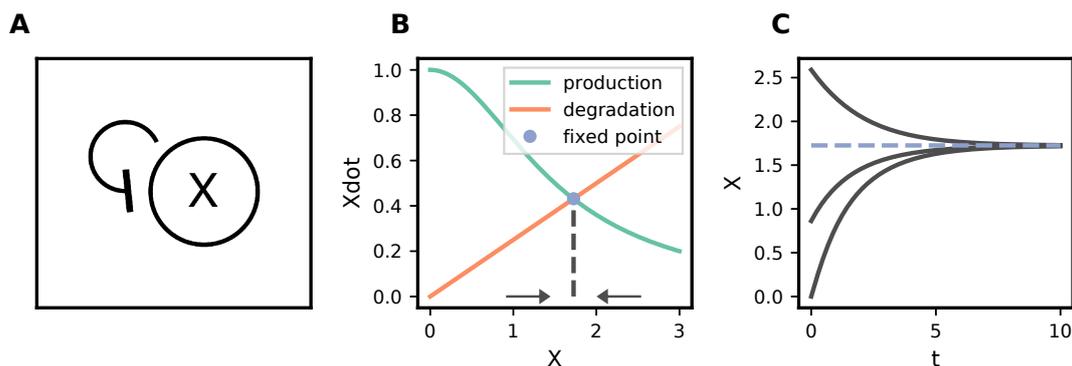


Figure 2.2: (A) Negative feedback loop with one component (Equation 2.7). (B) Rate plot showing the contribution of the production and degradation term to the rate of change. The intersection marks a stable fixed point, for points on the left or right of the fixed point the rate of change is determined by the difference between the production and degradation curve (arrows). (C) Time traces of example trajectories. As shown above, the system decays exponentially towards the fixed point. Parameters: $n = 2$, $K = 1.5$, $\alpha = 1$, $\mu = 0.25$.

introduction into the theory and models for biological oscillators is given by Forger (2017) [84]. In this section, we review the basic principles that explain the function of biological oscillators and can guide the design of synthetic oscillators [82, 89].

Principles of Biological Oscillators

Negative Feedback The most fundamental requirement for oscillations is *negative feedback*, which causes high concentrations to revert back to a lower starting point. The simplest motif containing negative feedback would be a single component that inhibits its own production. A rate plot (Figure 2.2 B), based on the following ODE, qualitatively explains the dynamics of this system:

$$\dot{X} = \alpha \frac{K^n}{K^n + X^n} - \mu X. \quad (2.7)$$

The intersection between the curve of the production and degradation term marks a stable fixed point (positive/negative total rate of change left/right of the fixed point). Therefore this system does not oscillate. As exemplified by the rate plot, the system cannot overshoot the fixed point in the one dimensional phase space (only monotonically approach it), hence this result is universally true for one component dynamical systems.

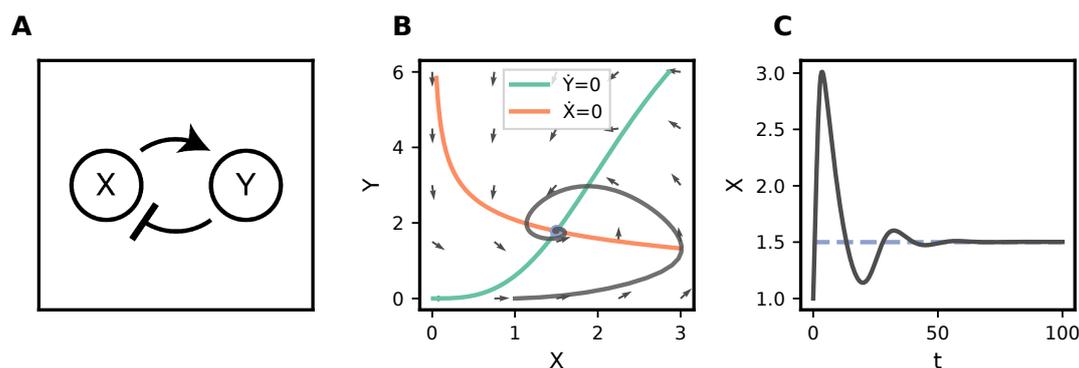


Figure 2.3: (A) Two component negative feedback loop with activator X and repressor Y (Equation 2.8). (B) Phase plane of the two components. The nullclines, marked with $\dot{X} = 0$ and $\dot{Y} = 0$, show where the production (depending on the other component) and the degradation of X or Y , respectively are in equilibrium and hence one component does not change. The intersection of the nullclines (red dot) marks the stable fixed point of the system. Quivers indicate the direction of change at different points in the phase space, suggesting spiral trajectories in the phase space. The trajectory of a solution from a random starting point, a collapsing spiral, is overlaid in black. (C) Time trace of the example solution trajectory. Oscillations quickly decay to the stable fixed point. Parameters: $n_X = n_Y = 3$, $K_X = 1$, $K_Y = 2.5$, $\alpha_X = \alpha_Y = 1$, $\mu_X = \mu_Y = 0.1$.

Delay In addition to negative feedback, there exists a second requirement for oscillations: The rates at the current point in time have to depend on a concentration at a previous point in time, requiring some sort of *memory*. This way, given the right parameters, the system will ‘overreact’ to a deflection from the fixed point. While approaching the fixed point after the deflection, the ‘resetting force’ is overestimated because it is based on old information about the past position, which was farther away. Hence, the system always overshoots the fixed point and instead orbits around it, leading to sustained oscillations. Such a memory feature can be realized as a delay caused by additional components in the negative feedback loop. Let’s consider a two component negative feedback motif, with activator X and inhibitor Y :

$$\begin{aligned}\dot{X} &= \alpha_X \frac{K_X^{n_X}}{K_X^{n_X} + Y^{n_X}} - \mu_X X \\ \dot{Y} &= \alpha_Y \frac{X^{n_Y}}{K_Y^{n_Y} + X^{n_Y}} - \mu_Y Y.\end{aligned}\tag{2.8}$$

Here, we modelled the interactions with nonlinear Hill input functions (note the different numerator for repression and activation) and linear degradation terms. This two component motif limited to one positive and one negative interaction does not produce

sustained oscillations. While simple motifs like the C1-FFL can be solved analytically, the nonlinear and coupled differential equations that describe more complex motifs can be difficult to solve with an analytical approach, but may be analysed with a graphically. In the phase plane, the two dimensional phase space of components X and Y , the solutions to the equation $\dot{X} = 0$ or $\dot{Y} = 0$ may be represented by a line. These *nullclines* mark points where the rate of change of one component is zero. The monotonic nullclines of equations (2.8), with suitable parameters, intersect in a single fixed point (Figure 2.3 B), where the rate of change is zero for both components. The quivers (little arrows) show how trajectories flow vertically close to the X -nullcline and horizontally close to the Y nullcline, overall hinting at circular or spiral trajectories. An exemplary numerical solution is attracted by the fixed point and spirals into it, which suggests that this fixed point is stable. We can prove the stability of the fixed point with linear stability analysis.

Linear stability analysis (see Appendix for an overview) of equations (2.8) shows that the two component negative feedback loop can have a non zero imaginary part. The partial derivatives of the right hand sides with respect to the other component $\partial_Y f$ and $\partial_X g$ have opposite signs, negative and positive respectively, because the interactions between X and Y are repression (negative) and activation (positive). Looking closer, the condition for a non-zero imaginary part (Appendix Equation (6.1) implies that the degradation rates (timescales) should be similar to minimize the left hand side and the slope of the activation/repression input functions should be steep to maximize the right hand side (corresponding to high Hill coefficients n). With the parameters from Figure 2.3, the discriminant of the characteristic polynomial is in fact negative, leading to oscillations around the fixed point. However the real part of the eigenvalues is given by $\text{Re}(\lambda) = -\mu_X - \mu_Y$ which is always negative, meaning that these oscillations are damped and will always decay to the fixed point over time³.

To get reliable oscillations out of this motif, we can add a third component to the negative feedback loop. To keep the resulting feedback loop negative, the total number of inhibitory interactions in a single feedback loop needs to be odd. Different combinations are possible. First, we will consider a simple extension of equations (2.8) with another component and corresponding positive interaction:

³With noise, this motif can lead to oscillations. [90]

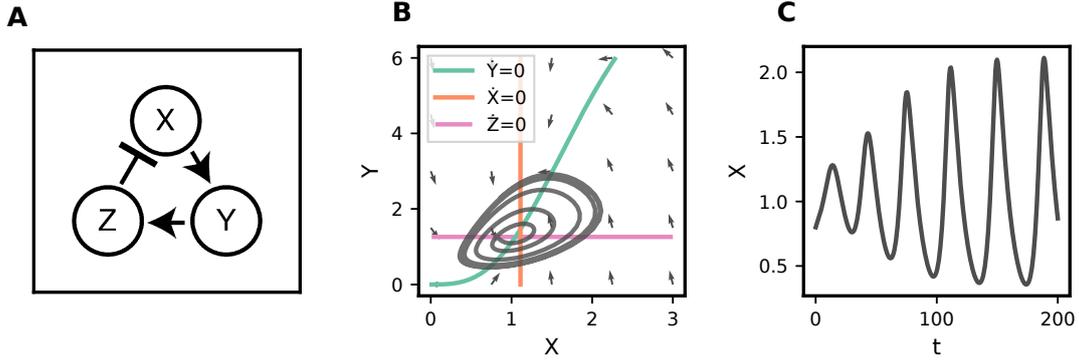


Figure 2.4: (A) Three component negative feedback loop with activators X and Y and repressor Z (Equation 2.9). (B) Slice through the 3D phase space at $Z = 2$. With $\dot{X} = 0$ and $\dot{Y} = 0$ and $\dot{Z} = 0$ nullclines. The intersection of the nullclines (red dot) marks the fixed point of the system. Quivers indicate the direction of change at different points in the phase space, omitting the change in Z direction. The trajectory of a solution from a random starting point, which describes an outwards spiral leading to a limit cycle, is overlaid in black. (C) Time trace of the example solution trajectory. Oscillations have a stable period that is sustained over time and quickly reach a stable amplitude. Parameters: $n_X = n_Y = n_Z = 3, K_X = 1, K_Y = K_Z = 2, \alpha_X = \alpha_Y = \alpha_Z = 1, \mu_X = \mu_Y = \mu_Z = 0.1$.

$$\begin{aligned}
 \dot{X} &= \alpha_X \frac{K_X^{n_X}}{K_X^{n_X} + Z^{n_X}} - \mu_X X \\
 \dot{Y} &= \alpha_Y \frac{X^{n_Y}}{K_Y^{n_Y} + X^{n_Y}} - \mu_Y Y \\
 \dot{Z} &= \alpha_Z \frac{Y^{n_Z}}{K_Z^{n_Z} + Y^{n_Z}} - \mu_Z Z
 \end{aligned} \tag{2.9}$$

The phase space for these systems is now three dimensional, making it harder to visualize. In a two dimensional slice view the X and Y nullclines can still be visualized (Figure 2.9). The X nullcline is vertical in this view, owing to the fact that the rate of change of X does no longer depend on Y . In contrast to the 2 component system, however, a trajectory starting close to the fixed point now spirals out before settling on a limit cycle; oscillating continuously over time with fixed period and amplitude. The delay between a change in one component and the resulting feedback is now sufficient to cause the trajectory to overshoot the fixed point. It should be noted that biological oscillators may consist of more components than might be naively assumed by modelling only proteins: Intermediate protein conformations or mRNA may be considered as separate

components; the same goes for a single component in a different spatial environment connected by active transport processes (e.g. nucleus and cytoplasm). This means that the delay in biological oscillators can be large enough to achieve long periods like 24 hours in circadian clocks [91].

The repressilator is a variant of the three component negative feedback loop where all interactions are inhibitory. Before the function of the repressilator was demonstrated with a synthetic genetic circuit in the ground breaking work by Elowitz and Leibler (2002) [3], theoretical studies had shown that this motif is capable of supporting sustained oscillations [92–94].

Alternative motifs that support oscillations are based on positive feedback and bistability, where hysteresis acts as a sort of memory and destabilizes the spiral trajectory towards the fixed point [89]. Given the motifs above that can result in oscillations, additional requirements constrain the parameters of the interactions. The input function needs sufficient *non-linearity*, the slope around the threshold should be large. Furthermore the *timescales* of the different components need to match. With timescales on different orders of magnitude, the system would follow the nullcline of the fast component, effectively reducing the dimensionality of the phase space.

2.3 Information in Biology

Information is a key concept in many engineering disciplines. The foundational theory of information by Claude Shannon was initially associated with the problem of reliable communication over noisy channels in an electrical engineering context and soon found countless applications in statistics and computer science [95].

But information theory also has relevant applications for biological studies. Optimization of information transmission has been suggested as a fundamental principle for living systems [96] and was examined in a variety of different contexts: In transcription networks [97], population dynamics [98] and neuroscience [99]. Historically, the notion of information (pre-dating the fully developed theory) has inspired the search for a genetic code [100].

In this thesis, circling back to engineering applications, we present the idea that information theory can also be useful to guide the optimization of biological engineering efforts. In this chapter, the basic theory and methods for the calculation of information theoretic quantities are discussed.

2.3.1 Introduction of Shannon Entropy

Information is generally a vague concept with diffuse colloquial interpretation. It was the achievement of Claude Shannon to find a mathematical answer to this ill defined question. To underline this, let us consider a comparison with the usual approach in the natural sciences: While, e.g., the discovery of a formula to describe the trajectory of a ball flying through the air could lean on established concepts such as velocity and its associated quantitative data, there was no data available that the theory of information could be built on. Instead it was primarily necessary to find a quantitative measure for a colloquial idea on which further theoretical structures could be built on afterwards. This section is loosely based on the books by Bialek (2012) [96] and MacKay (2003) [95].

The basic idea is that *information reduces uncertainty*. For example, a student asking for the result of an exam is faced with uncertainty whether he passed or failed. Based on their past experience, they may assign certain probabilities to the possible answers to their question. It seems intuitive that the probabilities of the outcomes are related to the amount of information that is transmitted. For a student that was very confident during the exam and is almost certain about the exam's result, the answer to the question will convey less information than for a student on the edge. When considering the possible answers to a question, are the exact choice of words relevant to the amount of

information transmitted? Let's consider a different example: A musician asks another which composition by Johann Sebastian Bach the other likes the most. The answer could of course be given by whistling a melody or speaking many words to describe the piece, but it is clear that the title of the piece would convey the same information. A very succinct answer would be the position of the piece in a catalogue of all of Bach's pieces (called the 'Bach-Werke-Verzeichnis'). The number of digits required for this catalogue number depends on the size of the catalogue. Looking for the shortest way to *encode* the possible answers to a question is the same task as compressing data to take up less memory space. In fact, Shannon answered the two questions about how much information is gained from an answer and what the potential for optimal data compression is both at the same time by introducing the Shannon entropy, in reference to statistical mechanics.⁴ In statistical mechanics, the entropy measures the uncertainty about the microscopic state of a system that remains after the macroscopic state is known. It is related to the number of the possible microscopic states and their relative probability. Similarly, the Shannon entropy of random variable X with N possible outcomes x_i :

$$H(X) = - \sum_i P(x_i) \log_2 P(x_i). \quad (2.10)$$

The associated probability for outcome x_i is given by $P(x_i)$, with $P(x_i) > 0$ and $\sum_i P(x_i) = 1$. The base for the logarithm is an arbitrary choice, base 2 corresponds to *bits* as unit for information and makes sense in the context of systems that have two different states, like a digital transistor (a single transistor can store 1 bit of information). The entropy $H(X)$ is maximal if all $P(x_i)$ are equal. Going back to our first example this case corresponds to a student who expects to pass or fail their exam with 50 % probability:

$$H = -(0.5 \log_2 0.5 + 0.5 \log_2 0.5) = 1 \text{ bit.}$$

The confident student with a, e.g., 90 % probability of passing receives less entropy:

$$H = -(0.9 \log_2 0.9 + 0.1 \log_2 0.1) \approx 0.47 \text{ bits.}$$

This definition of entropy as a measure of information fulfils a number of sensible

⁴My greatest concern was what to call it. I thought of calling it *information*, but the word was overly used, so I decided to call it *uncertainty*. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, "You should call it *entropy*, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage." ' - Claude Shannon [101]

requirements. If our student extends his question to ask about the result of more than one exam (with the same probabilities), surely the amount of information contained in the answer increases. Thus, $H(X)$ has to be a monotonically increasing function of N , the number of possible outcomes. Furthermore the information of two independent questions has to be additive, if the second question depends on the answer to the first it should be a weighted sum. Shannon proved that his definition of entropy is the only function that can fulfil these requirements [102].

To illustrate how the entropy is equal to the average length of an optimal encoding of a variable we will look at an example given by Shannon in his original publication [102]. Consider a random variable with 4 different values, the letters A, B, C and D with respective probabilities $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ and $\frac{1}{8}$. The entropy is calculated as:

$$H = - \left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{2}{8} \log_2 \frac{1}{8} \right) = \frac{7}{4}.$$

Clearly, the length of a binary number has to be an integer value, not a fraction. Strictly speaking, the entropy tells us we can find a binary code that encodes our variable with *average* length of $\frac{7}{4}$ binary digits. This is achieved by reserving a short code for the most common values, in this case the optimal code can be found as:

A \rightarrow 0
B \rightarrow 10
C \rightarrow 110
D \rightarrow 111

It should be noted that no ‘comma’ or other delimiter symbol is required to make a sequence of values using this code unambiguous (no code word must be the prefix of another code word for this to work).

The notion of clear answers that remove all uncertainty about a random variable X can rarely be applied to the real world, where incomplete evidence and noise are the norm. We get a more fitting description by considering the relationship between two random variables X and Y , where Y describes the data we gather about the state of the world X . If the two variables are not independent, we need to describe the probabilities of the

particular outcomes with joint, conditional and marginal distributions.⁵ In a Bayesian view, our uncertainty about the world is initially given by the prior distribution $P(X)$ and the corresponding entropy $H(X)$. After measuring data y_j , our hopefully improved knowledge about the world is given by the conditional probability distribution $P(X | y_j)$. Thus, our uncertainty about the world is reduced by the entropy of the conditional distribution. Now we need to introduce the different entropies for two random variables, closely derived from the original definition 2.10. The joint entropy of X, Y is defined as:

$$H(X, Y) = - \sum_{i,j} P(x_i, y_j) \log_2 P(x_i, y_j) \quad (2.11)$$

The conditional entropy of X given Y is also the conditional entropy of X given a specific value y_j averaged (weighted) over all possible y_j :

$$H(X | Y) = \sum_j P(y_j) H(X | y_j) = - \sum_{i,j} P(x_i, y_j) \log_2 P(x_i | y_j) \quad (2.12)$$

With this we can introduce the mutual information between X and Y :

$$\begin{aligned} I(X; Y) &= H(X) - H(X | Y) = \sum_{i,j} P(x_i, y_j) \log_2 \frac{P(x_i | y_j)}{P(x_i)} = \\ &= H(Y) - H(Y | X). \end{aligned} \quad (2.13)$$

2.3.2 Interpretations of Mutual Information

Mutual Information describes the reduction in uncertainty about X that can be obtained by observing Y . It can be noted that $I(X; Y) = I(Y; X)$, the Mutual Information is symmetric. The semicolon is used to distinguish $I(X; Y)$ from functions of the joint probability. The sketch in Figure 2.5 shows how the uncertainty about both variables (joint entropy) may be split up into constituents, in a symmetric way. While this sketch follows directly from the definition, it is useful to illustrate how the information from a conditional probability distribution plus the mutual information between the two variables adds up to the information contained in the marginal probability distribution. Note that, in relation to the above introduction where Y is collected data and X the

⁵

joint: $P(x_i, y_j)$, marginal: $P(x_i) = \sum_j P(x_i, y_j)$, conditional: $P(x_i | y_j) = \frac{P(x_i, y_j)}{P(y_j)}$

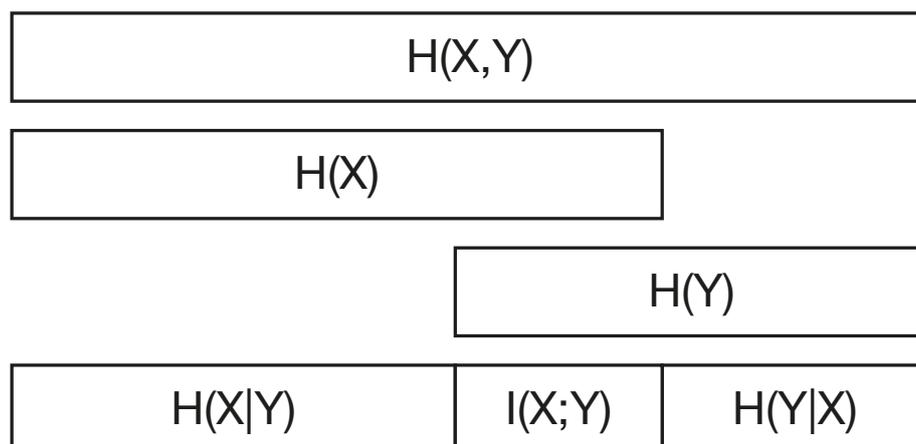


Figure 2.5: *The relationship between joint entropy (top row), marginal entropies (middle rows) and conditional entropies and mutual information (bottom row). The symmetric mutual information can be calculated in different ways by breaking down the joint, marginal or conditional entropies. Note, that the entropies drawn as blocks are not meant to represent sets as in a Venn diagram [95].*

unknown world, the latter version of equation (2.13) is the practical one. $H(Y)$ is the entropy of the distribution of our data (over the full dynamic range) and $H(Y | X)$ can be interpreted as the average noise of the data at all different points of X ; both quantities can be calculated from the experimental data.

When using mutual information to describe the interdependency of two random variables (not directly concerned with the underlying probability distributions), it may be compared to other, similar measures. The coefficient of correlation is a more common measure for the same purpose. As illustrated by examples in Figure 2.6, the coefficient of correlation only measures linear dependencies while mutual information is also suitable to describe non-linear dependencies.

How can a mutual information value in bits be interpreted? The example in Figure 2.6 A is a simple linear relationship with added noise. The calculated mutual information value of 2.4 bits gives the average length of the optimal encoding of the information about X gained from the knowledge of Y . Conversely, this information may be decoded in arbitrary ways. If the goal is to differentiate distinct intervals of X , the mutual information tells us that by observing Y we can distinguish no more than $2^I = 2^{2.4} \approx 5.3$ states due to the noise level.

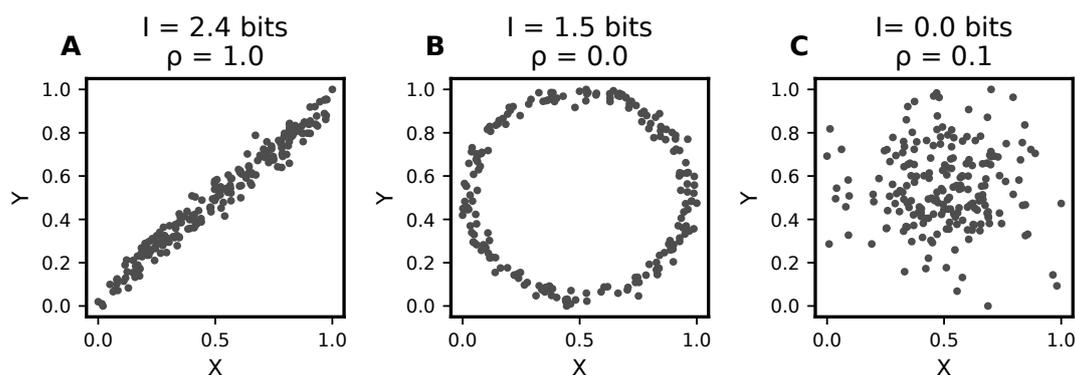


Figure 2.6: Samples of three different, joint probability distributions for random variables X and Y . For each set of samples, the mutual information and the Pearson correlation coefficient is calculated. To calculate mutual information, the underlying probability distribution is estimated from the samples with the method described in Figure 2.8. Unlike the linear interdependence (**A**), the non-linear interdependence between X and Y (**B**) can be quantified with mutual information but not with a correlation coefficient. Independent probability distributions for X and Y (**C**) leads to (near) zero values.

2.3.3 Measuring Information

Calculating information theoretic quantities (which depend on probability distributions) for experimental data brings a set of technical challenges. So far, in this chapter we introduced entropy definitions for probability distributions over discrete variables. In nature, quantities are usually continuous. All definitions can be readily adapted for continuous variables by swapping out the discrete sum with an integral. If the underlying probability distribution for a continuous variable however is not unknown, it needs to be estimated from experimental data. In experiments, we sample the continuous real world with discrete digital instruments. The dynamical range of measurement instruments usually dwarfs the number of samples N by orders of magnitude, which means that the discrete probability distribution over the space of possible measurements is very sparsely sampled. To estimate the underlying distribution of the variable, data points are aggregated into bins and their frequency is used for the estimate of the probability. Figure 2.7 shows an example of this process for a joint probability distribution of two variables.

The limited number of samples and the subsequent limited number of bins introduce random errors to the estimate of the probability distribution, which in turn leads to systematic errors for derived information theoretic quantities [99]. Different practical methods have been introduced to deal with these systematic errors (with prerequi-

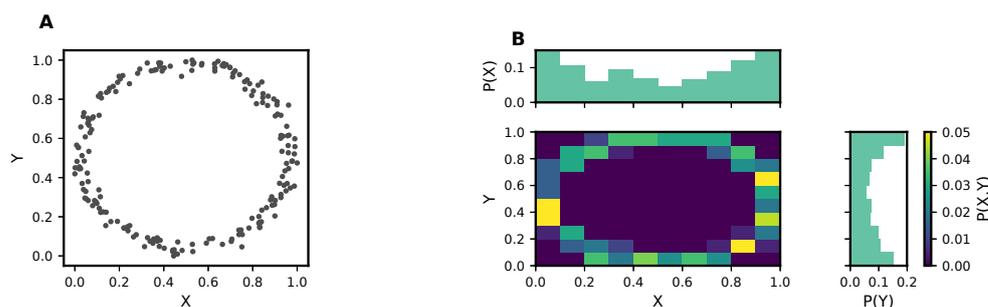


Figure 2.7: Sampling and density estimation for continuous variables. (A) Artificial example for a continuous joint probability distribution for two random variables X and Y sampled $N = 200$ times. (B) Estimation of the underlying probability distribution from the frequency of samples in 10×10 bins. Along the axes, the marginal histograms for each variable are shown.

sites) [103, 104]. The procedure shown in Figure 2.8 explains the approach to directly estimate the underlying distribution from the data and correct for systematic errors with a two step extrapolation procedure. Alternatively, if the data is not sufficient to estimate the underlying probability density, a Gaussian distribution based on the mean and variance of the data may be used instead. A Gaussian distribution can be shown to maximize the entropy for a given mean and variance [105], making an entropy value calculated using this approximation an upper bound to the real value. This approach is an example of the maximum entropy principle [106], which states that a distribution that makes the minimal amount of assumptions is the least biased possible estimate.

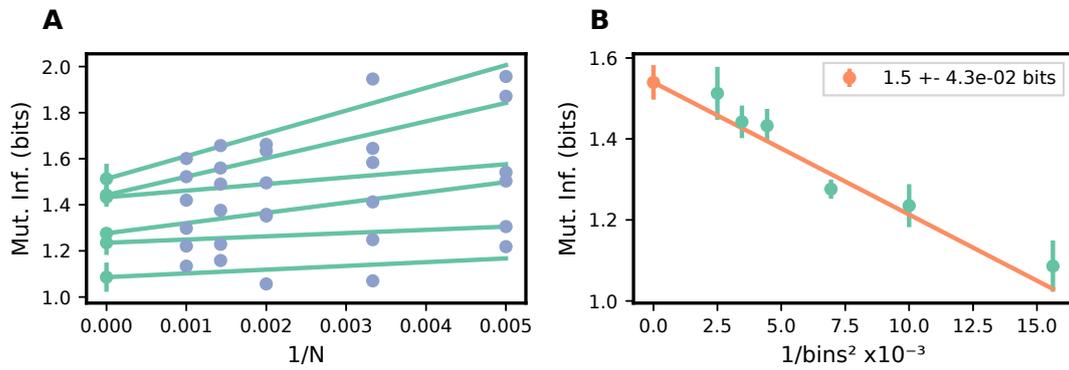


Figure 2.8: Extrapolation procedure for mutual information to deal with limited data. Data based on Figure 2.7. **(A)** Mutual information has a linear relation with the inverse of the number of samples. A value for an infinite number of samples is estimated by extrapolating from the available N samples. Data points for less than N samples are created by drawing randomly from the available samples. This procedure is repeated for a different number of bins. **(B)** Here, the mutual information has a linear relation with the inverse square of the number of samples. The estimates from **(A)** are used to extrapolate to a mutual information value for an infinite number of bins and infinite number of samples.

3 Results

3.1 A Synthetic Bacterial Clock with a Hybrid Feedback Loop Containing dCas9-sgRNA

The following section is based on a peer reviewed paper by the authors John Henningsen, Matthaeus Schwarz-Schilling, Andreas Leibl, Joaquin Gutiérrez, Sandra Sagredo and Friedrich C. Simmel [107].

Over the past two decades, researchers in synthetic biology have created a wide variety of synthetic gene circuits with designed dynamic, sensory or computational functions [4, 108]. Among these, synthetic genetic oscillators are particularly interesting, as they do not only require control over gene expression levels, but also over expression dynamics. In their seminal work, Elowitz and Leibler [3] created a genetic oscillator termed ‘repressilator’, whose negative feedback loop was composed of three genetic repressors (LacI, TetR, and λ -cI), which cyclically repressed each other’s expression. Later, a variety of other oscillator architectures were demonstrated, e.g., two-node oscillators containing an activating and a repressing link [109, 110], a mammalian oscillator with negative feedback based on antisense RNA transcription [111], and even a bacterial five-node oscillator containing five cyclically arranged repressor modules [112]. More recently, the dynamical properties of the original repressilator were analyzed in depth, leading to an improved circuit design, which resulted in long-term synchronized oscillations with reduced noise [65].

Oscillatory biochemical reactions naturally occur in a variety of biological contexts [89, 113, 114], e.g., in the control of bacterial cell division [115], the coordination of the eukaryotic cell cycle [116, 117], or in circadian rhythms [118]. Similarly, synthetic oscillator circuits could serve as pacemakers or timers for engineered processes. In fact, several potential applications of synthetic oscillators have been recently demonstrated, for instance, to synchronize lysis of therapeutic bacteria for *in vivo* delivery applications [119], or to probe bacterial growth dynamics in the gut [71].

Generally speaking, oscillatory dynamics are generated by non-linear negative feedback loops with a delay [89], where the delay may be achieved, e.g., by distributing the feedback over several intermediate nodes. Non-linearities are required to destabilize a system during its approach towards steady state, and shorter feedback loops have been shown to require stronger non-linearities than longer ones [17, 89]. Apart from non-linearities generated by cooperative binding, also thresholding by molecular titration can be employed [89], typically resulting in relaxation oscillations.

As an exciting alternative to gene regulation via transcription factors, recently the CRISPR interference (CRISPRi) technique has been established, which allows sequence-

programmable gene silencing based on the catalytically inactive mutant dCas9 of the CRISPR associated protein Cas9 [9, 120]. Until now, CRISPRi has only rarely been used in the context of dynamical gene circuits, however. Notable exceptions are two recent examples of oscillating networks, in which three single guide RNAs (sgRNAs) cyclically repressed each other's transcription in the presence of dCas9 [121, 122] or dCas12a [122]. In the present work, we took a hybrid approach - which we termed 'Rock Paper CRISPR' (RPC) oscillator-, in which only one of the transcriptional repressor links of a repressilator network was replaced by CRISPRi.

Long-term monitoring of more than 100 individual bacteria using a microfluidic 'mother machine' [33] allowed us to analyse the dynamic variability and stability of the CRISPRi-modified oscillator at the single cell level. Even though CRISPRi is not expected to provide highly non-linear repression, we demonstrate, both experimentally and via computational modelling, that our system is still capable of sustained temporal oscillations.

The integration of a CRISPRi node into the repressilator opens up the possibility to control arbitrary genomic functions of the bacteria by providing additional sgRNA molecules that target the corresponding sequences on the genome [10]. However, there may also be challenges associated with CRISPRi such as changes in the morphology and growth rate of the cells due to unspecific binding of dCas9 to genes on the genome [123, 124]. We therefore also investigated aspects such as de-synchronization and phase stability of the oscillator, and its relation to bacterial growth. Using an information-theoretic approach, we analyse its potential as a timekeeper or as a controller for temporal gene expression programs.

3.1.1 Circuit Design

The genetic circuit is based on the original repressilator, which contains the TetR, λ -phage cI and LacI transcriptional repressors in a negative feedback cycle. We replaced LacI with RNA-guided dCas9 and a suitable sgRNA that binds to the nontemplate strand of the *tetR* gene at about 190 base-pairs downstream of the promoter (Figure 3.1). Thus, the dCas9-sgRNA complex represses TetR and fulfills the same functional role as LacI in the original circuit. We replaced LacI as this is the only endogenous protein from *E. coli* and might interfere with the circuit (previous repressilators were operated in the *lacI*-deficient strain MC4100).

We tested different variations of this basic circuit. To enable operation of the oscillator in bacterial strains that contain LacI, we continued by exchanging the P_{LacO1} promoter controlling the *tetR* gene with different constitutive promoters. However, this led to a

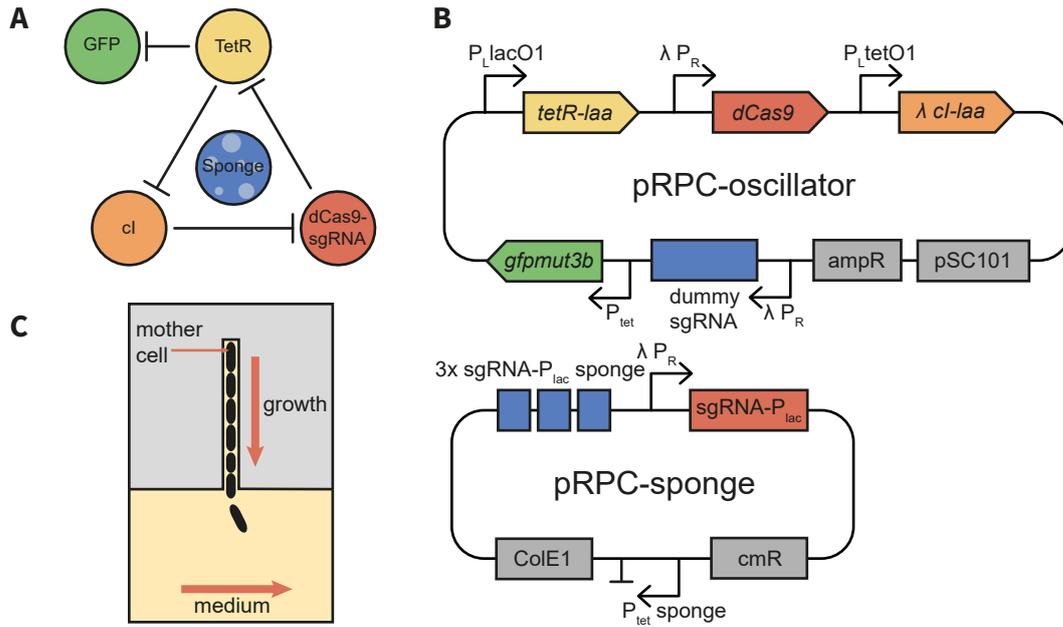


Figure 3.1: Design and microfluidic experiments. (A) Schematic representation of the repressilator circuit. A negative feedback loop of three repressors, including dCas9-sgRNA, generates oscillations, sponge elements stabilize the circuit dynamics and GFP is used as a fluorescent readout. (B) Schematic plasmid maps showing the realization of the genetic circuit on two plasmids with different copy numbers (pRPC-oscillator: ≈ 5 (pSC101), pRPC-sponge: $\approx 15 - 20$ (ColE1)). (C) Sketch of the mother machine microfluidic device. *E. coli* cells containing the plasmids are trapped in a small growth channel, where they are supplied with growth medium through a larger feed channel, allowing long-term observations of single cells in steady state growth conditions.

deterioration or even disappearance of the oscillations (Figure 3.2 C& D). We therefore kept the P_{LlacO1} promoter and either operated the oscillator in MC4100 (which lacks *lacI*, Figure 3.2 G) or in MG1655, in which case we supplemented the growth media with IPTG to alleviate the repression by endogenous LacI. We also did not observe any oscillations in a strain that lacked the protein degradation machinery. Furthermore, simulations of the genetic circuit suggested that the absence of degradation tags would significantly increase the period of the oscillator. We thus left the degradation tags on λ -cI and TetR and used strains with an active degradation machinery.

Following the approach by Potvin-Trottier et al. [65], we introduced ‘sponge elements’ for the repressors, which were previously found to stabilize the dynamics against noise occurring at low copy numbers. To this end, we employed decoy binding sites for TetR and the dCas9-sgRNA complex on a secondary sponge plasmid (with a higher copy number than the oscillator plasmid). In order to test the capacity of the oscillator to

control the expression of another gene as a secondary process, we expressed an additional ‘dummy’ sgRNA from a λ -pR promoter on the oscillator plasmid, which does not have a corresponding target sequence on the chromosome or on the plasmids. The ‘dummy’ sgRNA thus simulates the regulation of an external secondary target. Furthermore, the λ -promoter acted as an additional sponge for λ -cI, while the dummy sgRNA competed with the formation of functional dCas9-sgRNA complexes (for experiments without sponge or dummy sgRNA see Figure 3.2 C–E). We placed the dummy sgRNA on the oscillator plasmid so that its copy number is smaller than that of the sgRNA which is involved in the oscillator and encoded on the sponge plasmid. The state of the oscillator can be tracked by a fluorescent protein (GFPmut3b), integrated into the oscillator plasmid, that is in phase with the λ -cI repressor.

3.1.2 Microfluidics Experiments

The oscillator circuit was analyzed on a single cell level in a suitable microfluidic device termed the ‘mother machine’ (Figure 3.1 C). Since the oscillator is not synchronized across the bacteria population, the oscillation of the fluorescence intensity of individual cells averages out in bulk. Visual proof of oscillation may be obtained from the growth of bacteria in larger microfluidic chambers, but for the quantitative study of the oscillator circuit a continuous observation of single cells over time is necessary, either by trapping or tracking across timelapse frames. In the mother machine individual bacteria cells are trapped in $1\ \mu\text{m} \times 1\ \mu\text{m}$ cross section growth channels. The cell at the end of the growth channel, the ‘mother cell’, is held in place and can be continuously observed while cell elongation and division causes ‘daughter cells’ to be pushed out of the growth channel and be washed away. Furthermore, the cells should be in a steady state with regards to the growth conditions. While such a state may be approximately assumed during the exponential growth phase in batch culture, this approach is limited to short time periods and only suitable for the computation of static logic functions, not dynamic circuits. In the ‘mother machine’ microfluidic chip a steady nutrient supply pumped through a larger feed channel ($20\ \mu\text{m} \times 50\ \mu\text{m}$ cross section) perpendicular to the growth channels enables steady state growth over time periods in the order days.

The data in Figure 3.2 shows a typical example of a single growth channel filled with *E. coli* cells containing the two oscillator plasmids. A kymograph (Figure 3.2 B), displaying the fluorescence intensity of the selected growth channel over time, visualizes the periodic, oscillator-controlled GFP expression in the cells. The apparent synchronization of the cells within a growth channel is caused by their close relationship, as they descend from

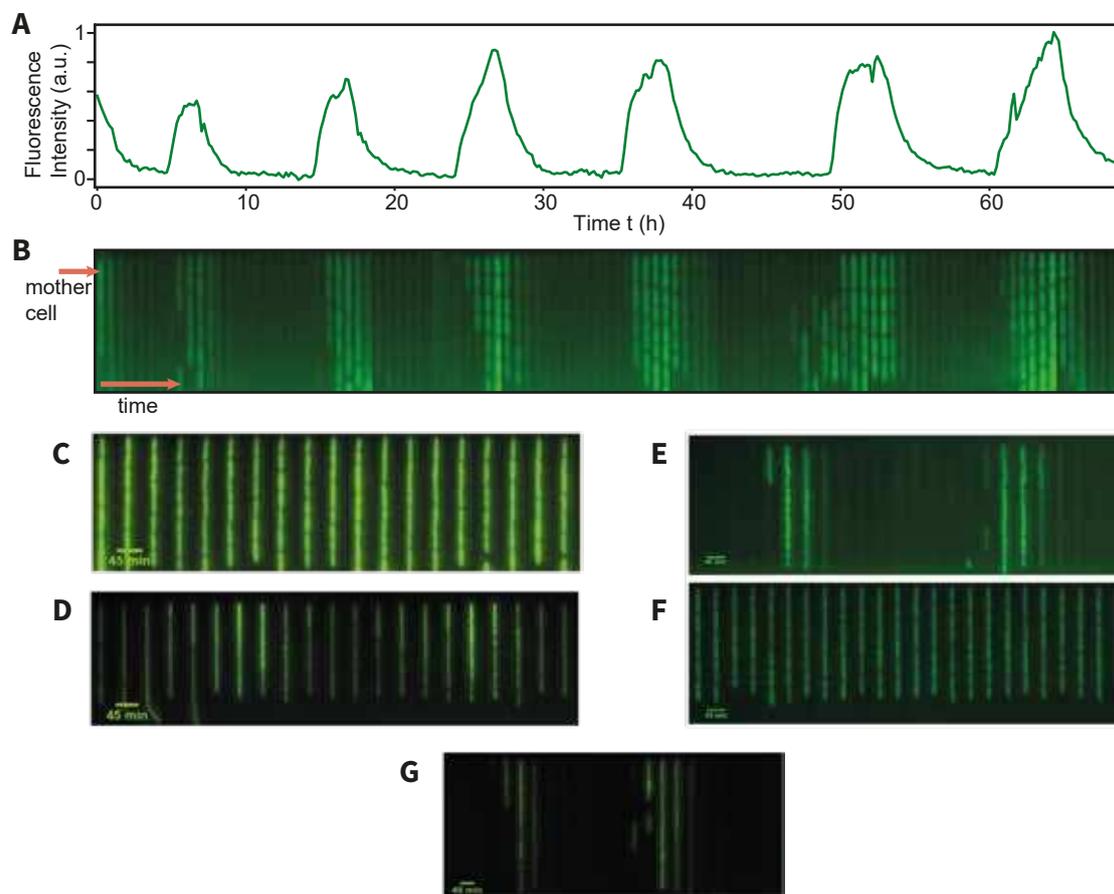


Figure 3.2: GFP fluorescence data. (A) The fluorescence intensity (F.I.) of the mother cell (shown in B) over time is extracted from the microscopy videos by detecting the mother cell area and taking the average of the fluorescence intensity of that area (B) Kymograph (a single vertical growth channel over time) of cells displaying stable oscillations over almost 70 h. All following kymographs are examples of variations from the final RPCi construct (Figure 3.1 B) that were tested: (C) P_{LlacO1} promoter replaced with the constitutive promoter J23101 and no dummy sgRNA, no oscillations observed; (D) P_{LlacO1} promoter replaced with the constitutive promoter J23106 and no dummy sgRNA, irregular oscillations observed; (E) no dummy sgRNA in *E. coli* strain MC4100 (no LacI) instead of MG1655, stable oscillations observed; (F) negative control without any sgRNA on the oscillator plasmid and without sponge plasmid in MC4100, no oscillations observed; (G) RPCi in MC4100, stable oscillations observed.

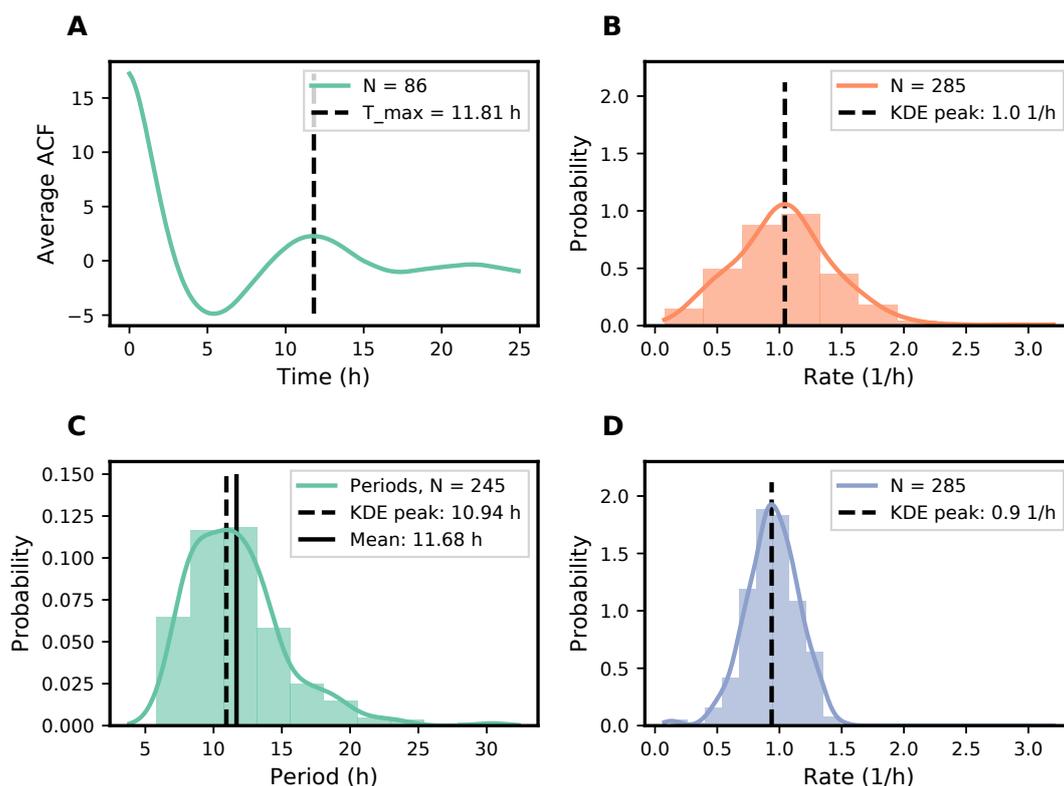


Figure 3.3: Analysis of single cell data. 86 channels measured in parallel over 50 h are analyzed and compared. (A) Autocorrelation function of the fluorescence intensity of single cells averaged over all mother cells. The first peak suggests that the time traces are periodic with a period of roughly 12 h. (C) Distribution of periods (peak to peak distance) in all time traces (histogram with kernel density estimation (KDE)) shows the variation of the oscillator frequency. (B) F.I. peaks decay exponentially. The distribution of the rate (base e) over a fixed time interval is shown in on top. (D) Comparison to the similar distribution of growth rates during the same interval suggests that dilution is responsible for protein decay.

the same mother cell and are thus expected to have similar cellular composition. The vertical position of the cells in the channel changes over time due to mechanical forces arising from the growth of the individual cells, which conversely also limits their growth rate [125]. For this growth channel, the extracted quantitative fluorescence intensity data for the mother cell (Figure 3.2 A, see following section for method) shows a total of six regular periods are visible over 69 hours of experiment.

3.1.3 Single Cell Data Analysis

The observation of many mother cells in parallel (up to 86 in one experiment) together with statistical analysis can give insight into the circuit and host dynamics that would be obscured in simpler experiments. For this purpose, the microscopy timelapse, with bright field (phase contrast) (BF) and fluorescence intensity (FI) channels, was processed to yield time traces describing the fluorescence intensity $I_i(t)$ and cell area for individual mother cells over time. To obtain the time traces, the bright field channel was used to discard outliers and to classify pixels into cell and background categories. The analysis was cut off after 50 hours of experiment time since after two or three days many cells stopped dividing. Mother cells that stopped regular cell division earlier were excluded as outliers [33].

In order to be able to better investigate the periodic portion of the noisy signal, we calculated the autocorrelation function (ACF) $A_i(\tau) = \langle I_i(t + \tau)I_i(t) \rangle_t$ for each mother cell. The ACF of the fluorescence intensity averaged over our set of single cell time traces, (i.e., $\langle A \rangle(\tau) = 1/N \sum_i A_i(\tau)$) (Figure 3.3 A) shows that the analyzed cells oscillate on average. The first peak of the average ACF indicates an average oscillator period of 11.7 ± 0.4 hours (standard error of the mean, bootstrapped (SEM)). A histogram of individual oscillation periods (calculated from the peak-to-peak distances T_j for all full periods in $\{I_i(t)\}$) also shows a mean of 11.7 h and a standard deviation of 3.5 h (Figure 3.3 C). The observed distribution has positive skewness (sample skewness $g_1 = 1.3$), corresponding to the observation that most cells oscillate with a period in a relatively narrow interval around the mean (75% between 8h and 15h). Only a few oscillators were observed that ran very slowly, but no fast oscillations were recorded. Out of all mother cells that showed normal growth, only 5 out of 86 cells did not oscillate while for the remaining 81 cells 2 periods or more were registered. As shown in Figure 3.4, within the same cell a given period T_j does not correlate with the previous period T_{j-1} , indicating that the variance of the distribution is not exclusively explained by cell-to-cell variations, but also by fluctuations inherent to the synthetic circuit. We did not observe an overall slowing down of the oscillations, however, which conceivably could occur due to ageing of the mother cell or other deterioration effects.

The asymmetric shape of the peaks in the oscillator time traces suggests that our circuit behaves as a relaxation oscillator, which is consistent with previous findings for the repressilator v2.0 [65]. In a relaxation oscillator, a build-up phase of active protein production is followed by a decay phase, in which protein production has stopped and

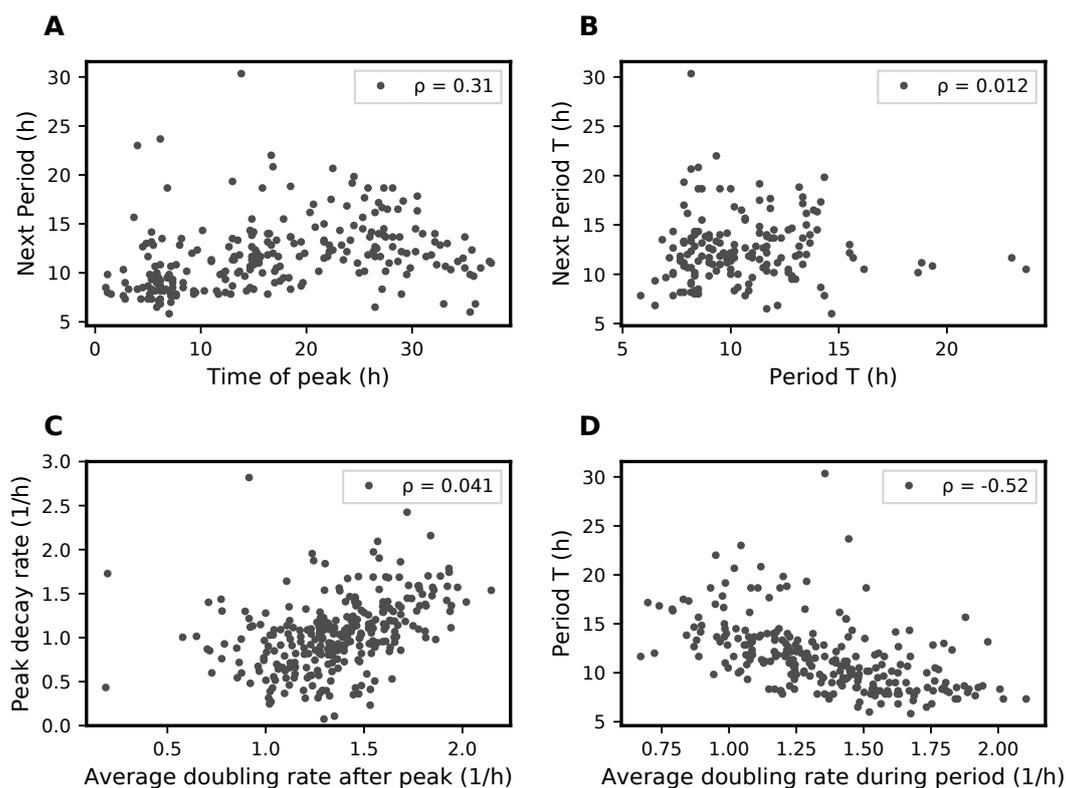


Figure 3.4: Scatter plots showing the relationships between different quantities calculated from the single cell data (dataset 1). Pearson correlation coefficient ρ is given in the legend. (A) Time of a F.I. peak against the duration of the following period. A slowdown of the oscillator is not observed. (B) Length of a period against the length of the following period in the same mother cell. (C) Average doubling rate after a peak (1h - 4.5h) against the peak decay rate in the same interval. (D) Average doubling rate during an entire period against the length of the period.

proteins are diluted or degraded. We find that the fluorescence time traces in the decay phases 1h - 4.5h after each peak are indeed closely matched by exponential fits (with a mean coefficient of determination $\langle R^2 \rangle = 0.94$), as would be expected for relaxation oscillations.

In our oscillator circuit, the two repressor proteins are equipped with degradation tags, while the dCas9-sgRNA complexes are not actively degraded. We therefore expect dilution of dCas9-sgRNA (and potentially removal of the tightly bound complexes from their target sites by the replisome [126]) to be the slowest steps in the feedback circuit. Note that removal of dCas9-sgRNA would activate TetR, which in turn would suppress GFP production.

To elucidate the relation between the observed decay phase and dilution due to bacterial

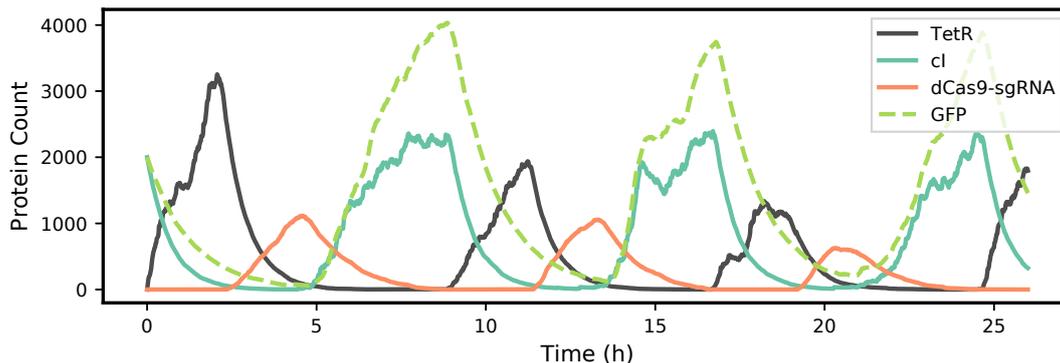


Figure 3.5: Stochastic simulations match experimental data and reveal parameter constraints of oscillatory regime.

growth, we calculated an instantaneous growth rate $G_i(t)$ of the mother cell from division events observed in the microfluidic device. From $G_i(t)$ we then obtained average growth rates for each decay period (i.e., 1h - 4.5h after each peak), and compared them to the observed fluorescence decay rates. Figure 3.3 B&D shows histograms for both of these rates. The peaks decay with most likely rate $\lambda_{dec} = 1.0 \pm 0.03\text{h}^{-1}$ (SEM) while the cells grow with most likely rate of $\lambda_{growth} = 0.9 \pm 0.03\text{h}^{-1}$ (SEM). The decay of the fluorescence signal thus clearly occurs on the same time-scale as cell growth, but with a larger variability observed for the oscillator decay ($\text{CV} = 40.4\%$) than for the growth rate ($\text{CV} = 22.3\%$).

3.1.4 Simulation

In order to check our assumptions about the dynamics of the RPC oscillator, we performed computer simulations using a stochastic model of the system. Using physically plausible parameters [127], the model results in time traces that closely recapitulate the experimental data (Figure 3.5). In particular the peculiar form of the oscillations and the time scale are matched very well. The model demonstrates that GFP production is out of phase with the presence of dCas9-sgRNA and roughly in phase with the production of λ -cI, as would be naively expected from the circuit topology. Accordingly, the period between peaks in GFP expression is determined by the decay phase of TetR, while the build-up phase of GFP is mainly determined by the dilution phase of dCas9-sgRNA.

It has been shown that dCas9-sgRNA binds to its target with very low off-rates [128, 129], and we therefore investigated the effect of such quasi-irreversible binding within

our model. Our simulations suggest that irreversible binding of dCas9-sgRNA prolongs the dilution phase of dCas9-sgRNA, which increases the period of the oscillations. The dilution of dCas9-sgRNA in turn is coupled to the replication and dilution dynamics of the circuit plasmids.

The model is thus consistent with our assumption that removal of the dCas9-sgRNA complex is a rate-determining step of the oscillator. Furthermore, our model confirms that strong cooperativity is not required for the oscillations, but titration against the decoy binding sites on the sponge is important.

3.1.5 Synchronization and Phase Stability

After experimentally synchronizing a set of oscillating cells, we recorded how their subsequent loss of synchrony in order to understand their collective phase stability. From this dataset we can also use methods from information theory to draw theoretical conclusions about the viability of our circuit for timekeeping applications. While our oscillator circuit has no internal mode of synchronization, it can be synchronized with the addition of external inducer chemicals to the growth medium. The addition of aTc to the medium inhibited the function of TetR and the omission of IPTG meant that endogenous LacI could be active. Thus the oscillator circuit was stopped while growing in batch culture. After subsequent loading into the mother machine the cells were supplied with normal medium conducive to oscillations and the experiment conducted as before. As the single cell oscillators are locked into the same static state at the beginning of the experiment, the oscillators are expected to start oscillating in phase and go out of phase progressively due to stochastic effects.

Figure 3.6 B shows the fluorescence intensity time traces of all cells used in the analysis. At the beginning of the traces, the cells appear to be in the decay phase after a (first) GFP maximum (this initial lag is explained by the time delay caused by transferring bacteria from culture to the mother machine and the start of timelapse acquisition (≈ 150 min)). To quantify dephasing among the oscillators, an instantaneous phase $\phi_i(t)$ was calculated for every time point of every trace $I_i(t)$ by applying a Hilbert transform to the data [130]. From the distribution of the phase angles, we then obtained their mean and standard deviation at every time point (Figure 3.6 C). The mean increases approximately linearly, with a slope corresponding to an average period length of 12.0 h. The standard deviation of the phase distribution steadily increases over time, and reaches half a period after 40 h, indicating a complete desynchronization of the observed set of oscillators in this time period.

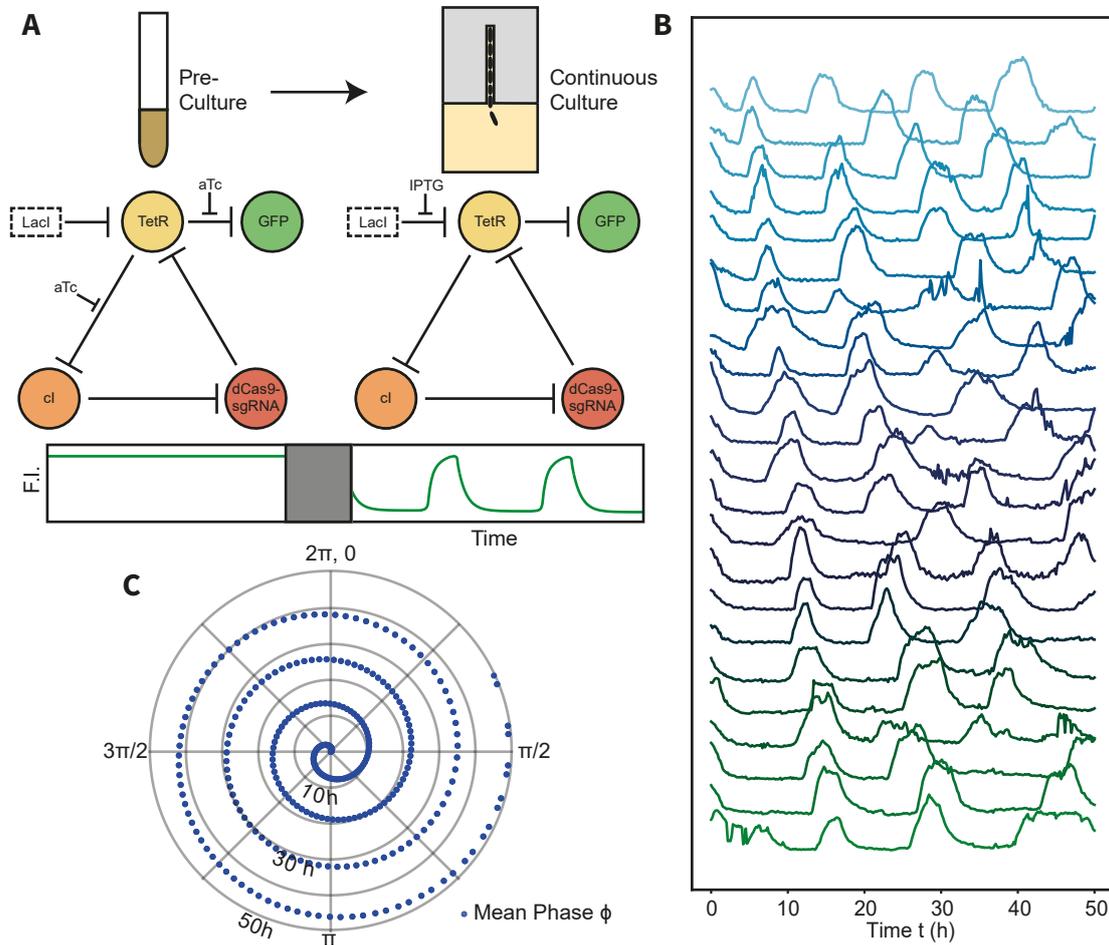


Figure 3.6: Dephasing behavior of the oscillator. **(A)** Cells are synchronized by fixing the phase in the bulk preculture before the start of the microfluidic experiment. *aTc* is added to inhibit *TetR* and no *IPTG* is added to inhibit background of endogenous *LacI*, thus fixing the circuit in a state with high *cl* repressor and *GFP* concentration. After a time delay for experimental setup, the inducers are reversed in the microfluidic experiment the inducers are reversed and the cells begins oscillating. **(B)** Summary of single cell fluorescence intensity traces after phase synchronization. All traces start in phase at the tail end of a peak and dephase over time. **(C)** The instantaneous phase of the time traces is quantified with the Hilbert transform. The plot shows the average phase of all time traces on a circle with a radial time axis.

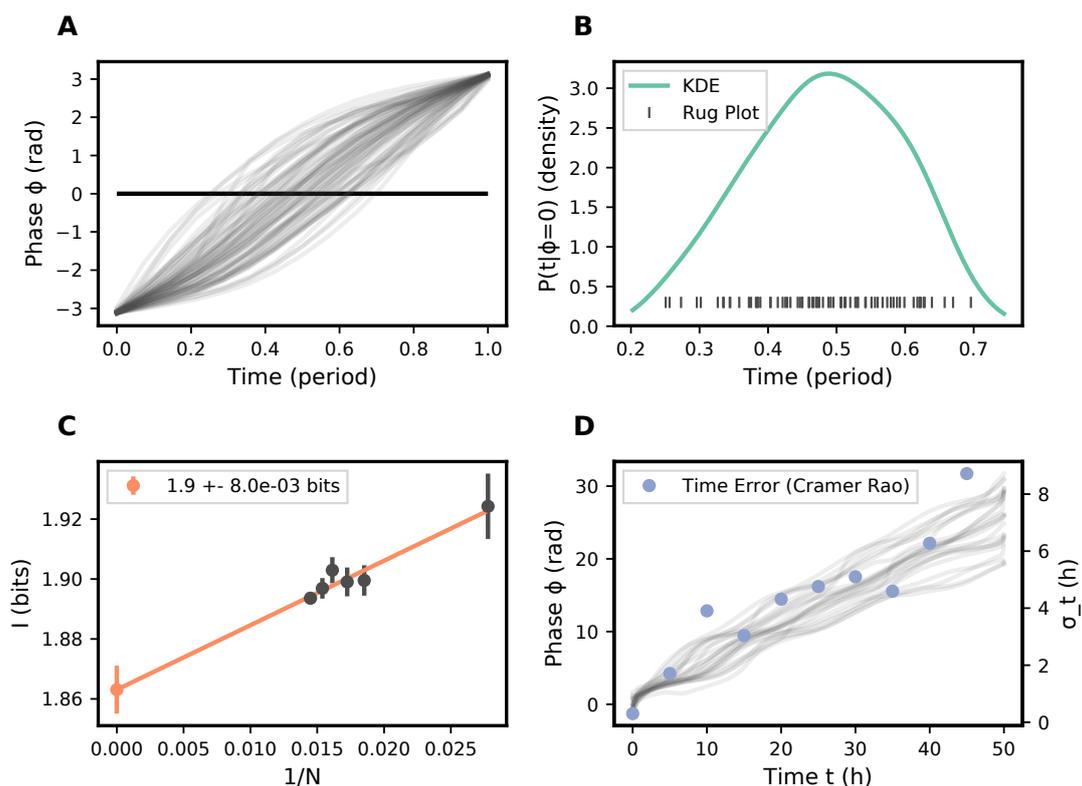


Figure 3.7: (A) Variance of the instant. phase during a period is shown for all periods in the data set, with time axis scaled by respective period length. (B) Width of the conditional probability $P(t|\phi = 0)$ (KDE and rug plot) determines how well different time states can be distinguished after measuring, e.g., $\phi = 0$ (w.l.o.g.). (C) Extrapolation of the value for the mutual information to an infinite number of samples. The mutual information is calculated for subsets of the original data of different sizes (random resampling without replacement). Each size of subset is resampled 100 times to calculate mean and standard error. The y-intercept of a linear fit corresponds to an extrapolated value for an infinite number of samples. (D) Estimation of elapsed time after synchronization reduces in accuracy with loss of synchrony. Set of instant phase traces over the entire duration of the experiment shown in blue. Lower bound (Cramer-Rao) for the error in the time estimation at different time points shown in red.

3.1.6 Temporal Information and Reliability as Clock

In a further step, we want to consider potential applications of our circuit. Since sgRNAs can be designed for multiple targets in parallel, the CRISPR node of the oscillator system could be used to control secondary processes. We consider two subtly different concepts: Control of the order of execution of downstream processes or as literal clock keeping track of time delays. The former could be compared to a clock signal controlling a synchronous electronic circuit or the conductor of an orchestra. The latter view is more compatible

with circadian oscillators found in nature or simply an alarm clock.

For the first concept, a consistent gene expression pattern of one period is required for the robust differentiation of distinct states. To this end, we analyzed all full periods in the phase trace dataset and quantified the variability. Variations in period length were equalized, since for now we are only concerned with the order of events and not their exact timing. It has to be noted that this approach implicitly assumes that the bacteria ‘know’ when an oscillation period starts or ends (corresponding to a peak in protein expression), therefore have no uncertainty about the relative timing of these points. To understand how to quantify how well we can differentiate distinct states of the oscillation, we can think about the *information* the phase signal confers about time. Let us consider the information gained by a single phase measurement. Before the measurement of the phase, the time within one period is completely unknown (corresponding to a uniform prior distribution $P(t)$). After measuring a specific phase signal ϕ^* , the probability to be at a specific time point is then given by the conditional probability (i.e., posterior distribution) $P(t|\phi^*)$ (Figure 3.7 B). The narrower this distribution, the more uncertainty about the point in time is removed and hence more information is gained. This intuitive idea can be formalized using the concept of mutual information $I(t; \phi)$, which describes the amount of information about one variable (in bits) that is obtained from observing the other. From our experiments, we computed a value of $1.9 \pm (5.8 \cdot 10^{-3})$ bits (Figure 3.7 C), which indicates that theoretically a maximum of $2^{1.9} \approx 3.7$ distinct temporal states per period could be reliably distinguished from the output of a single node in an individual instantiation of the oscillator circuit.

Secondly we want to quantify how precisely one could tell the absolute time passed since the initial synchronization of the bacterial oscillators by observing a single instance. Experimentally we can observe how the variability in period length leads to a distribution of phases that spreads out over time. Implicit in this representation of phases is the requirement for a counter (otherwise the accumulated phase $\phi(t)$ is not accessible, but only $\phi(t) \bmod 2\pi$). A counter system has been built before but is not in reach for our system. When the absolute time is estimated from the measurement of the phase, the variability in the phase signal, quantified by its standard deviation $\sigma_\phi(t)$ leads to uncertainty of the time estimate $\sigma_t(t)$. A lower bound for $\sigma_t(t)$ is given by the Cramér-Rao bound $\sigma_t(t) \geq 1/\sqrt{\mathcal{I}(t)}$, where $\mathcal{I}(t)$ is the Fisher information calculated from the phase angle distribution. As the oscillators go out of phase, the uncertainty about the absolute point in time increases (Figure 3.7 D), i.e., the accuracy of the clock decreases. While the absolute error increases almost four-fold over the course of the experiment (compare

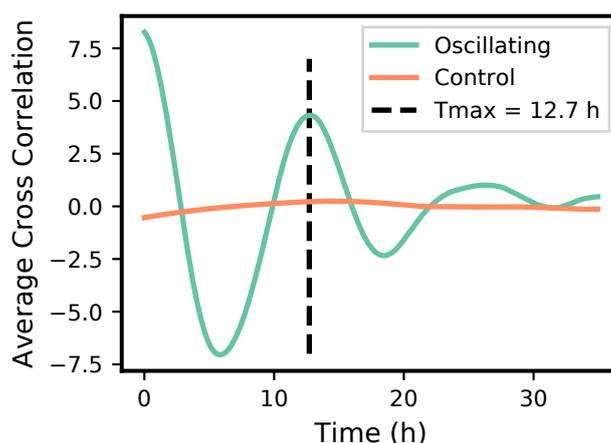


Figure 3.8: Cross-correlation of fluorescence intensity $I(t)$ and instantaneous growth rate $G(t)$ averaged over the set of mother cells shows that the growth rate oscillates in phase with the fluorescence intensity. A control experiment with strongly suppressed oscillations (aTc added, no IPTG, see Figure 3.6) shows essentially no correlation between fluorescence intensity and instantaneous growth rate.

$\sigma_t(t = 5h) \geq 1.7h$ with $\sigma_t(t = 40h) \geq 6.5h$, the error relative to the runtime of the experiment actually goes down.

The above considerations should be interpreted as theoretical estimations that can serve as a point of reference or perhaps upper bound for the capabilities of the bacterial oscillator to act as a timekeeping module. The computations in this chapter are somewhat artificially based on the phase data extracted from the fluorescence signal generated by the oscillator. In order to actually ‘read out’ the oscillator phase, the bacteria would have to make use of instantaneous protein levels and potentially their rates of change. On the other hand, the temporal information from several of the oscillator nodes (rather than from just one) might improve the ability to determine a point in time within an oscillation. In general, information theoretic computations as shown here would have to be adapted to the molecular details of different potential timekeeping applications.

3.1.7 Interplay with Bacterial Growth

Gene expression activity is intimately linked to the bacterial growth rate [68, 69], and it can therefore be expected that genetic oscillations will be correlated with periodic changes in cell growth. Indeed, our single cell analysis revealed that the instantaneous growth rate $G_i(t)$ of each mother cell oscillates in phase with and with the same frequency

as the fluorescence intensity. Strong periodic fluctuations in the cell growth are already visible in the single cell growth data. To deal with the relatively noisy growth data, we computed the cross-correlation $C_i(\tau) = \langle G_i(t + \tau)I_i(I(t)) \rangle_t$ between the growth and fluorescence traces for each single cell and then averaged over the whole set of cells. $C_i(\tau)$ measures the similarity of the two signals as a function of the relative time shift τ between them. Ideally, the cross-correlation of two synchronous signals peaks at $\tau = 0$ and at the common period $\tau = T$. Indeed, the ACF of the fluorescence intensity for our synchronized cells oscillates with a period of 12.4 h and the average cross-correlation data in Figure 3.8 peaks at $t = 0$ h and $t = 12.7$ h, which demonstrates that the instantaneous growth rate oscillates in phase with the fluorescence intensity.

We point out that bacterial cell division occurs on a different time-scale than these oscillations - the average instantaneous growth rate in the experiments was $1.13 \pm 0.44 \text{ h}^{-1}$, corresponding to a doubling time of 53 min with a $CV = 39\%$. We also performed control experiments with bacteria, in which the oscillations were permanently suppressed as shown in Figure 3.6 A. In this case, no correlation was obtained between growth rate and fluorescence (Figure 3.8). further supporting that the experimentally observed dynamic changes in growth rate are indeed caused by the oscillations of the circuit.

There are manifold ways in which oscillator dynamics and growth could be coupled. In our oscillator circuit, the production of GFP is in phase with the production of the λ -cI repressor, which represses the expression of sgRNA and dCas9. Since high GFP fluorescence is correlated with a high instantaneous growth rate, an influence of dCas9-sgRNA expression and impeded growth appears likely [123, 131]. It is further conceivable, that interactions between the circuit and bacterial growth feed back on the oscillator dynamics in a complex way: a reduction in growth speed would also reduce the dilution rate of the oscillator components, which in turn would affect the timing of the relaxation oscillations.

3.1.8 Discussion and Conclusion

In this work, we have integrated transcriptional repression via CRISPR interference into a three-node repressilator network (termed the ‘RPC oscillator’), which resulted in stable genetic oscillations at the single cell level. A statistical analysis of a large number of bacterial oscillators showed considerable variability in the periods among the cells, but also fluctuations of the periods within individual cells. The shape of the oscillatory fluorescence readout signal indicates that the RPC oscillator functions as a relaxation oscillator. Stochastic numerical simulations suggest that titration of the

oscillator components against decoy binding sites as well as the dilution or removal of dCas9-sgRNA complexes from their target sites due to cell growth play a decisive role in its dynamics.

With a period of $T \approx 12\text{h}$, the RPC oscillator is much slower than the original repressilator [3], and slower than the repressilator v2.0 [65]. In terms of bacterial generations (based on an experimentally determined cell doubling time of 53 minutes), however, we obtain an oscillation period between 13 and 14 generations, which actually is very close to the value obtained for the repressilator v2.0 without active degradation [65], and slightly faster than a CRISPRi-only oscillator [121]. Importantly, in our RPC circuit the two conventional repressors were actively degraded as in [3], while the dCas9-sgRNA complexes were not, suggesting that the prolongation of the period is mainly due to the presence of CRISPRi in the circuit and potentially due to its effect on cell growth.

Quite notably, we found that the genetic oscillations of the RPC oscillator were accompanied by periodic changes of the cellular growth rate with the same period. We surmise that this effect might be caused by the periodic cellular load generated by dCas9-sgRNA expression during operation of oscillator, indicating that the use of CRISPRi in dynamical circuits requires special care.

On the other hand, the sequence-programmable knockdown of gene expression via CRISPRi opens up a wide range of potential applications, as secondary processes controlled by this synthetic clock need not be constrained to gene expression from plasmids. It is easily conceivable that the CRISPRi-coupled oscillators could be used to control endogenous, chromosomal genes by simply supplying the corresponding sgRNAs. For this purpose, the lower copy number and expression levels of endogenous genes have to be considered and may require further tuning of the coupled oscillator-load circuit. The use of multiple different sgRNAs in parallel may also lead to dCas9 ‘bottlenecking’ effects [131, 132].

Quite generally, when coupling the output of a gene regulatory circuit to a downstream load process, potential ‘retroactivity’ effects may have to be taken into account [56, 133]. In the context of our circuit, for instance, connecting the CRISPRi node to a downstream node will change the concentration of free dCas9 and thus ‘retroactively’ affect the upstream oscillator circuit. As we included a ‘dummy’ sgRNA in our circuit as a placeholder for the potential control of an endogenous gene, in a sense the retroactivity effect is already accounted for in our circuit. Further, a theoretical analysis of retroactivity effects [56] suggests that the RPC oscillator actually should be quite robust in this respect because of the high expression levels of the repressors observed in our simulations

compared to the relatively low number of the respective repressor binding sites.

In order to explore the potential of the RPC circuit to act as a temporal controller or timekeeper, we analyzed the cell-to-cell variability of the oscillator phase within an information-theoretic framework. The results suggest that the output of a single oscillator node could be used to reliably distinguish three to four phases within a single oscillator period, which could be used, e.g., to control the order of the corresponding number of downstream processes. The oscillator could also be used as a real ‘clock’ to measure ‘absolute time’ starting from a defined point in time. We found that due to fluctuations in the oscillator phase, an initially synchronized population of oscillators loses synchrony over time - after about 40 hours (or four oscillation periods) the error in time is on the order of 6 hours (or half a period).

Based on the above considerations, future studies will aim at further improving the dynamics of the RPC oscillator and explore potential applications. The load on the cells could be reduced by utilizing a genome-integrated gene for dCas9, which potentially would reduce the observed influence on cellular growth. An interesting opportunity could be the use of sgRNA-aptamer fusions [134], which would make the CRISPRi process dependent on the presence of the aptamer’s target molecule. This could be potentially used to tune the oscillator frequency, and also to reset or entrain the synthetic clock.

Rather than using repression by CRISPRi, one could also use CRISPR activation (CRISPRa) [135–137] based on dCas9-activator fusions, both as part of the oscillator circuit or for the control of secondary processes. The most exciting prospect would be to control bacterial gene expression of chromosomal genes with the oscillator, which could be achieved by targeting these genes with the corresponding sgRNAs. This could be utilized, e.g., to couple the oscillator to metabolic processes such as the utilization of certain metabolites. By supplying these metabolites in a rhythmic manner, this in turn could be used to evolve ‘circadian’ oscillators with a desired frequency.

3.2 CRISPRi-Induced Auxotrophy as Selection Principle in Bacterial Cultures

Following up on the RPCi oscillator project, we want to explore ways to couple the state of such a dynamic circuit to the environment. This way, a selection pressure can be imposed on the host of the circuit and mediate a change in relative fitness depending on whether the circuit state is adapted to environment state or not. Over the long term, desirable properties related to the circuit selected can be evolved. More concretely, in the case of the RPCi oscillator, selection pressure could be applied periodically to entrain the oscillator to a desired frequency. We therefore need a way to impose a selection pressure that is controlled by the state of the circuit and the state of environment. Amino acid auxotrophy, meaning the inability to synthesize an amino acid required for growth, is an attractive candidate for this purpose for two reasons: CRISPRi, which is already integrated into the oscillator circuit, can be used to make a strain auxotrophic depending on the circuit state by adding a suitable guide RNA [138]. Furthermore, the presence of the relevant compound in the medium, which can be controlled externally, decides whether the auxotrophic strain can grow. Alternative options such as a killer gene controlled by externally added inducer [139] or switchable antibiotic resistance would require major modifications to the oscillator circuit.

Auxotrophy has been used as a selection principle in other contexts: In knockout strains, amino acid auxotrophy is observed for knockouts targeting specific biosynthesis pathway genes [140]. With a suitable knockout strain, reintroducing the knocked out gene on a plasmid can serve as a selection marker [141] (as an alternative to antibiotic resistance). The selective growth of auxotrophic strains can be used to sense the presence of a non-synthesized compound [142, 143]. Escape from biocontainment can be precluded for auxotrophs that require a nonstandard amino acid for growth [144].

3.2.1 Growth in Different Bacterial Culture Systems

Laboratory evolution procedures for bacteria usually require cultivation over long periods of time [145]. We developed a number of different culturing strategies that are suited for long running growth experiments with *E. coli*. Here we will present two different implementations for automated chemostats (Figure 3.10). The first implementation, a converted bioreactor with added feed inlet and waste outlet, represents the conventional approach for chemostats. It is based on a commercial bioreactor, where the culture is heated, stirred and aerated in a large vessel with optional measurement and control of

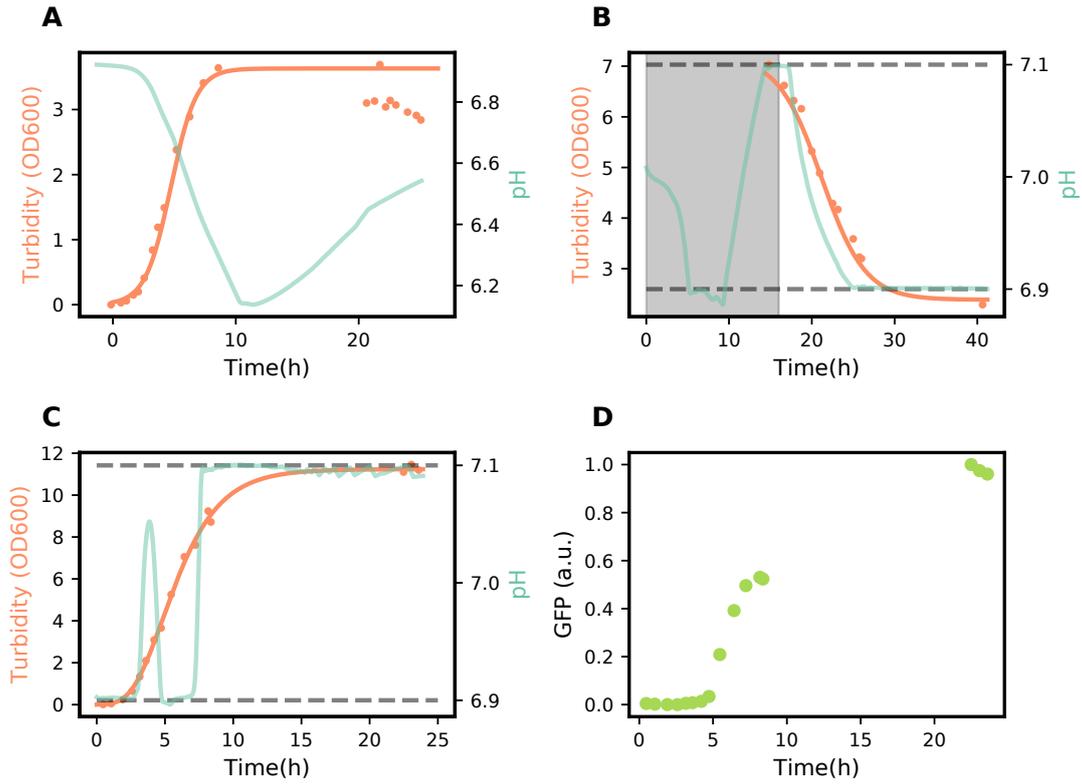


Figure 3.9: Bioreactor growth experiments. Culture turbidity and fluorescence intensity was measured for extracted samples, pH measured automatically. Growth curves were fit with a Gompertz model. **(A)** Continuous culture setup with M9 medium (20 mM glucose), dilution rate $D = 0.1 \text{ h}^{-1}$. A steady state slightly above 3 OD_{600} is reached after 10 h. **(B)** Batch culture for 16 h, then continuous culture with M9 medium (10 mM glucose), dilution rate $D = 0.2 \text{ h}^{-1}$, pH regulation active with 0.1 pH tolerance (dotted horizontal lines). Culture density decreased from 7 OD_{600} to 2 OD_{600} after dilution was turned on. **(C)** Batch growth of auto-induced sender-receiver strain induced with rhamnose (Section 3.4) in LB medium, pH regulation active. **(D)** GFP fluorescence measurements over the duration of C. A delay of 5 h between exponential growth and an exponential increase in fluorescence is observed.

pH and oxygen partial pressure. Additional medium is added at a fixed rate with a software controlled pump. Culture that rises above a fixed fill height is removed with a fixed outlet pipe connected to an external pump (with a higher pump rate than the feed pump). This way, the dilution rate can be controlled accurately and, if required, dynamically.

Figure 3.9 shows example data from growth experiments in the bioreactor chemostat setup. Inlet and outlet flow was set up as described above, resulting in growth going through the same phases as in batch growth but reaching a lower steady state. The

turbidity of the culture was measured externally in a spectrometer, manual samples had to be taken. The best fit for the turbidity datapoints, which follow a sigmoidal shape, was achieved with a Gompertz function [146]. Sensors integrated into the bioreactor provide continuous data points for the partial oxygen pressure and pH of the culture that give additional insight into the metabolism of the culture (pH changes shown in Figure 3.9). The carbon source used by *E. coli* affects the pH in the following way: Use of carbohydrates (preferred) leads to acidification, use of amino acids as carbon source leads to alkalization. Based on the pH data we can conclude that in Figure 3.9 A, the cells metabolize the glucose contained in the M9 minimal medium; the beginning of the stationary phase coincides with the end of carbohydrate metabolism, presumably due to depletion. Subfigure B displays the same phenomenon but carbohydrates become available again when dilution is started, leading to a fall of pH. In this experiment, the built in control system of the bioreactor is used to regulate the pH within set bounds (0.1 around 7 pH) by addition of acid or base when demanded. This system (as well as varied stir rate and air supply in response to measurements of a pO₂ sensor) can remedy some effects adverse to growth that are a consequence of high density bacterial culture and ultimately lead to more biomass at the same nutrient concentration. As a result of this control system the pH value stays fixed at the lower/upper set point as long as the culture tends to acidify/alkalize. The effect of this control system can also be observed in subfigure C, where bacteria were grown in LB medium with added rhamnose inducer using batch mode: Two distinct phases of acidification can be observed, interrupted by a short alkalization phase. It can be theorized that initially glucose is metabolized, after which the less preferred rhamnose inducer is used with a short lag time inbetween [147]. Around the end of exponential phase, as is expected from LB medium, amino acids are metabolized leading to alkalization [146]. The saw pattern produced by the pH control system gets slower over time, this indicates the slow down of metabolism during stationary phase. The pH data thus shows that this is not a steady state, marking the difference between batch mode and chemostat mode which is not immediately clear from turbidity data alone.

In practice the bioreactor chemostat setup suffered from a number of limitations. The minimum volume required to operate the bioreactor put limitations on the achievable dilution rate. A dilution rate D on the order of 1 culture volume per hour means that the software controlled pump needs sufficient throughput and that the fresh medium stock has to hold more than 10 culture volumes for overnight culturing (minimum culture volume of 0.6 L). Smaller bioreactor setups with the potential for parallelization could enable

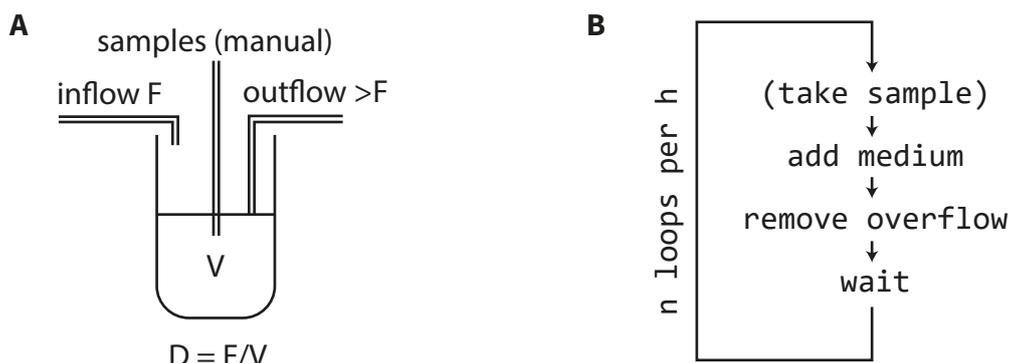


Figure 3.10: Sketch of different chemostat implementations. **(A)** Continuous inflow of fresh medium with rate F into a culture with volume V . Outflow is regulated to a set fill height. The dilution rate is $D = F/V$. Samples can be removed, but require manual handling. **(B)** Dilution in discrete steps with a pipetting robot. Steps are performed in sequential order and looped n times per hour. For each step, f fraction of the culture volume V is exchanged. For high n , this setup approaches continuous culture chemostat **(A)** with dilution rate $D = nf/V$. The advantage of this setup is that it allows the collection and handling (e.g. fixation) of distinct samples.

much higher experimental throughput with a higher variation of possible experimental conditions. The setup was also limited in the capability to record data relevant for synthetic bacterial cultures. We evaluated a light scattering based sensor that can measure the optical density of the culture from outside of the culture vessel and but deemed it to be unsuitable due to limited sensitivity at densities below 1 OD_{600} . A DIY optical density sensor or a fluorescence spectrometer could in principal be attached to an outflow, but biofilm formation would present a significant engineering challenge for long term measurements [37]. Other downstream analysis techniques, such as flow cytometry or qPCR, depend on the collection of separate samples, which cannot be realised with (variable) flow through fixed connections but requires some sort of robotic actuation to switch from sample to sample.

Motivated by the limitations of the bioreactor based chemostat, we developed an alternative approach making use of an Opentrons liquid handling robot. A liquid handling robot or pipetting robot can move and actuate an adapted pipette along 3 axes in space and can be programmed to transfer liquids between arbitrary vessels and tubes (i.e. culture, reservoir or samples). Figure 3.14 B shows a sketch of our implementation of a chemostat in an Opentrons robot, termed ‘chemotrons’. A culture (20 mL) is heated and stirred with a hotplate and magnetic stirrer, placed in the pipetting robot enclosure. The culture is diluted in discrete steps by pipetting fresh medium with fraction f of

the overall culture volume from a reservoir to the culture and subsequently removing culture. This dilution step is carried out n times per hour using a 1000 μL pipette, with the primary limitation of limited pipette tip supply. In regular intervals, samples were transferred from the culture to separate tubes for later analysis with flow cytometry.

The scattering flow cytometer data was irregular for samples stored over night in the device, presumably due to cells bursting. To counteract this, different fixation protocols based on formaldehyde crosslinking were tested to investigate the feasibility of protocols that do not require cooling or washing steps, which are difficult to implement with a pipetting robot. Figure 3.11 shows the protocols in detail and for each protocol the distribution of GFP fluorescence of a sample of a culture taken directly after induction and 1 h after induction. An unsuitable fixation protocol might result in spurious production of GFP after the sample is removed from the culture. The data shows that all fixation protocols resulted in minimal differences to a baseline of samples preserved at -80°C . For further experiments, a protocol including formaldehyde and kanamycin was chosen, the latter serving to interrupt protein biosynthesis and prevent any changes in the cellular proteome.

The number of dilution steps per hour determines the extent of the difference between a continuous flow based chemostat setup and a chemostat setup with discrete dilution steps (Figure 3.10). Instead of a completely stable steady state, discrete dilution steps will result in a see saw pattern for the culture density, sharply falling after every dilution event and slowly rising after the addition of fresh medium until overshooting the steady state. Due to the non linear decline in growth rate for decreasing substrate concentration, this overshooting of the steady state leads to a lower average growth rate and thus a lowered steady state. An extreme example would be the daily passaging of cultures, which can be considered as growth of a single culture with dilution steps spaced out 24 h. There, bacteria may spend the majority of time in stationary phase, resulting in reduced average growth rate (compared to continuous culture with the same overall dilution). We examined this effect with an ODE growth model (Section 2.1.1) with realistic parameters. The result in Figure 3.12 B shows that for less than one dilution step per hour, the growth rate and thus the steady state density of the culture is significantly reduced. But for more than two dilution steps per hour, the effect can be considered as small and the chemotrons can be directly compared to a flow based continuous culture with dilution rate $D = nf/V$. Figure 3.12 A shows an example of the growth of a culture in the ‘chemotrons’. A steady state, with some fluctuation, is reached after 6 h. Together with the validated sampling protocol, this proves that this automated setup is suited for

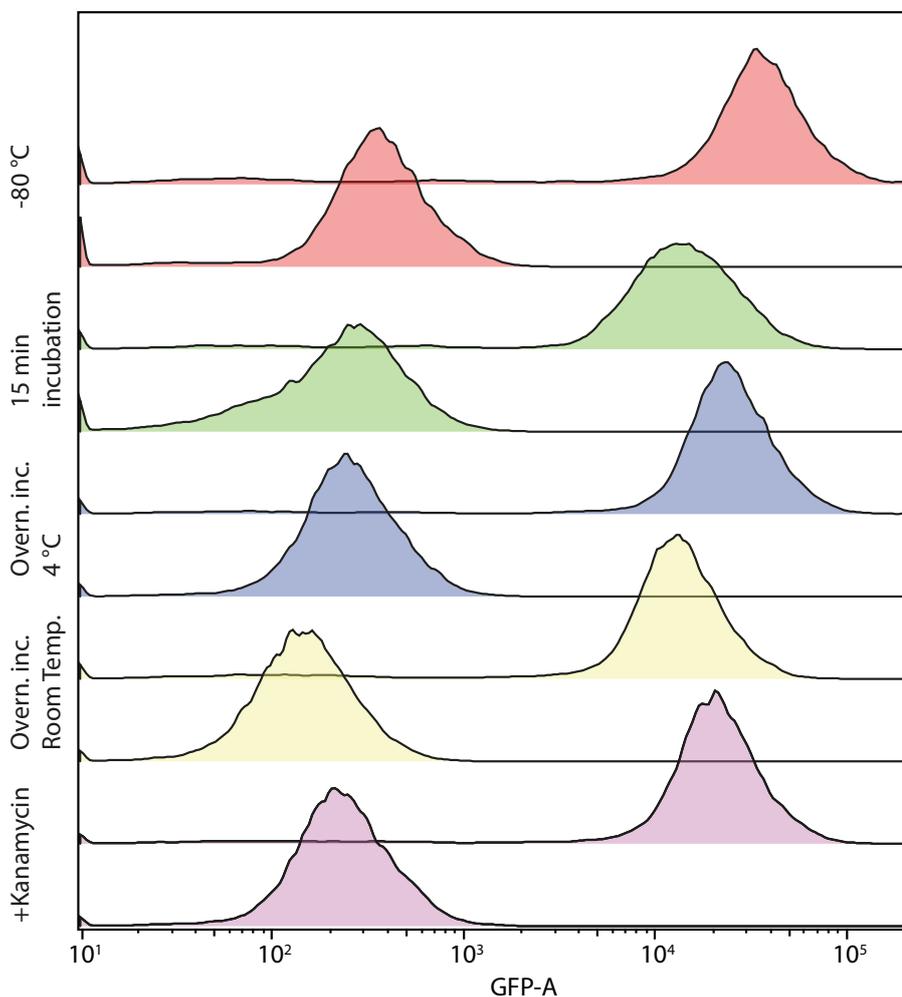


Figure 3.11: Comparison of fixation protocols for flow cytometry of *E. coli* with implementation feasibility in a pipetting robot in mind. The GFP fluorescence of samples taken immediately after induction of GFP expression and 1 h afterwards are compared. Forward and side scatter data looked very similar for all protocols. Immediate freezing at $-80\text{ }^{\circ}\text{C}$ (with 25% glycerol added) was used as baseline. All following protocols are based on formaldehyde fixation. The second protocol includes a initial washing step in PBS, 15 min incubation with formaldehyde (2%), and finally two washing steps in PBS (washing steps are not possible in a basic pipetting robot). The next two protocols provide for 1:1 mixing of culture and 4% formaldehyde and overnight incubation at fridge or room temperature respectively. In the final protocol, formaldehyde was added to the culture up to 0.8%, then mixed 1:1 with PBS and $200\text{ }\mu\text{g mL}^{-1}$ kanamycin added (kanamycin stops protein biosynthesis). The kanamycin protocol was used in subsequent experiments.

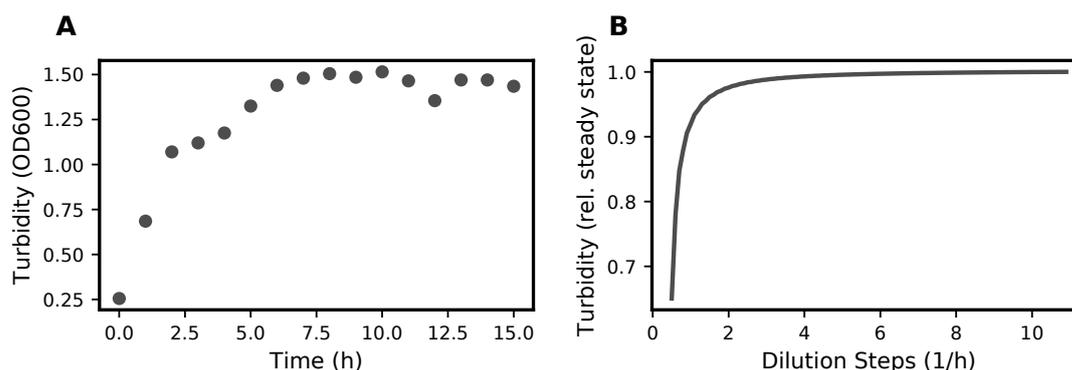


Figure 3.12: Growth in the “chemotrons”: Example data and theoretical model. **(A)** Growth of the RPCi repressilator (Figure 3.13) in LB medium with 0.8 h^{-1} dilution rate and 2 dilution steps per hour. A steady state is reached after about 6 h. **(B)** The effect of non continuous dilution on the steady state level of a chemostat. Based on a realistic model for *E. coli* growth, dividing continuous dilution into discrete steps lowers the steady state culture turbidity. For a high number of dilution steps, the difference becomes negligible.

the time resolved analysis of long term growth under stable conditions with informative techniques such as flow cytometry.

Figure 3.13 shows the population fluorescence of the RPCi bacterial oscillator grown with the ‘chemotrons’ setup. The oscillator is synchronized at the start of the experiment as described in section 3.1 and subsequently grown for 15 h (the experiment was ended prematurely due to instability of the setup). At the start of the experiment, the population of oscillators has a high fluorescence level as a direct consequence of the initial induction of the pTet promoter of the GFP reporter (with the purpose of synchronization). We know from microfluidics experiments in section 3.1 that this genetic circuit does in fact lead to oscillations of individual bacterial cells, which can in principle not be determined from the population level data gained from flow cytometry. However the time resolved data in Figure 3.13 suggests that over the course of the experiment, the synchronized oscillators split into two subpopulations. After the fluorescence of the entire populations decreases over the first 4 h, for one subpopulation the fluorescence switches to a high state again starting at around 9 h. Another subpopulation remains at a low fluorescence level, leaving to a bimodal distribution. A plausible interpretation for these results is that after successful initial synchronization, one part of the population did not show any dynamic activity and remained in an off state while the other part of the population did show synchronized, population level oscillations. This interpretation is consistent with the results from microfluidic experiments, where oscillations were observed on a similar

time scale and examples of cells with non functioning oscillator were observed.

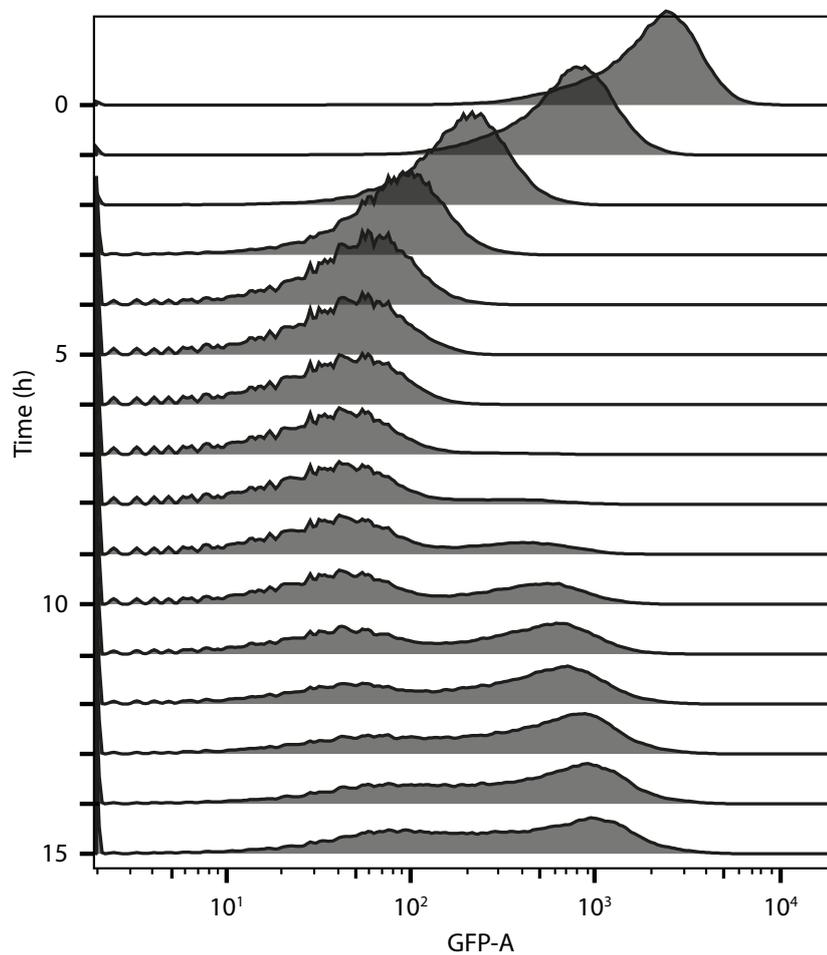


Figure 3.13: Pipetting robot chemostat: FACS measurement of a continuously sampled RPCi oscillator with 0.8 h^{-1} dilution rate and one sample per hour. GFP fluorescence population distribution measured with flow cytometry. Before the start of the experiment the oscillators are synchronized by adding aTc, consequently GFP is high at the beginning. After 3 h the oscillators turn off and after around 9 h, the majority fraction of the population fraction turns on again. The experiment was ended prematurely when the culture vessel was knocked over by the pipetting robot arm.

3.2.2 Amino Acid Auxotrophy Mediated by CRISPRi

This section presents a proof of concept for amino acid auxotrophy mediated by CRISPRi with suitable guide RNAs. As sketched by Figure 3.14, the CRISPRi system depends on the guide RNA sequence and is switchable with an inducer. If CRISPR interference is effective, growth depends on the presence of the targeted amino acid in the medium. To validate the principle of CRISPRi induced auxotrophy, *E. coli* strain MG1655 containing plasmids for sgRNA and dCas9 expression under a pTet promoter were cultured in with all three different inputs (guide RNA sequence, inducer, amino acid supplemented to the medium). We initially screened 6 different guide RNAs targeting the synthesis pathways of 4 different amino acids, resulting in 2 candidates for tryptophan and glycine auxotrophy, respectively. Target genes were chosen based on auxotrophs previously identified in knockout studies [140, 148].

Figure 3.15 shows the density of aux cultures reached after a fixed time, A and B show the results for two different targeted amino acids, tryptophan and glycine. In both cases, the lack of growth for the induced sample and the in contrast to the uninduced sample with the targeted amino acid added to the medium shows that CRISPRi can cause auxotrophy. The growth of the control samples with a non targeting, random, guide RNA proves that this effect is target specific. Furthermore, the growth of the uninduced samples shows that CRISPRi remains switchable for this application.

The difference between subfigures A and B lies in the base medium: For A, M9 medium with glucose and without any amino acids added was utilized (common M9 recipes include undefined amino acid sources such as casein). This lead to slow overall growth, with the fastest growing sample reached only a moderate turbidity of 0.3 OD₆₀₀ after 24 h of growth. Since slow growth would be disadvantageous for downstream applications, Neidhardt Rich Defined Medium (RDM) [149, 150], modified to exclude relevant amino acids, was utilized for the experiment shown in subfigure B. This medium supports growth rates similar to rich media as LB and the concentration of all ingredients is exactly known (and can be varied). The results in Figure 3.15 show that RDM supports more than ten times denser growth over a shorter time period (albeit comparing different targeted amino acids). In all subsequent auxotrophy experiments RDM was utilized. The downsides of RDM are a laborious preparation procedure, especially if amino acid contents are varied, and its heat sensitivity that renders it unsuitable for autoclaving.

Additional experiments for the glycine auxotroph in RD medium hinted at the fact that growth may be rescued with the external addition to the medium of not only glycine,

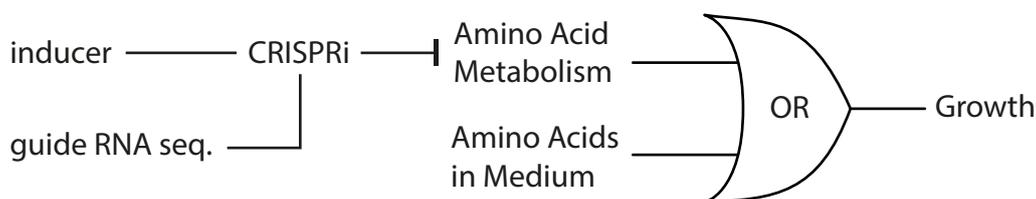


Figure 3.14: Sketch of CRISPRi induced auxotrophy. The cell needs essential amino acids to grow, they can come from the medium or the internal synthesis machinery. Only if neither source is available the cell will not grow, here described with OR logic. Variable inputs to the system are the sequence of the guide RNA (target binding or random sequence control), inducer for dCas9 expression and the amino acids in the medium. Growth is the observable output. Mediated by induced metabolic constraints, the state of the environment and the genetic circuit are coupled to affect the fitness of the cell. This way, modification of the environment can be employed to select for a specific circuit state.

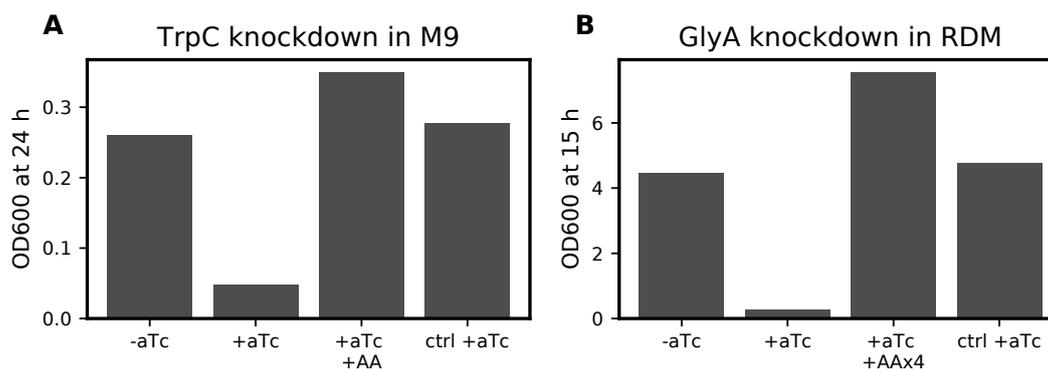


Figure 3.15: CRISPRi Auxotrophy proof of principle. The base medium does not contain the respective amino acid which is targeted. **(A)** Tryptophan auxotrophy in M9 minimal medium, growth experiment in a shaking flask. CRISPRi with TrpC targeting sgRNA prevents growth, switchable with inducer. Supplementation of tryptophan to the medium (+AA) rescues growth. A random sequence sgRNA does not prevent growth. No sample grew to high densities in this medium over 24 h. **(B)** Glycine auxotrophy in rich defined medium without 4 amino acids (glycine, alanine, serine, threonine, all other amino acids are contained in the medium), growth experiment in a plate reader. CRISPRi with GlyA targeting sgRNA works similarly as **A**. Instead of one, 4 amino acids are missing from the medium are added to rescue growth. This medium supports faster growth than M9.

but potentially other amino acids. Considering the metabolic pathways for the synthesis of glycine, compensation for the loss of the targeted *glyA* gene by a number of backup pathways related to other amino acids seemed plausible [151]. To elucidate this effect, the growth of the glycine auxotroph was tested with RDM containing all but glycine and a combination of three other amino acids: Alanine, serine and threonine. Figure 3.16 shows the result for all possible combinations: All combinations that are not lacking serine did grow in spite of missing glycine. On the other hand, media not containing serine and glycine did not grow. The presence or absence of threonine or alanine showed no effect.

The results of this experiment show that, in addition to glycine, the concentration of serine also is an important factor for the growth of a *glyA* deficient strain. For future experiments, such unexpected interactions should be expected due to the complexity of the network of metabolic pathways. In a natural microbial environment or a synthetic microbial system, even more complex cross-feeding interactions between species may have to be considered. With sufficient knowledge of relevant microbiology and systems biology, it may be possible to predict these unforeseen interactions and potentially utilize them to engineer novel behaviour. Otherwise broad empirical trials may be prudent to catch any effects that emerge from the interaction with a complex system such as bacterial metabolism.

In a next step, we evaluated the possibility to extend the capabilities of the CRISPRi auxotrophy system with an additional possible input through the modification of the single guide RNA with theophylline aptamers. Figure 3.17 shows the growth resulting from three different sgRNA variants targeting *glyA*: The standard sgRNA without modifications and sgRNAs with a theophylline aptamer supposed to be activated (*ligRNA+*) or deactivated (*ligRNA-*) by the presence of the ligand molecule [134]. The first 5 panels in Figure 3.17 in principal show the same result as previous experiments, namely that the CRISPRi *glyA* target causes auxotrophy that is switchable with an inducer. The time course data reveals that growth only starts after a long lag time for the uninduced and control sgRNA sample. The addition of theophylline has no effect on the sample with standard sgRNA.

The aptamer sgRNA samples show mixed results: The *ligRNA* sample-, auxotroph in the absence of the ligand, grows differently depending on the addition of theophylline. However the auxotroph shows more growth than the standard sgRNA auxotroph, perhaps indicating that the modified sgRNA has lower affinity to the target. The *ligRNA+* sample did not show a response to the ligand.

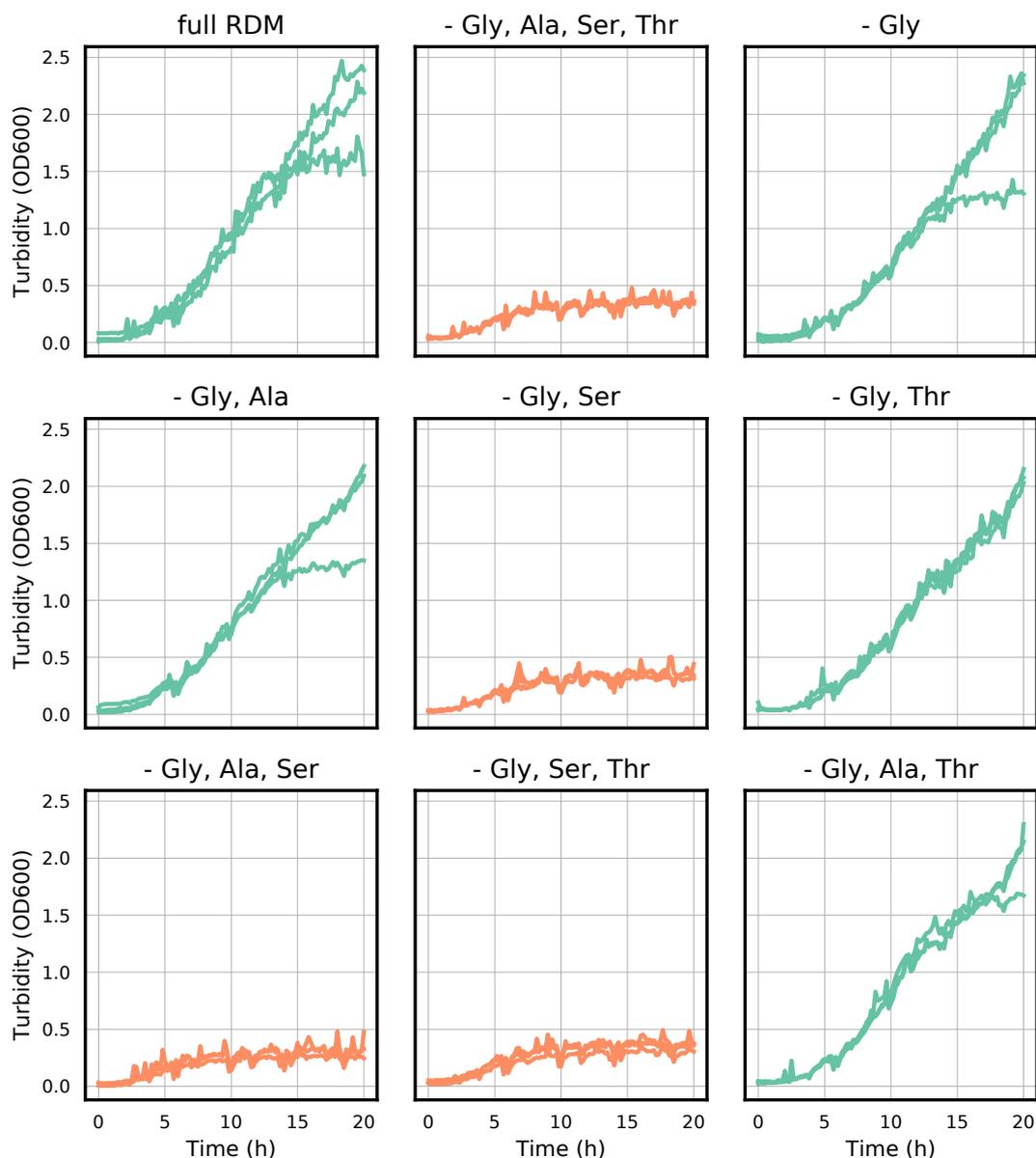


Figure 3.16: *CRISPRi glycine auxotrophy: Growth experiment in a plate reader with added inducer, base medium contains all amino acids except glycine or alanine, serine, threonine, or any combination of those 3. Glycine, whose synthesis is targeted, is expected to be required for growth. No growth is observed in all samples lacking glycine and additionally serine. Samples containing serine show growth. This suggests that glycine synthesis is possible from serine when GlyA is targeted. After about 30 h, samples that did not grow initially also started to grow (Appendix Figure 6.5).*

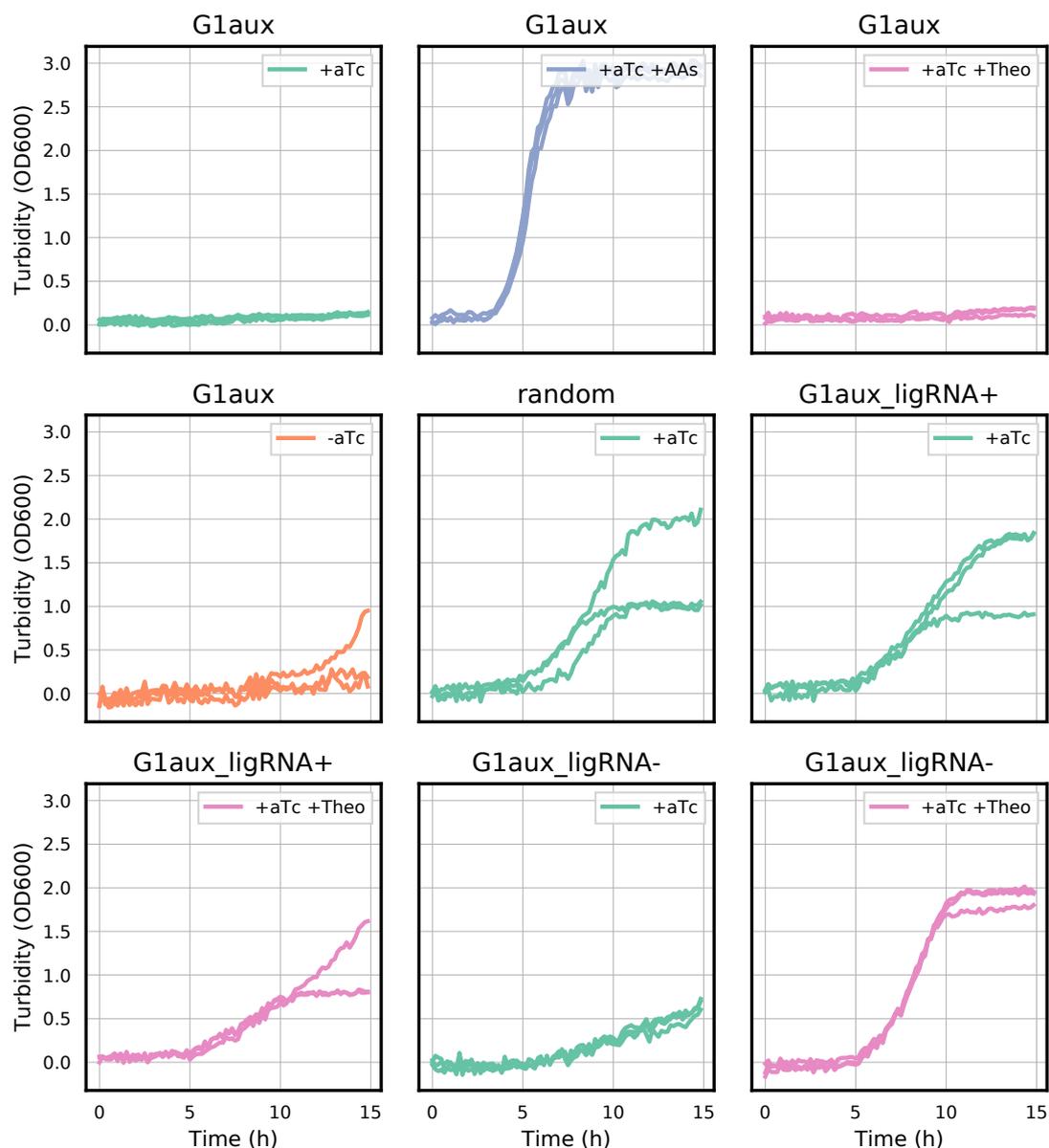


Figure 3.17: Growth experiment in a plate reader testing *sgRNA* variations for glycine auxotrophy (triplicates). Theophylline (*theo*) *sgRNA* aptamers [134] are compared to normal *GlyA* targeting *sgRNA*. Rich defined medium without 4 AAs is the base medium (Figure 3.15). Long lag times are observed for uninduced or control *sgRNA* samples. The addition of theophylline has no effect on the sample with normal *sgRNA*. The *sgRNA* with *ligRNA*⁻ aptamer (active in the absence of *theo*) can be switched to prevent growth, although weaker than the normal *sgRNA*. The *sgRNA* with *ligRNA*⁺ aptamer (active in the presence of *theo*) does not have a (switchable) effect on growth.

3.2.3 Outlook

In a next step, a selection procedure based on the presented CRISPRi auxotrophy system should be tested. By co-culturing the strain with GlyA targeting sgRNA and the random sgRNA control strain under selective conditions (glycine and serine deficient medium, with inducer), the difference in growth rates should lead to an enrichment of the non auxotrophic strain. The chemotrons setup could be employed to continuously sample a long-running chemostat culture. Preliminary experiments have shown that quantitative PCR (qPCR) with primers annealing to the sgRNA target sequence can be utilized to measure the relative concentration of a strain in a co-culture. Together, the chemotrons setup and downstream qPCR analysis would offer insight into the population dynamics of mixed cultures depending on varied auxotrophy properties. For the upstream integration of CRISPRi auxotrophy into the RPCi oscillator, a theoretical model based on growth rate differences could inform further experiments. The dependence of the oscillator period time on the growth rate [65] and the general interplay of growth and the oscillator circuit could lead to unexpected effects if the growth rate is altered with periodic interference with the host metabolism. With a suitable choice of timescales for the external control of selection pressure, the addition of the glyA targeting sgRNA to the RPCi oscillator could allow the tuning of the oscillator frequency with an entrainment procedures [152].

3.3 Fabrication of Sub-Micron Scale Microfluidic Chips

This chapter covers the development of sub-micron scale microfluidic chips for bacteria experiments. Recent advances in nanofabrication have been utilized for the quantitative study of bacterial cells in microfluidic chips [153], covered in detail in a review by Hol and Dekker (2014) [26]. On one hand, small scale microfluidic traps allow the confinement of bacterial cells and in turn their observation at a single cell level. The so called ‘mother machine’ developed by Wang et al. [33], where cells get trapped in growth channels whose cross-section fits only a single cell, allows for continuous, long-term growth with defined inheritance following cell-division. The well defined structure of such a device aids the quantitative analysis of genetic circuits in bacteria. One focus of this chapter is the development of such a chip for the long-term observation of dynamic genetic circuits in *E. coli*, requiring a resolution of 1 μm (Figure 3.18).

On the other hand, interactions of bacteria with their environment or interactions between bacteria mediated by their environment can be studied under controlled conditions in a synthetic spatial environment. As an example, Phan et al. (2020) [154] explored how bacteria grow and move in microfluidic channels arranged in a complex maze or fractal pattern (similar work by Tokárová et al. (2021) [155]) and Gupta et al. (2020) [156] investigated the dynamics of microbial consortia with diffusion-mediated interactions in spatially structured environments. Such experiments can help to close the gap between microbiology experiments in featureless spatial environments like agar plates or bulk liquid culture and the intricate and complex natural environments of bacteria. Spatially separated but connected communities can interact and cooperate in a synthetic ecosystem, whose stability depends on the spatial structure [157]. For the study of such systems sub-micron channels may be utilized, specifically because they are too small for *E. coli* cells to fit but still allow the exchange of small molecules such as autoinducers or metabolites, thus acting as a filter.

In the following subsection, the fabrication of a mould for sub-micron structures via lithographic patterning of SU-8 photoresist is covered. Furthermore, selected relevant steps in the chip production and setup of long-running microfluidic experiments and finally optimization strategies are discussed.

3.3.1 Fabrication

Lithographic patterning of micrometer features lies at the lower resolution limit of many fabrication methods. The limiting step is usually the exposure of the pattern, although

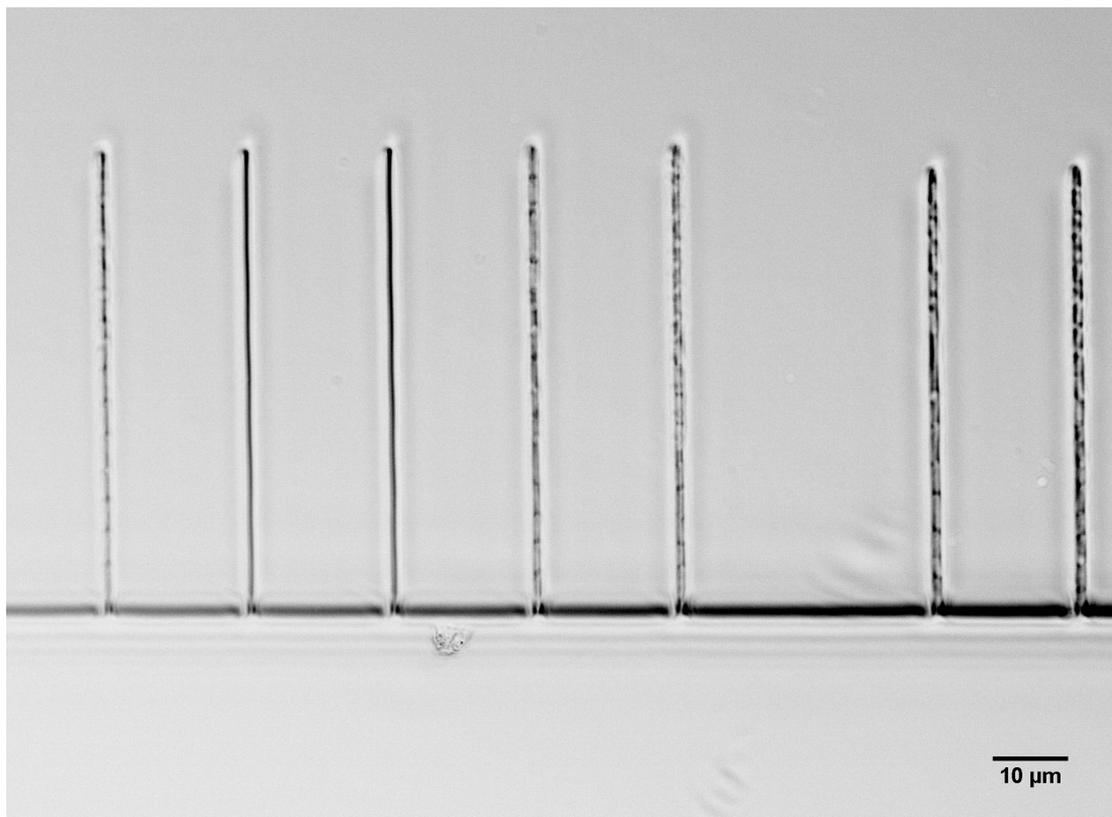


Figure 3.18: Microfluidic device for the observation of single cells during continuous growth (‘mother machine’). Image shows 7 growth channels, 5 of which are loaded with bacteria. Growth channels are about 1 μm wide. This specific chip was found to be not suitable for long-term experiments due to bad cell loading and irregular growth, presumably because the growth channels are too long.

	Manual Mask Aligner	Maskless Aligner	EBL
Illumination	UV lamp	UV LED	electron beam
Patterning	photomask	DMD chip	beam scanning
Resolution	1 μm	1 μm	0.01 μm
Exposure	seconds	minutes	hours
Design iteration	weeks	hours	days
Alignment	10 μm	sub μm	sub μm
Reproducibility	low	high	medium
Setup	manual tinkering	minimal	slow

Table 3.1: Comparison of a manual mask aligner, maskless aligner and electron beam lithography as exposure method for the fabrication of micrometer scale features in SU-8 resist. Quantities are meant as order of magnitude estimate.

photoresist properties may also be a factor. Here we compare three different exposure methods for this specific application (overview in Table 3.1): Photomask based exposure with a manual mask aligner, exposure with a maskless aligner (MLA) and electron beam lithography. This comparison is based on experience with the specific devices available in the ZNN clean room at TUM (see Methods), but should generalize to similar equipment in other academic research labs.

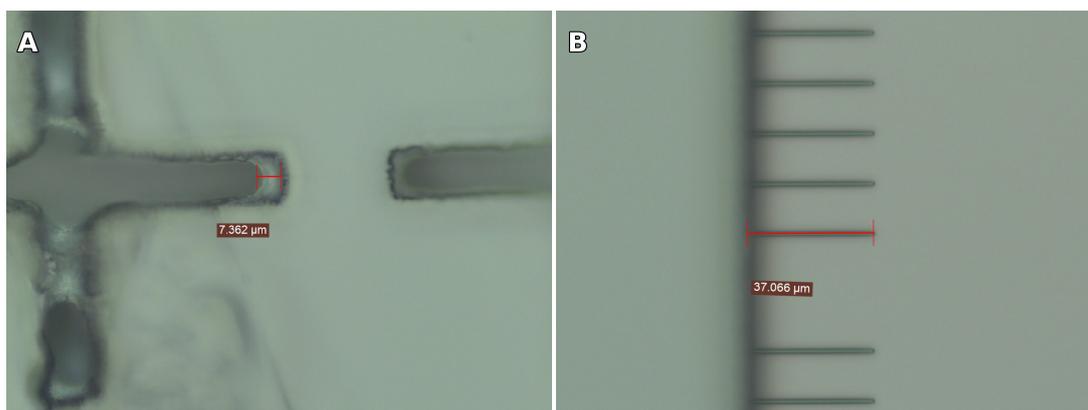


Figure 3.19: *Small errors from manual alignment have a significant impact on the function of the structure. (A) Manual alignment with the help of cross shaped alignment markers. After structure development an offset of around 7 μm was observed. (B) Offset between growth channel layer and feed channel layer result in 25 % error in growth channel length, significantly impacting the function of the channel.*

In a manual mask aligner, a coated wafer is exposed with a UV lamp (mercury-vapor) through a photomask to project the desired pattern onto a surface. Because the whole wafer is illuminated at the same time, the duration of the exposure procedure can be very short (for this reason related methods are used in the mass production of semiconductors). The resolution is limited by the projection fidelity (distance and alignment between photomask and substrate) and the resolution of the photomask. One principal disadvantage of this method is the necessity to produce or order a photomask for every design variation, leading to iteration times in the order of weeks. For established fabrication procedures, this disadvantage does not apply. To achieve alignment between layers, the relative position of the stages holding photomask and coated wafer is manipulated manually with micrometer screws. The correctness of the alignment can be gauged visually through a microscope by comparing the deviation between alignment markers on wafer and photomask. Given the stages are parallel and wafer and photomask are dimensionally stable, two translation and one rotation degree of freedom need to be adjusted. Structures

with features in the micrometer range usually require alignment between a thin/small layer and a thicker layer that connects the small structures. Small mistakes in the manual alignment process can affect critical dimensions of the small structure, such as the length of a growth channel (Figure 3.19). Therefore, precision on a similar scale as the smallest structure is usually required for alignment, placing high demands on the precision of both operator and the mechanical mechanism. In principle the alignment of two lines can be judged with high precision by human vision. Offsets between two lines much smaller than the ‘resolution’ of the retina may be detected, an effect called *hyperacuity* [158]. Alignment markers may be further optimized with this in mind, e.g. with a vernier scale pattern. On the other hand, alignment markers themselves are subject to similar problems as other small, delicate micro-structures, namely exposure quality and possible mechanical damage. Imperfect alignment marker structures make accurate alignment very unlikely. In conclusion, while chips with micrometer scale structures may be fabricated using a manual mask aligner, the limitations of this method cause such a procedure to be prone to errors and poorly reproducible.

Maskless lithography is a promising alternative to photomask based lithography that can be achieved with different techniques. In the maskless aligner device (MLA, Heidelberg Instruments), a desired pattern is projected onto the substrate by selective reflection off a digital micromirror device (DMD or alternatively DLP). A digital micromirror device is a micro electromechanical system (MEMS) which consists of a grid of mirrors that can be individually tilted in order to reflect incident light either onto a target or away from it (to a beam dump). In this way, a raster of pixels can be projected at once with the UV LED light source. Since the DMD chip is not large enough to pattern a large substrate with sufficient resolution, the desired pattern is converted into separate sections that are exposed in sequential order (with automatic step motion of the substrate stage between exposures). The direct conversion of user generated CAD patterns into machine instructions for the DMD chip and stepper motors allows for rapid design iteration. The resolution of a DMD based maskless aligner depends on the specifications of the device, the ‘pixel’ size of the individual mirrors and the projection optics (with the diffraction limit of UV light as lower bound for the resolution).

Figure 3.20 shows examples of structures with a line width down to $0.6\ \mu\text{m}$ and line spacing down to $1.8\ \mu\text{m}$ that were exposed with the MLA. The resulting line width can deviate from the CAD pattern. The variable automated patterning helps to compensate this effect through rapid screening of test structures with varied line width and other parameters. High resolution cameras allow for a precise and reproducible semi-automatic

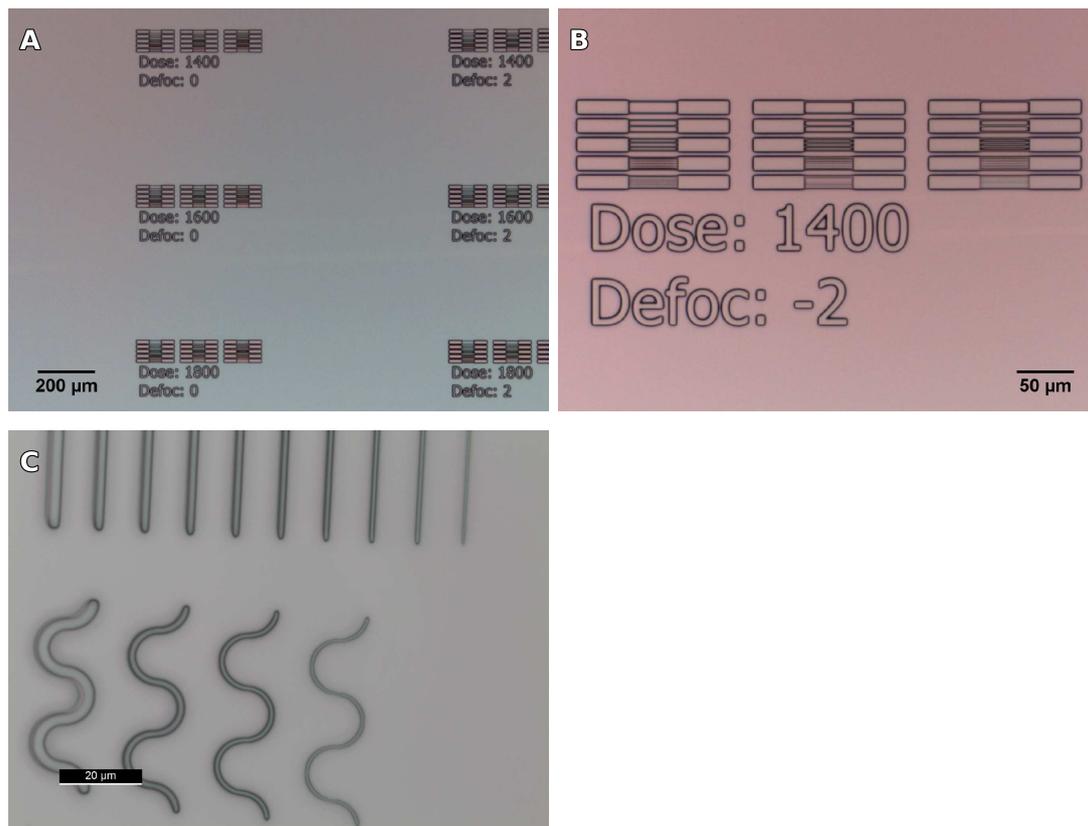


Figure 3.20: Test structures are essential to find suitable design and exposure parameters. Automated patterning with different exposure parameters using a maskless aligner and test structures featuring varying line widths allow a comprehensive parameter screen in a single experiment. Sub-micron resolution in SU-8 can reliably be achieved with a maskless aligner. (A) A test structure is exposed repeatedly on a grid with a maskless aligner, with varying parameters for dose and focus. (B) Closeup of the test structure: Two large channels are connected with sub micron lines with varying width and spacing. For the bottom-most structures, the fine spacing is not sufficient to distinguish the thin lines. The third row shows that four parallel channels across 10 μm can be fabricated reliably (down to 1.8 μm line spacing). (C) A test structure for straight and curved lines of varying width. The thinnest line width determined from the microscopy image is 0.6 μm .

alignment procedure. The location of alignment markers is detected in the camera image and referenced to the stage position. By scanning four alignment markers in this way, the device software can automatically adjust the projection of the pattern to achieve alignment between layers.

A different approach to maskless lithography is patterning with a focused beam. To expose an area, the beam is scanned along a suitable path. Here, we discuss electron beam lithography (EBL), but similar systems exist based on UV lasers or focused ion

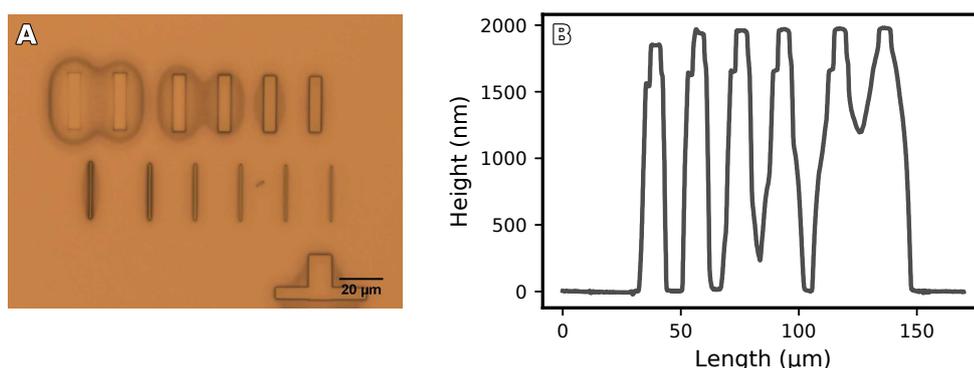


Figure 3.21: Proximity effects of electron beam dose. (A) Microscope image of an e-beam lithography test pattern: Thick and thin lines exposed with a range of doses. Around the thick lines on the left, proximity effects from over exposure are visible. (B) Height profile through the thick lines reveals poor line separation at high beam dose.

beams. The use of a focused electron beam results in very fine resolution more than sufficient for the stated goal of sub-micron SU-8 structures. One major disadvantage of the scanning approach compared to raster exposure is the long exposure time per area. Exposing a few hundred $1\ \mu\text{m} \times 40\ \mu\text{m}$ squares (around $0.01\ \text{mm}^2$) of SU-8 2000.5 usually takes around 1 hour. This means that exposing wafer scale features (cm^2) with e-beam lithography is not practical.

The details of how the dose of the electrons is deposited in the resist are important. The primary electrons will scatter and produce a cascade of secondary electrons with lower energy in the proximity of the incident location. While SU-8 resist is designed to trigger a chemical polymerization reaction in response to UV light, these scattering events are found to be sufficient for the same effect [159]. Thus the fidelity of the process depends not only on the beam properties (acceleration voltage, focus, aperture, stigmation), but also on the substrate and photoresist properties and, more critically, on the specific pattern (e.g. the dose required to adequately expose a specific feature depends on the neighbouring features due to proximity effects). For this reason, the micro structures that can be achieved with EBL and the suitable parameters (such as the dose) is subject to experimental trial and error (commercial simulation software can be used to predict proximity effects). This fact increases the time required for iterative optimization compared to more predictable UV based exposure and makes isolated screening for suitable parameters for specific features prudent. Figure 3.21 shows an example of a dose screening for a simple line structure, where the proximity effect can be clearly observed

for high doses. Examples in the literature show that SU-8 structures for single cell studies with line widths of 100 nm can be achieved with EBL [160]. Another challenge for EBL can be the setup procedure required for every session. A number of parameters related to the electron beam, focus and alignment of the substrate have to be adjusted with the EBL device operated in scanning electron microscope mode (SEM). The variance and potential for user error of the setup process can be a limiting factor for the reproducibility of EBL exposure and the required expenditure of time may make other methods more appealing if the specific benefits of EBL are not strictly necessary.

The development of the fabrication process is centered around meeting the feature width and height specified in the design phase (and adapting the specifications when necessary). Firstly it has to be established that the chosen exposure method and process is in principal suited for the width and height requirements. The variation between different devices and setups means that theoretical approaches or protocols from different labs have to be verified and adapted to the specifics of the particular facilities used.

The width of features in developed photoresist depends on the exposure pattern width, exposure dose and other secondary process properties. Systematic screenings of those two parameters can be evaluated under a conventional microscope or alternatively with scanning electron microscopy (SEM). Figure 3.20 A and B show an example of a systematic screening of exposure and focus parameters for a test structure. The test structure includes lines of varying width and spacing. The cross section profile of a pattern (ideally vertical side walls in most cases) may depend on the depth of focus of the exposure and photoresist properties (and secondary electron effects in the case of e-beam lithography); it can be evaluated with a stylus profilometer or tilted-view SEM.

The height of a structure made out of epoxy based photoresist like SU-8 is determined by the layer thickness that is achieved in the coating step prior to exposure. In the case of spin coating, the layer thickness is mainly a function of the resist dilution (and thus the viscosity) and the spin coating rotation speed. Systematic screens of these parameters or a datasheet supplied by the photoresist vendor allow the approximate interpolation of parameters suitable for the desired height. Such approximations can be suitable for less sensitive applications; to meet exact height specifications the more easily varied spin coating rotation speed can be fine tuned. Figure 3.22 shows example data of systematic screening of layer heights in the 1 micron range for SU-8 2000 and SU-8 3000 resist.

The ratio of width and height specifications, the aspect ratio, mainly poses a challenge for post exposure processing steps. Thin, tall structures suffer from poor adhesion to the substrate and mechanical fragility. The adhesion can be improved with a uniform

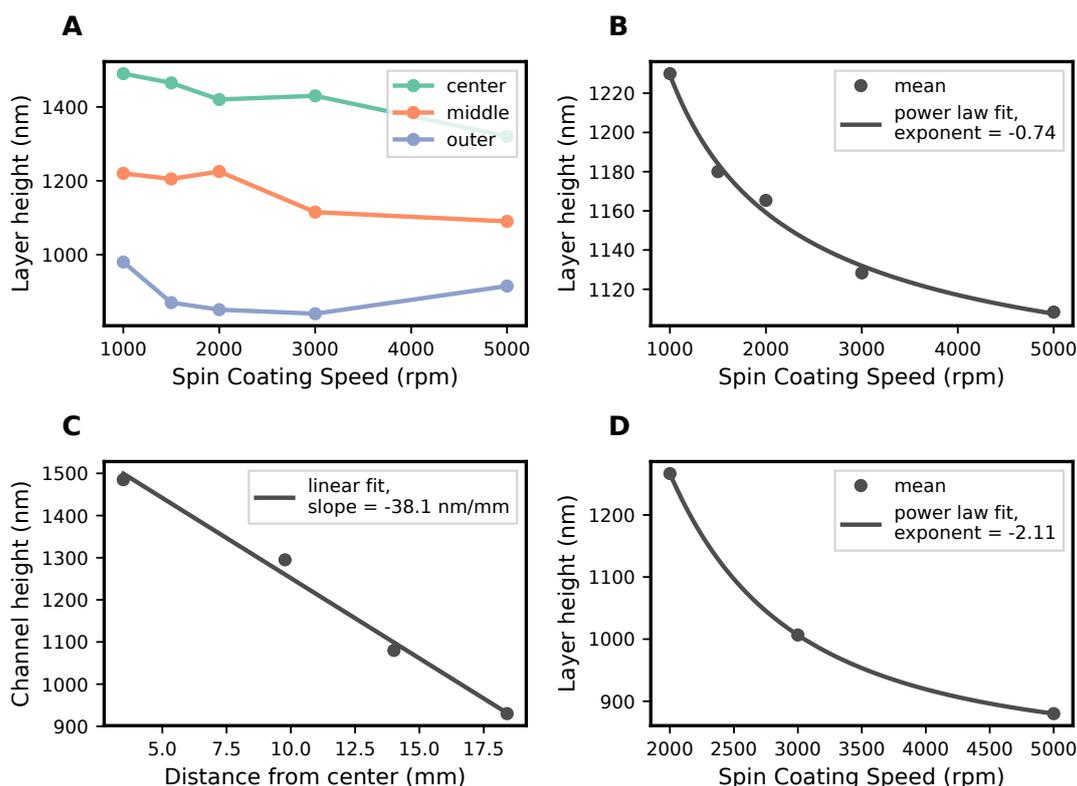


Figure 3.22: Thickness of spin coated SU-8 photoresist layers on a 2-inch wafer, measurements performed with a profilometer. (A) Layer height of SU-8 2000.5 depends on the spin rotation speed. Thickness in the range of 1 μm can be achieved with this dilution, depending also on the distance from the substrate center. (B) Results from A averaged over position more clearly shows the relationship between layer height and rotation speed. (C) Measured height of exposed and developed growth channels. The height decreases linearly with distance from the center (centrifugal force). (D) Layer height of SU-8 3000 resist with 19% solid content. Similar heights to SU-8 2000.5 can be achieved despite not officially supported in the datasheet.

SU-8 base layer. Mechanically fragile structures have to be processed with care, strong agitation during development and washing steps as well as the mechanical force exerted by a profilometer stylus may damage the structure. Prior to exposing the structure to the mechanical strain of the PDMS casting process, a hard bake step is essential to extend the lifetime of fragile structures. Figure 3.23 A and B shows examples of mechanically deformed or otherwise damaged structures. The 1 micron wide channels in Figure 3.23 A have broken after repeated PDMS casting procedures while the sub 1 micron channels in subfigure B presumably collapsed during development of the exposed structure, creating an interesting wave-like pattern. Figure 3.23 C shows an example of fractures in a large

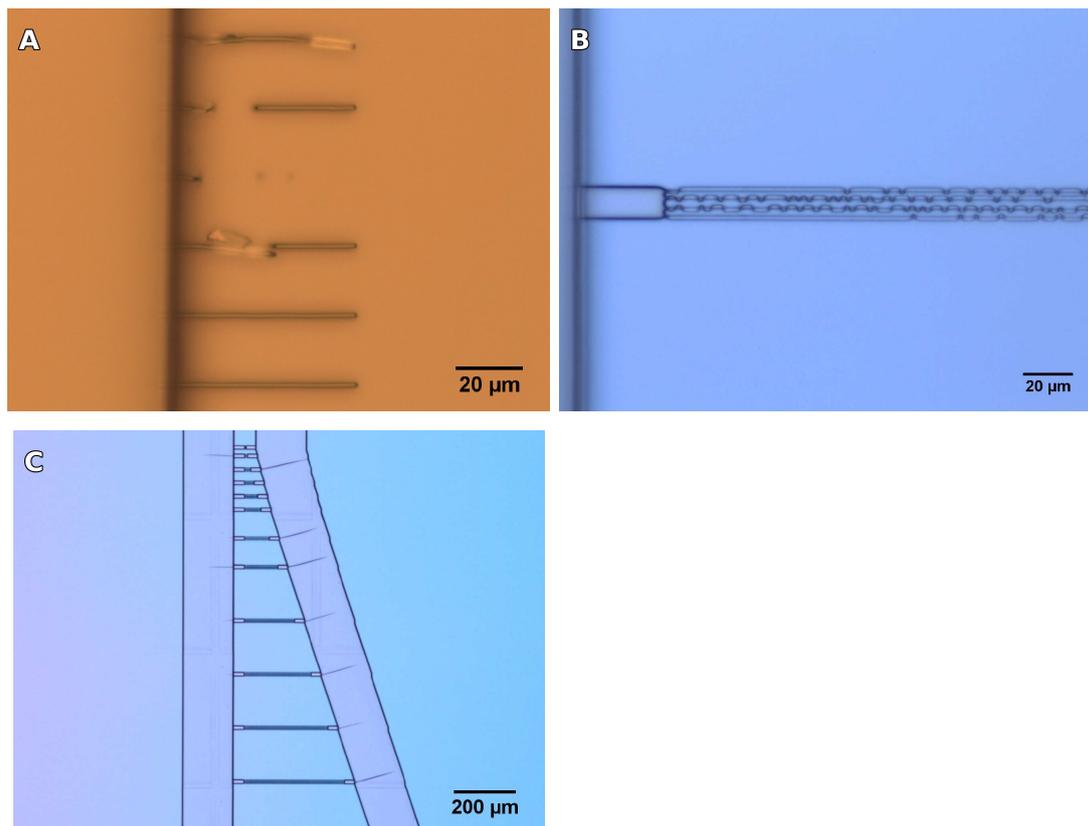


Figure 3.23: Mechanical damage and residual stress of thin and shallow channels (A) Around $1\ \mu\text{m}$ channels are partially damaged after repeated PDMS casting. (B) Sub micron channels collapsed in a wavy pattern during post exposure processing. (C) The right diagonal channel contains regular fractures located at sharp bends in the channel geometry.

channel ($>10\ \mu\text{m}$) located at kinks in the pattern. These fractures usually have no impact on the resulting PDMS chip but appear regularly in structure geometries that include sharp edges. It seems likely that these fractures stem from residual stress that appears during the post exposure process, where the structure is subjected to repeated heating and cooling steps (an alternative explanation could lie in similar anisotropic effects that occur during the chemical polymerization process of the exposed photoresist). The thermal properties of a structure depend on its geometry (e.g. local surface to volume ratio). Thus sharp bends in the geometry can lead to local temperature differences during fast heating or cooling of the structure. The resulting uneven thermal expansion or contraction can therefore explain the occurrence of internal stress and ultimately fractures in the material. To avoid such fractures, it is therefore important to slowly ramp hot plate temperature for baking steps. Also, structure designs can be adapted to

include gradual curves instead of sharp bends. Literature values for the coefficient of thermal expansion of SU-8 photoresist are in the range of 50 to $100 \times 10^{-6} \text{ K}^{-1}$ [161] [162] [163], compared to just $3 \times 10^{-6} \text{ K}^{-1}$ for the silicon substrate [164]. This fact hints at temperature gradients being a factor for adhesion issues between SU-8 structures and silicon wafer substrates.

3.3.2 PDMS Chip Integrity

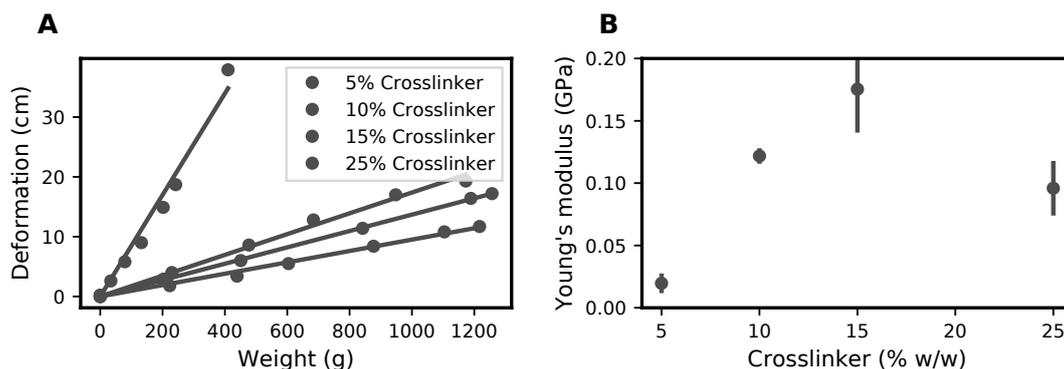


Figure 3.24: Elastic properties of cured PDMS depending on cross-linker concentration. (A) Strips of cured PDMS show linear spring response to variable loads. (B) The Young's modulus of PDMS has an inverse U relationship with the cross-linker concentration. Errorbars (fitting error) magnified 10x.

The following paragraph discusses application specific issues that had to be solved after successful clean room fabrication. When a fabricated SU-8 mold has passed visual inspection under a microscope, a PDMS chip can be cast, processed and bonded to a glass coverslip to yield a finished chip ready for microfluidics experiments. Only during the bonding step issues specific to the design features around sub micron structures were found to occur. Depending on the aspect ratio (height:width) of individual design features, the ceiling of wide and flat channels may sag and inadvertently get bonded to the glass surface, which then constitutes a blockade in the channel. To avoid such defects, the pressure exerted when placing the PDMS chip on the plasma cleaned coverslip should be kept as low as possible. Furthermore the sagging defect clearly depends on the elasticity properties of the PDMS polymer. A range of experiments was conducted to understand the influence of the amount of cross-linker added to the PDMS prior to curing on the elasticity of the material. Strips of cured PDMS were loaded with variable weight and the resulting displacement was recorded (Figure 3.24). The data in subfigure A shows that,

for all cross-linker concentrations, a linear response was observed with a varying spring constants. Taking into account the dimensions of the strip, the Young's modulus can be calculated. With increasing cross-linker concentration, the Young's modulus increases up to at least 15 %. For the highest measured cross-linker concentration of 25 %, the Young's modulus is significantly lower than the maximum value. These results suggest that while cross-linker is responsible for the elastic stiffness of the material, at a certain cross-linker concentration the maximum Young's modulus is reached and any additional cross-linker dilutes the material. Based on this experiment, 15 % cross-linker concentration instead of the default 10 % was used with success to eliminate the bonding defects described above. Generally speaking, knowledge of the properties of the materials involved in microfluidics can be helpful to identify a promising approach for solving material related problems and also to predict the feasibility of new designs with a mechanical model.

3.3.3 Loading Bacteria

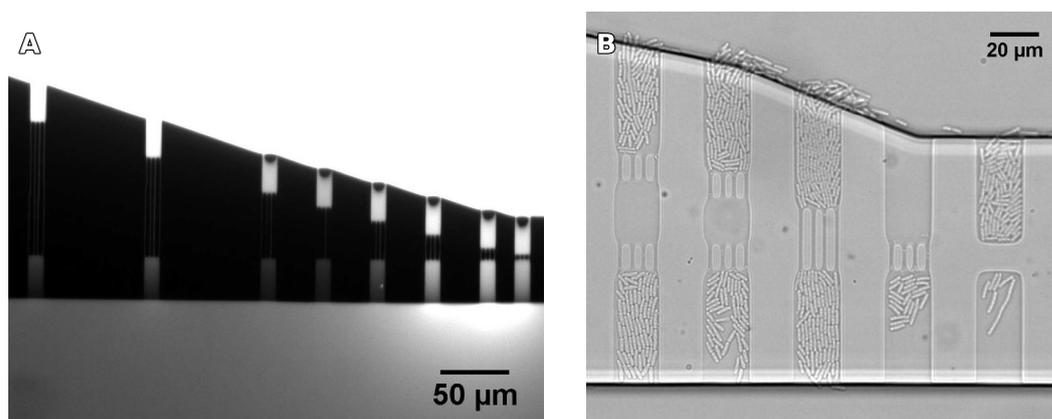


Figure 3.25: Initial tests of microfluidic chips designed to test the communication between connected growth chambers. (A) Fluorescence microscopy image showing diffusion of a dye (Cy5) from the top channel to the bottom channel through growth chambers connected with sub-micron channels of varying length. Only the shortest connections allow significant amounts of dye to pass. Some growth chambers display a collapsed roof inadvertently bonded to the glass substrate. Contrast adjusted to show variation in bottom channel. (B) Bacteria loaded into growth chambers connected with sub-micron channels. The bacterial cells are too large to pass into through the small channels.

For the observation of bacteria in microfluidic chips, the initial loading of bacteria into the chip is a critical step. For the experiments shown in this thesis, the general approach was to load the empty chip with very dense culture (exponential culture concentrated roughly 100-fold), letting cells settle into traps in the chip, then flushing the dense culture

with fresh medium. In the example shown in Figure 3.25 B, 9/10 growth chambers were successfully loaded with bacteria, eventually growing to fill the whole chamber. The cells in the bottom right show an example of filamentous morphology, where despite successful loading the bacteria displayed irregular growth and failed to fill the chamber. The loading of such growth channels and chambers, especially those with dimensions around 1 micron, required specific strategies for acceptable results. Firstly, the loading is very sensitive to the exact dimensions of a growth channel (width and height). While a small overshoot of height or width only results in undesirable growth patterns, undershooting the dimensions of single cells mean that cells are unlikely to fit into a growth channel. The layer height of spin coated photoresist decreases with distance from the center, leading to variations of growth channel height (Figure 3.22 C). These variations in height are sufficient to affect the success of bacteria loading, making it necessary to empirically test loading efficiency for every individual chip that can be produced from a mold. Other dimensional issues make it necessary to return to clean room fabrication with adjusted parameters. The long time to feedback between clean room fabrication to PDMS chip and finally microfluidics experiment can make this iterative process very time consuming. A second point crucial for successful loading is the pressure difference between inlet and outlet of the microfluidic chip. While full pressure equalization is very hard to achieve [154], it helps to limit the pressure differential through careful adjustment of an attached syringe. That way, the flow speed in the main channel is low and the likelihood for cells to enter small channels is higher. In general, highly repeated structures are advisable because it allows for an experiment to proceed even if the loading efficiency is relatively low or other stochastic errors occur during a timelapse.

3.3.4 Timelapse Setup

Timelapse microfluidics experiments, especially if run for more than 24 hours, require a highly stable setup surrounding the microfluidic chip. Critical points that have to be optimized are the stability of microscope focus, the feed medium pump and tubing connecting pump to microfluidic chip. We can use a simple mathematical model to see how stochastic failure events affect experiments that are run for an extended period of time. Assuming constant failure rate λ and an exponentially distributed time to first failure $f(t) = \lambda e^{-\lambda t}$ (waiting time between Poisson process events), we can estimate how the requirements for the failure rate depends on the duration of the experiment. The

probability for a failure to have occurred after time T is given by:

$$F(T) = \int_0^T f(t)dt = 1 - e^{-\lambda T}.$$

We can choose an arbitrary acceptable failure probability for the entire experiment p_a , this gives us a maximum acceptable failure rate per hour dependent on the experiment runtime T :

$$\lambda_a(T) < -\frac{\ln(1 - p_a)}{T}.$$

This simple model illustrates how the requirements for the acceptable failure rate (and thus the stability requirements for the experimental setup) get stricter for longer timelapse experiments but eventually level off as the hyperbola approaches a horizontal asymptote. As an example, this model shows that for a timelapse experiment that is supposed to run for 50 hours with an acceptable failure probability of 50 %, the hourly failure rate needs to be below $1.4\% \text{ h}^{-1}$. Since the single failure rate used in this model in reality represents many concurrent processes with individual failure rates, if long run times are necessary, the complexity of an experiment should be kept as low as possible to make a small failure rate. On the other end, it may be prudent to minimize the timelapse duration required with experiment design or data analysis strategies (e.g. parallel instead of sequential measurements) or by “speeding up” the experiment (e.g. with faster growth of a bacterial system).

3.3.5 Optimization

From an abstract perspective, the establishment of a working fabrication protocol can be understood as an optimization problem for a number of parameters (with varying degrees of dependence between parameters). Like any other experimental procedure, clean room fabrication requires ideas, creative problem solving on one hand and rote parameter optimization and iterations on the other. However in fabrication, the latter part is quite prominent and while it is definitely necessary to bring the appropriate amount patience, it is prudent to choose an approach that limits this part as much as possible. Exhaustive screenings of parameters, sometimes called grid search to denote the regular intervals between tested parameters, scales exponentially with the number of parameters considered. Thus it is only suitable for a small set of parameters independent with the rest of the process. A good example are exposure dose and simple design parameters such

as line width, which can be tested with high throughput in a single fabrication process. Maskless lithography systems can automatically expose a test pattern (containing a range of design parameters to be tested) with a range of exposure parameters (Figure 3.20). Such an approach is less suited for other process parameters such as e.g. baking times, were every parameter test requires individual hands on time. Furthermore it may be prudent to isolate one part of the fabrication process and figure out suitable parameters before moving on to downstream processes. For example, once the fabrication of mechanically stable sub-micron SU-8 channels is solved, the optimization PDMS chips molded off such channels becomes an independent optimization problem. If such channels would have to be optimized for stability and successful PDMS chip casting at the same time, downstream failures would be harder to assign to a choice of parameters and troubleshooting complexity would rise.

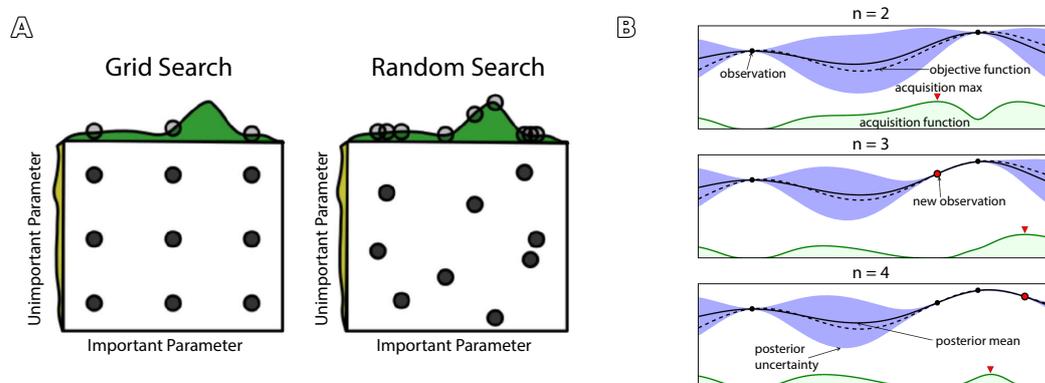


Figure 3.26: Theoretical considerations about optimization strategies for a given objective. (A) Two independent parameters are sampled 9 times. With a grid search strategy, 3 distinct values sampled per parameter is insufficient to resolve the objective landscape. Random search can serve to distinguish the unimportant parameter with a flat objective landscape from the important parameter. Adapted from Bergstra et al. [165]. (B) The unknown 1D objective function (to be optimized, dotted line) is approximated with a probabilistic model (solid line and blue shading). Given the probabilistic model, which is based on a prior distribution and previous observations, the acquisition function evaluates the utility of candidates for the next observation. The acquisition function is elevated where the objective is predicted to be high (exploitation) and where the uncertainty about the objective is high (exploration). The next observation is picked at the maximum of the acquisition function, resulting in an updated model. Adapted from Shahriari et al. [166].

Theoretical considerations suggest that for a complex optimization landscape of many parameters, a random search may be the best strategy [165]. By randomly sampling the parameter space, the number of distinct parameters tested is maximized, allowing for the identification of parameters with little influence for the optimization result (Figure 3.26

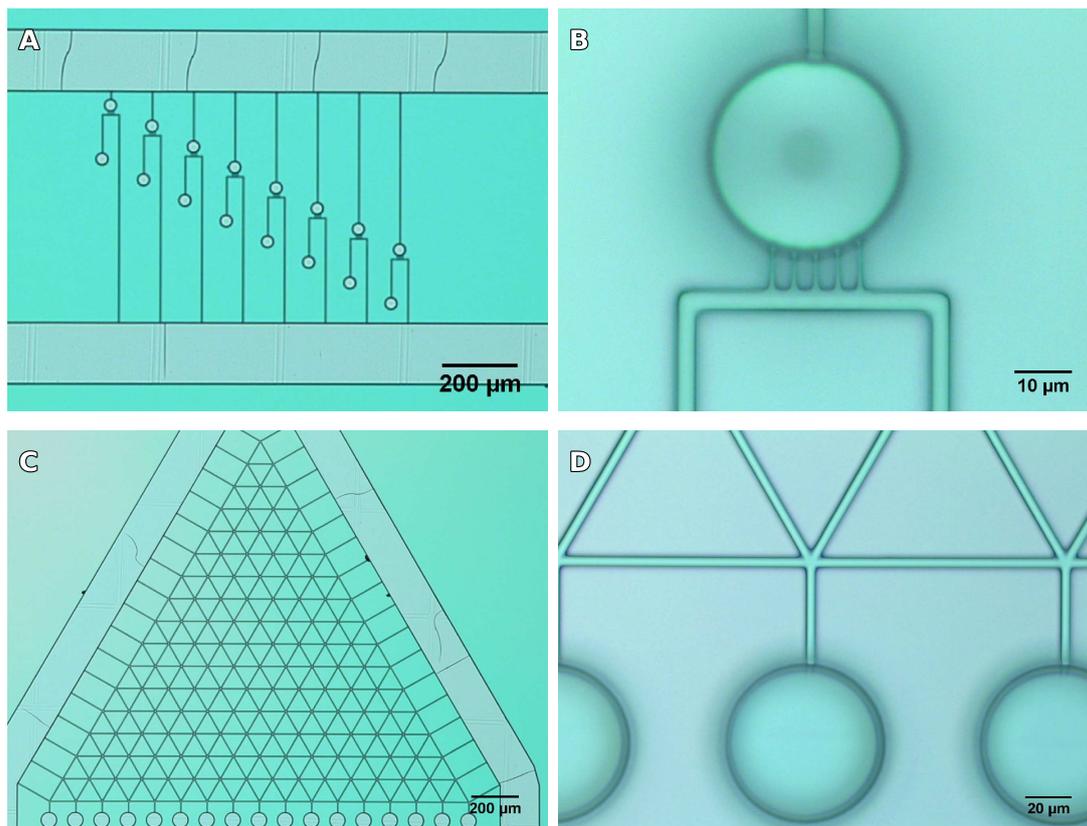


Figure 3.27: *Prototype designs. (A) Two circular reservoirs are each connected to a large channel via a $1\ \mu\text{m}$ access channel with varying length. The top reservoir is connected to the opposite access channel via sub-micron channels, potentially impacting the transmission through the access channel. (B) Closeup of A, showing 5 sub-micron connections. (C) $1\ \mu\text{m}$ channels connected to two large channels and a number of reservoirs in a triangular grid. (D) Closeup of C, showing the accurate grid junctions and connection to circular reservoirs made from a thicker layer.*

A). In the context of fabrication of microfluidics molds, this strategy suggests that naive exploration of complicated processes such as e-beam setup should not be undervalued. An intuitive search approach based less on systematic search but on prior knowledge from experience or literature may be formalized as Bayesian optimization [166] (Figure 3.26 B). In this abstract, mathematical form a search strategy can also guide the design of experiments, making it possible to close a design build test loop with a high degree of automation [167].

3.3.6 Outlook

In summary, the tools established for the fabrication of mother machine devices can be extended to resolutions lower than the dimensions of *E. coli* cells. Through the use of modern maskless technologies with high reproducibility, successful production of a microfluidic chip with a new design can be accomplished in as little as one working day from clean room to finished chip. In contrast to silicon etching based processes for the fabrication of sub-micron structures [168], SU-8 photoresist allows the direct fabrication of a negative mold without difficult etching steps or intermediate molds. With these tools, exciting projects are possible in the future. Prototype designs shown in Figure 3.25 could be utilized to study the transmission of autoinducer signals between colonies separated by sub-micron channels depending on the distance between them. We investigated preliminary ideas about how the passage of bacteria through a channel may be affected by contact with a secondary colony or how bacteria respond to environments with interesting geometry (Figure 3.27).

In the future, the potential for fast and reliable fabrication of design variants could be utilized to integrate a bacterial culture and its spatial environment into a hybrid dynamical system. A new microfluidic chip could be created based on the result of an experiment with the previous pattern and so on. If such a sequence of experiments is regarded as a single, connected, experiment, the modification of the synthetic spatial environment could be regarded as a type of actuation of an *in silico* controller. By adapting the pattern of a microfluidic chip (and rapidly manufacturing the new chip) to the result of an experiment with the previous pattern, the spatial environment can serve as a type of actuation for an *in silico* control scheme (with experiment iterations as discrete steps). A simple example of feedback control would be the adjustment of the width of a nutrient supply channel based on the deviation of the observed growth rate of bacteria from a desired set point. In a similar way, one could control the interaction strength of a connection between two competing or cooperating strains (e.g. predator prey system [51]) cultured in parallel in the same chip and connected via sub-micron channels. A more complex spatial environment like e.g. a maze could be increased or decreased in difficulty depending on whether a bacterial culture can find the exit in the desired time. For this application, the maze pattern could be designed by an automated script with a limited number of input parameters. In a next level, the bacteria closest to the exit of the maze could be extracted and used to inoculate the next round. More generalized, this could be interpreted as a continuous culture with an evolution procedure

towards a complex objective [60, 64]. Such a framework could be studied as an example of a living system under evolutionary pressure interacting with an *in silico* control system.

3.4 Sampling the Distance Dependence of Quorum Sensing Between Bacterial Colonies

A signaling link between bacterial colonies mediated by autoinducer diffusion can provide information about the distance between the colonies. The understanding of synthetic cell to cell communication processes can be a crucial component to engineer living materials with programmable environmental responses [169, 170] or to recreate complex multicellular functions such as morphogenesis [171, 172]. In this series of experiments we want to sample the joint probability distribution of the distance between colonies that can produce or receive autoinducer and the GFP signal triggered by the autoinducer. Bacterial colonies were either a receiver (with GFP reporter) or a sender and receiver based on the Lux quorum sensing system from *Vibrio fischeri*. In order to collect data for a range of different distances between bacterial colonies, we initially tested various experimental strategies to accurately control the distance between colonies. Depending on the effort, different levels of accuracy can be achieved (manual pipetting < pipetting with purpose designed jigs < CNC control with a pipetting robot or a printer with a syringe print head [173]). However, all of these approaches have in common that they are limited by throughput. In this chapter we present an alternative approach: The placement of colonies is left to random chance with a standard plating procedure. A suitable density of colony forming units (CFU) ensures good sampling of the space of relevant distances. The throughput afforded by the high number of colonies that fit a single petri dish allows for the measurement of a statistical distribution in a single experiment. Generally, a sufficiently sampled joint probability distribution allows the computation of information theoretic quantities such as mutual information [104].

Bulk Characterization We constructed a sender and a receiver plasmid, making use of the well established Lux quorum sensing system. The receiver plasmid consists of *LuxR* under a constitutive promoter and a GFP reporter under the p_{Lux} promoter on a high copy number plasmid. In addition to receiver module, the sender plasmid contains *LuxI* under a p_{Rha} promoter, leading to the production of AHL in the presence of Rhamnose (see plasmid maps in Appendix Figure 6.1). Characterization in liquid culture verified the function of the sender and receiver constructs. In both the sender and receiver plasmid, the constitutively expressed LuxR activates the expression of the GFP reporter in the presence of a relatively low concentration of externally added AHL signal (App. Figure 6.2 A, B). By expressing the LuxI protein, controlled by the presence of rhamnose as

inducer, the sender plasmid can also synthesize AHL signal. The process from LuxI expression to AHL synthesis and finally GFP reporter expression introduces a time delay compared to the previous experiments with externally added AHL signal. After adding rhamnose to the culture medium, we measured a lag time of about 3 hours between the start of exponential growth (which normally coincides with reporter production) and the measurement of GFP signal (App. Figure 6.2 C). In a comparable experiment conducted in a bioreactor, we measured a delay of 5 h (see Fig. 3.9 D). In the absence of externally added AHL, the strength of the GFP reporter expression depends on the rhamnose concentration (App. Figure 6.2 D). In all following experiments with the sender construct, 100 mM rhamnose was added to induce the production of AHL signal.

AHL Mixed into Agar Moving on from characterization in bulk medium, control experiments were performed to understand the response of AHL receiver colonies to defined AHL concentrations and gradients. Firstly we plated a strain with the receiver plasmid and a plasmid with a RFP reporter under a constitutive promoter on LB agar plates containing varying concentrations of AHL inducer to verify the bulk experiments in separate colonies. The data (Figure 3.28 A, B) shows that the GFP fluorescence (pLux promoter) has a sigmoidal dependence on the AHL concentration. The highest AHL concentration resulted in zero colonies formed. With a Hill function fit, the K_d can be calculated as 24 nM, about twice as high as in bulk experiments (with a single plasmid). A possible explanation for this higher measured K_d is that the AHL inducer was partially degraded during the production of the agar plates. Meanwhile the RFP fluorescence) decreases with rising AHL concentration. While the GFP transfer function qualitatively behaves as expected from the bulk measurements, the RFP expression would naively not be expected to vary with AHL concentration.

In a next step, we wanted to understand how AHL inducer spreads in spatially varied agar environments. Spatial variation of inducer concentrations was realised with multi layer agar plates (with equal concentrations of nutrients and antibiotics everywhere). For Figure 3.28 C and D, the petri dish was equally divided into a region with (100 nM) and without AHL. For Figure 3.28 E and F, a linear initial gradient of AHL was produced (see Methods). Even spreading of the AHL receiver plus RFP strain on these agar plates resulted in colonies with fluorescence intensities depending on their position. Again, very high AHL levels appear to inhibit colony formation. For the equally divided plate, full induction was observed in colonies 1 cm away from the divider and falling off continuously with further distance. For the linear gradient, induction was observed in areas where the

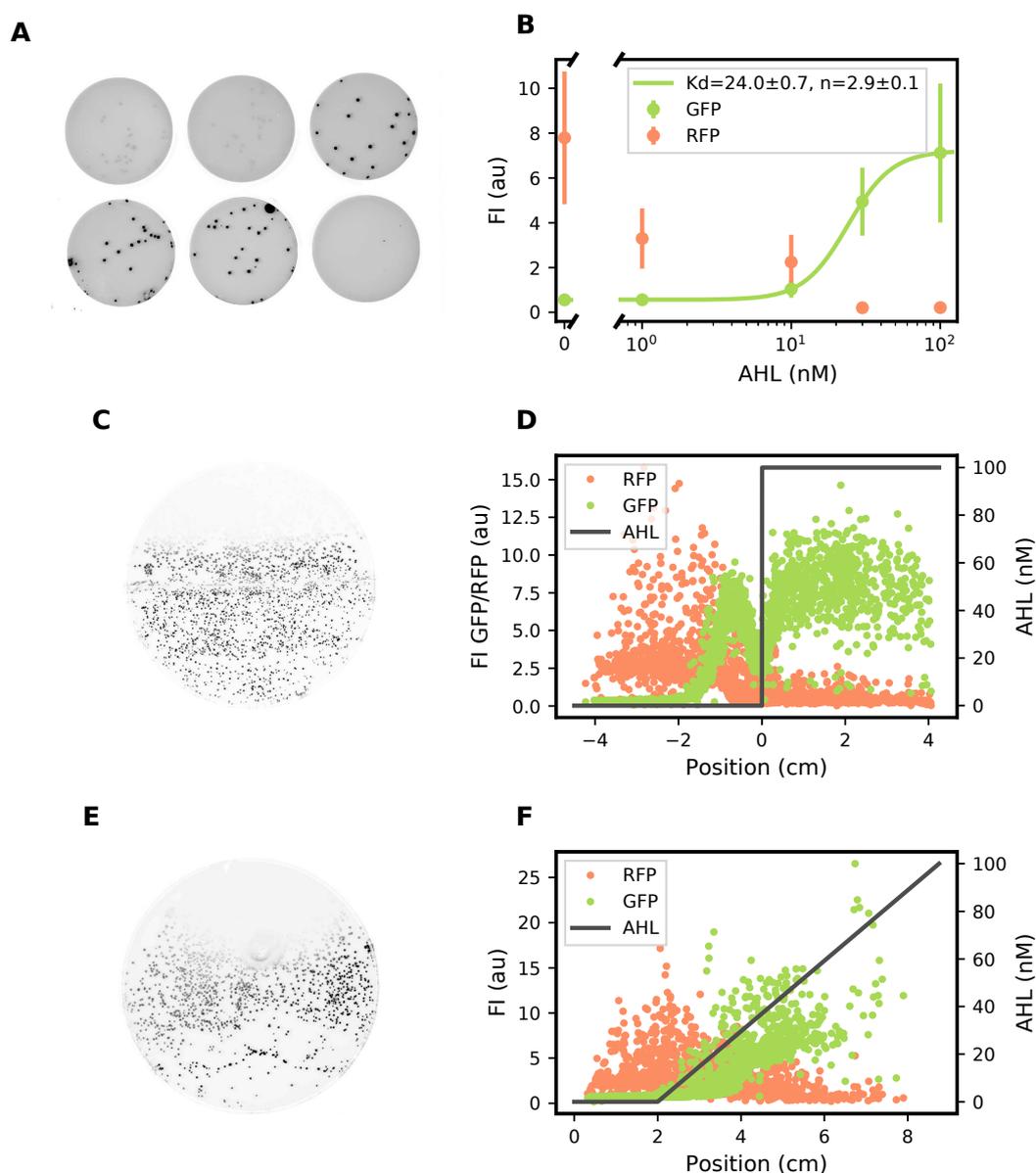


Figure 3.28: AHL receiver plasmid (GFP) and constitutively expressed RFP reporter co-transformed and plated evenly on LB agar plates. (A) GFP expression of colonies on agar plates containing varying AHL concentrations (increasing from top left to bottom right). GFP fluorescence intensity (FI) increases with AHL concentration. No growth is observed for the highest AHL concentration. (B) Analysis of mean GFP and RFP FI of colonies shown in A. GFP follows Hill curve similar to bulk experiment (App. Figure 6.2). RFP produced from a constitutive promoter decreases unexpectedly with increasing AHL concentration. (C) GFP expression of colonies on a multi-layer agar plate: The lower layer of the petri dish is divided in half, only one side with AHL added to the agar (lower half). A top layer of agar (without AHL) is poured over the divider to produce a smooth, even surface. GFP fluorescence can be observed in colonies over the AHL containing half and a limited distance away from the divider. (D) Scatter plot of GFP and RFP fluorescence over position relative to the divider. The black line indicates the initial AHL concentration. FI is lowered around the divider potentially due to irregularities in the agar. Aside from this artifact, GFP FI smoothly declines with distance from the divider. RFP FI (constitutive) is strong in regions with low GFP FI. (E) A ‘wedge’ of AHL containing agar creates a linear (initial) gradient. Fewer colonies grew per area in the region with the highest AHL concentration. (F) GFP FI decreases roughly linearly and goes to zero below an (initial) concentration of 20 nM.

initial AHL level was above a threshold of approximately 20 nM. In general, high variance of the GFP fluorescence intensity was observed between colonies in areas with high AHL inducer concentration. The levels of constitutive RFP reporter in subfigures D and F show an inverse relationship with the GFP fluorescence intensity consistent with the control B.

These results suggest that it is possible to measure the concentration of externally added AHL in agar with AHL receiver colonies. At a sufficiently high concentration, the diffusion of AHL in agar can be observed over a distance of approx. 1 cm. The variation of the constitutive RFP reporter appears to be a consequence of limited cellular resources for endogenous protein expression that are shared between the receiver plasmid and the RFP plasmids.

Interaction of Sender Colonies Supported by the previous experiments with externally added AHL, the transmission of information over space with quorum sensing signals was investigated by random plating of sender colonies. The sender strain containing the AHL sender plasmid and a constitutive RFP label was plated on a LB agar plate without AHL and with 100 mM. Can AHL – produced by the senders – activate GFP expression under a pLux promoter and does the distance between sender colonies affect the strength of the GFP signal? After overnight growth, the position and fluorescence intensity of each individual colony was extracted from a high resolution fluorescence image with a basic image processing script. Figure 3.29 shows a statistical summary of 729 sender colonies randomly spread over a plate. The distribution of the GFP fluorescence intensity of all colonies is asymmetrical with a coefficient of variation (CV) of 27 % (Figure 3.29 A). A histogram of all combinatorial colony to colony distances shows that only a small fraction of these connections are in the 1 cm range for possible interaction through AHL diffusion. Therefore, we will approximate by only considering the distances to the respective nearest neighbouring colony (NN-distance) instead of all possible connections. Does the distance to the nearest neighbour sender colony have an effect on the GFP signal of a colony? A scatter plot between those two variables reveals no correlation between the NN-distance and GFP FI, even though the majority of neighbours are only around 1 mm apart. Moving beyond the nearest neighbour, we considered the diffusion of AHL from every sender colony as point source (2-dimensional gaussian profile) and used this simple model to calculate an uncalibrated AHL concentration at the location of each colony. However, this value also does not show any correlation with the reporter signal (Figure 3.29 D).

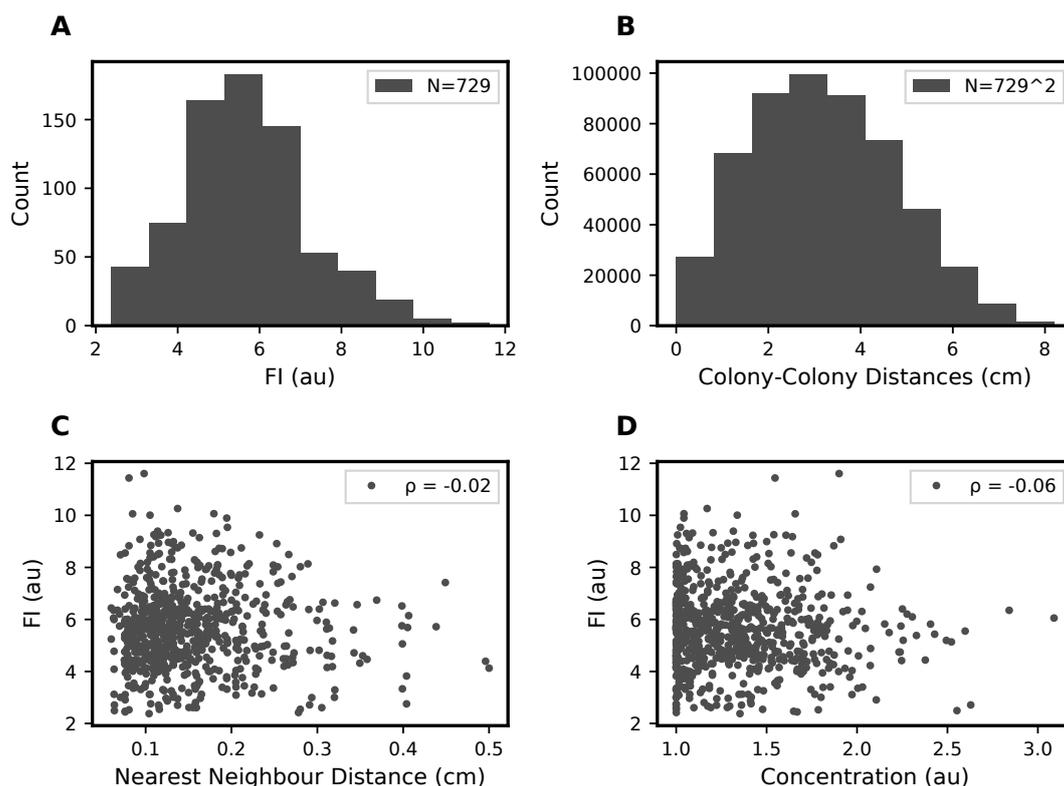


Figure 3.29: Randomly distributed colonies containing AHL sender and constitutive RFP label on an agar plate. Statistical analysis based on the GFP fluorescence intensity (FI) and location of each colony. The data for a replicate of this experiment is shown in Appendix Figure 6.3. (A) Histogram of the GFP FI of 729 sender colonies. (B) A histogram of all combinatorial distances between colonies shows the spatial distribution of the colonies. Only a small fraction of these connections is below a threshold of 1 cm. (C) Scatter plot showing no relationship between the distance of a sender colony to the nearest neighbour sender colony and FI (ρ is the Pearson correlation coefficient). (D) A theoretical AHL concentration of a colony is calculated from the distances to other sender colonies (uncalibrated). A scatter plot shows no relationship to FI.

Interaction of Sender and Receiver Colonies We adapted the experiments by co-plating a sender strain and a receiver strain. Labeling the strains with plasmids for the constitutive expression of RFP and CFP, respectively, allows for the extension of the image processing method to discriminate between colonies belonging to either strain. A 1:1 mix of sender and receiver colonies was prepared. Most receiver colonies show no significant GFP fluorescence intensity (Figure 3.30 A); it should be noted that the arbitrary units between Figure 3.29 and 3.30 are consistent and that the specific filter setup in this experiment causes background fluorescence from dense colonies. This low FI is observed despite close distances in the range of millimeters between receivers and nearby sender colonies.

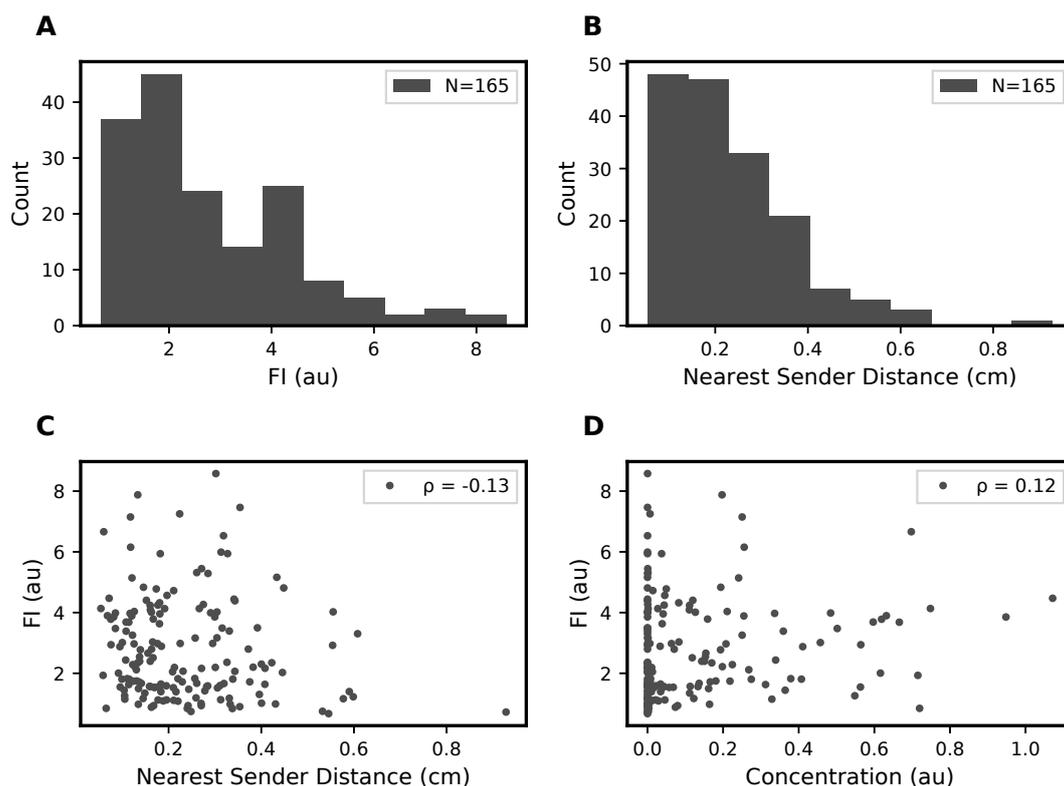


Figure 3.30: Randomly distributed colonies containing either an AHL sender and constitutive CFP label or an AHL receiver and constitutive RFP label on an agar plate. Statistical analysis based on the GFP fluorescence intensity (FI) of receiver colonies and the location of sender colonies. The data for a replicate of this experiment is shown in Appendix Figure 6.4. (A) Histogram of the GFP FI of all 165 receiver colonies. Arbitrary units are consistent with data shown in Fig. 3.29. (B) Histogram showing the distance of the nearest sender colony for every receiver colony. (C) Scatter plot showing no correlation between the distance of the nearest AHL sender. (D) Scatter plot of the theoretical AHL concentration at the location of a receiver colony and the corresponding FI shows no correlation.

Again, the data shows no correlation between minimum distance to a sender colony and fluorescence intensity of receiver colonies (Figure 3.30 B,C). The point source model suggests that despite small minimum distances between receiver and sender colonies, the theoretical local concentration of AHL is close to zero at the location of most receiver colonies. In this experiments, distances under 1 mm could only be observed in receiver colonies that grew directly next to sender colonies. While the data from non-separated colonies is somewhat ambiguous, a fluorescence intensity gradient could be observed through the diameter of those receiver colonies, hinting at an effective induction range under 1 mm.

In conclusion, experiments based on the analysis of individual colonies is a useful approach to achieve moderately high throughput. The variations in spatial context generated by the random spatial distribution of colonies allows insight beyond mean values into the statistical properties of populations. However, the limitations stemming from the size of individual, separate colonies prevent a sampling of distances that is useful to investigate the spatial information transmission by quorum sensing communication. Unless the AHL sensitivity of the reporter can be greatly increased, it is necessary to extend the principle to smaller scales, e.g. in a microfluidic chip environment [174]. To fully understand the spatial distribution of AHL in the different experiments shown here, a more sophisticated model is needed. A simple 2D view of the surface or 1D profiles and distance measures are simplifications that may be inadequate. A better model would be described by a reaction diffusion equation with source and decay terms. In the case of sender colonies, they can be modelled as a continuously producing point source, whereas in the case of AHL mixed into agar, the agar can be treated as a constant reservoir of AHL [175, 176]. This distinction makes it clear that these two cases are very different and the observed distance relationship with AHL receivers cannot be directly compared, which is confirmed by the divergent experimental results.

4 Conclusion

In this work, we developed a bacterial oscillator that integrates the programmable repressor dCas9. In a next step, dCas9 can be used to couple the oscillator to native systems of the bacterial host and thus control its temporal behaviour. We observed a dynamic interdependence of the state of the genetic circuit and the growth rate of the host, which in turn is known to influence the dynamics of the circuit. Based on the analysis of the stochastic variability of the circuit, we present a quantitative upper bound for the fidelity of the oscillator working as a synthetic clock to organize secondary processes. For the design and improvements of future dynamic synthetic biological systems, these information theoretic quantities may be a useful optimization goal.

We also showed that CRISPR interference can be used to couple the growth of bacteria to environmental conditions by inducing amino acid auxotrophy. In conjunction with the established continuous culture techniques, this may be used as selection principle to tune the function of dynamic CRISPRi-based genetic circuits.

The rapid development of CRISPR technology in recent years opens up exciting opportunities to improve and extend the functionality of the presented systems. For example, the programming of multiple targets simultaneously becomes possible with multiplexing strategies [177, 178], dynamic repression/derepression can be achieved with competing guide RNAs [179] and negative feedback stabilizes the availability of unbound dCas9 [180].

We established fabrication techniques for sub-micron microfluidics chips that enable single cell experiments. The reliable and rapid manufacturing of these chips opens up the possibility of a fabrication-experiment-analysis loop that offers new insight into the interaction of engineered cells with a complex spatial environment. Furthermore, the coordination of bacterial colonies in space with cell-to-cell communication was investigated.

In the future, engineered cells with CRISPR modules may be integrated into more complex environments such as microbial communities, in which they could perform the function of a biological clock or react to the state of the environment with programmed logic. For in vivo applications outside of defined lab conditions, the principles of growth and evolution need to be considered and integrated into the design process. With the societal challenges of our time in view, we can achieve sustainable technological progress by harnessing the ‘advanced nanotechnology’ found in biology.

5 Methods

5.1 Bacterial Oscillator Experiments

Cloning The two plasmids used in this section were constructed using standard PCR, restriction ligation and Gibson assembly [181]. DNA sequences are listed in the SI. The pRPC-oscillator plasmid is based on pZS1-ITlrLLtCL [3] (Addgene #26489) with the dCas9 gene taken from pdCas9-bacteria [9] (Addgene #44249). The pRPC-sponge plasmid is based on pLPT41 [65]. Cloning steps were performed in *DH5 α* .

Bacterial Cell Culture Bacterial cultures (K-12 wild type strain *MG1655* containing the plasmids described above) were grown overnight in LB medium, diluted 1 to 100 into 10 ml LB medium and grown for another 3 hours (37 °C, 250 rpm). For loading of the microfluidic device, the cultures were spun down (5000 rcf for 5 min) and resuspended in 150 μ l flow medium (LB medium supplemented with 1.1 ‰ (w/v) bovine serum albumin (BSA) for passivation). All medium was supplemented with carbenicillin (50 μ g/ml) and chloramphenicol (12.5 μ g/ml) antibiotics. If specified, the inducers isopropyl β -D-1-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc) were added (500 μ M and 214 nM, respectively).

Microfluidics To construct the microfluidic devices, PDMS and curing agent (Dow Corning, Sylgard 184) were mixed (1:10) and poured on the master wafer in an aluminium foil container, degassed for 30 min and baked for 80 min at 80 °C. The cured PDMS was detached from the master wafer and individual chips were cut out. The channel side of the chip was immediately covered with Scotch magic tape to keep it clean. Inlet and outlet holes were punched with a biopsy puncher. Cover slips (170 μ m thickness, Carl Roth, LH26.1) were cleaned with Hellmanex III (2 ‰, Hellma) and rinsed with ddH₂O. PDMS chips were bonded to the cover slips after O₂ plasma treatment (0.7 min, Diener Femto) and baked for 60 min at 80°C to complete the bonding. The complete microfluidic devices were loaded with highly concentrated bacterial culture with a syringe through a PTFE tube (0.8 mm inner diameter, Bola, S1810-10) attached to the outlet. The loading was observed under a microscope and the pressure on the syringe adjusted so that the flow of the culture stopped. The bacteria were left to settle for 30 min (at 37 °C). After that the loading syringe was cut off, the free end of the tube connected to the outlet was placed into a tube for waste collection and the dense loading culture was flushed out with flow medium from a 20 ml syringe through a second silicone tube attached to the inlet. A continuous flow of flow medium was set up with a syringe pump (0.5 ml/h, TSE Systems) for the duration of the time lapse. Syringes were replaced as

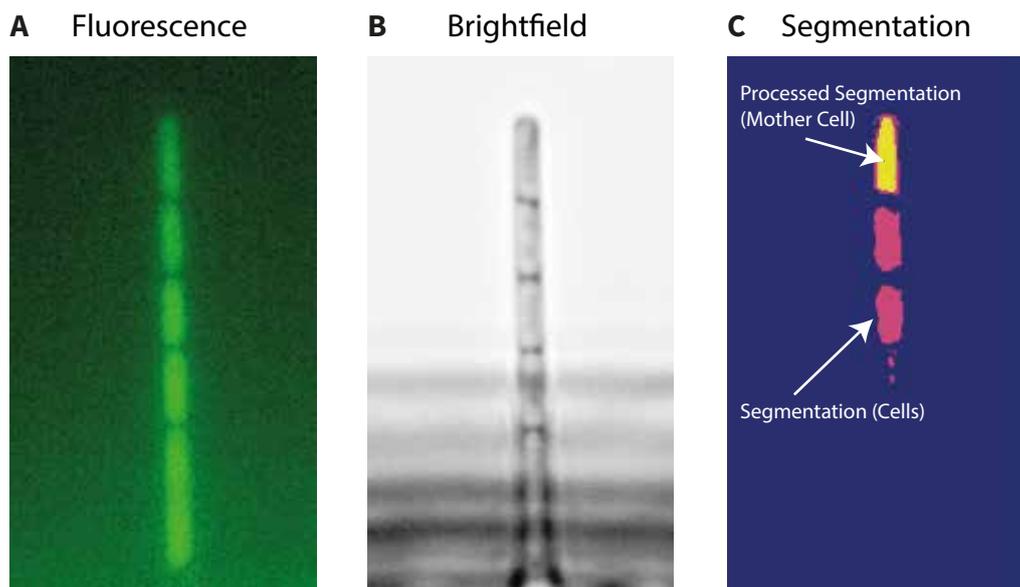


Figure 5.1: Sketch of the segmentation process used to extract single cell data from microscopy videos. The brightfield channel (middle) is used to create a segmentation mask corresponding to cells using *ilastik* (cells in the lower part of the growth channel are ignored). The initial segmentation is processed to refine the mask and detect the mother cell. The resulting mother cell binary mask is applied to the fluorescence intensity channel and values for the average over the cell area are extracted for each frame.

needed taking care not to introduce air bubbles. In the case that a buildup of bacteria stuck in the feed channel was observed in the feed channel during a timelapse, a manual flush was performed periodically.

Microscopy Time lapse videos were recorded with a Nikon Ti-2E microscope using a 60x plan apochromat oil objective (NA 1.40), 1.5x magnification, a SOLA SM II LED light source, Andor NEO 5.5 camera and NIS elements software. The temperature in the microscope enclosure was controlled at 37 °C, the microfluidic chip and tubing were fixed in place with custom 3D printed clamps for optimal stability. The Nikon Perfect Focus System was used to account for focus drift. For the timelapse, up to 40 different positions were imaged in phase contrast brightfield and GFP fluorescence channels for 200 ms every 10 minutes for up to 70 hours.

Data Analysis Timelapse videos were split per imaged position and cropped to selected growth channels with custom *imageJ* [182] macros (selection criterium: successful loading

and continuous growth observed the brightfield channel over the course of the timelapse). A segmentation mask for trapped cells was created with the software *ilastik* [183] (pixel classification using a random forest classifier, all input features selected). Training of the classifier was done by specifying the class (cell or background) of sample pixels. To achieve good results, examples of all relevant image features were included in the training, while taking care not to overtrain the classifier on specific examples. All following data analysis steps were done with Python (Version 3.7, Python Software Foundation) using the *numpy*, *scipy* and *pandas* libraries (Figure 5.1). Segmentation masks were cleaned up with binary morphological operations, connected pixels corresponding to the mother cell were determined by position in the channel, fluorescence intensity values were averaged over mother cell pixels and corrected with reference to the should-be constant value of PDMS background. In cases where the mother cell stopped dividing during the timelapse, the first daughter cell below the mother cell was used as substitute. The instantaneous growth rate was calculated by detecting cell division events and differentiating smoothed (Savitzky-Golay filter) cumulative division event data. To calculate the instantaneous phase, the phase angle of the analytical representation of the fluorescence intensity data was calculated and unwrapped. This (complex) analytical signal was constructed with the smoothed fluorescence intensity signal as real part and the Hilbert transform thereof as imaginary part [130, 184]:

$$\zeta(t) = s(t) + i s_H(t).$$

The Hilbert transform of a real signal is defined by:

$$s_H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau,$$

where the integral is understood as its Cauchy principal value. The analytical signal has a positive spectrum with no negative frequency component, and it can be represented by complex vectors rotating in one direction. The instantaneous phase is then readily available as the angle of the analytical signal: $\phi(t) = \arg[\zeta(t)]$

The concept of mutual information is based on Shannon's information entropy [102]:

$$S[P(x)] = - \sum_{x_i} P(x_i) \log_2 P(x_i),$$

which measures the uncertainty about a random variable $x \in \{x_i\}$ with probability distribution $P(x_i)$. The mutual information quantifies the statistical dependence between

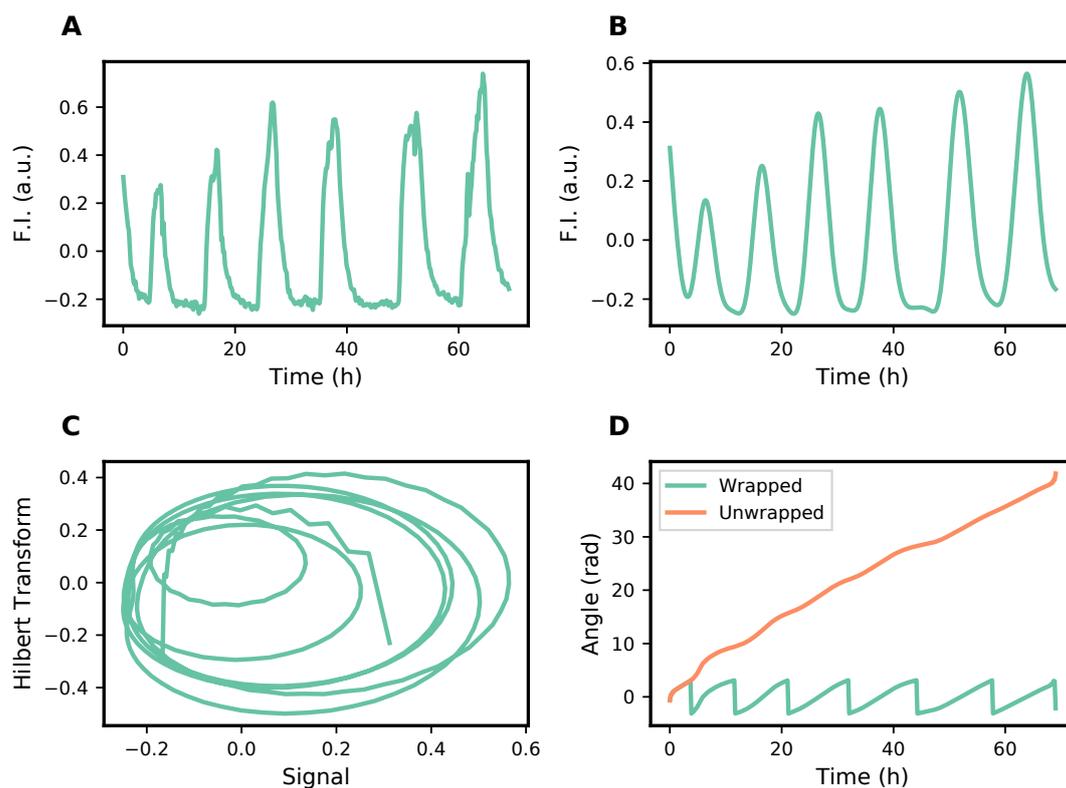


Figure 5.2: Illustration of the calculation of the instantaneous phase. The oscillating signal (A) is smoothed (B) with a Savitzky-Golay filter. The Hilbert transform is used to construct a complex analytical signal. A visualisation in the complex plane (C) shows the circular trace of the analytical signal vector. The instantaneous phase is the angle of the complex vector (D). The phase angle ($\text{mod } 2\pi$) can be unwrapped to measure the progression over multiple periods.

two random variables and their associated probability distributions and thus the amount of information that can be gained about one variable by observing the other, which can be expressed as:

$$I(t; \phi) = S[P_\phi(\phi)] - \langle S[P(\phi|t)] \rangle_t.$$

Written in this way, it can be interpreted as the information contained in the dynamic range of the measured signal minus the average information entropy of the variability of the signal at fixed time points, i.e., noise. To estimate a value for the mutual information from our experimental data, we used the *second Gaussian approximation* described by Tkačik et al. [104]. Briefly, the conditional probability distributions $P(\phi|t)$ were approximated as Gaussian and the marginal distribution $P_\phi(\phi)$ is obtained by integrating the conditional prob. dist. over time. The effect of the limited number of samples is dealt

with by extrapolating to infinite samples from randomly sampled subsets of the available data (cf. SI Fig. S16).

The Cramér-Rao bound [185] gives a local value about the error of the time estimate from knowledge of the phase (compare to the global value of the mutual information). It states:

$$\sigma_t(t) \geq \frac{1}{\sqrt{\mathcal{I}(t)}},$$

with the Fisher information $\mathcal{I}(t)$ (for a Gaussian $P(\phi|t)$) given by

$$\mathcal{I}(t) = \left(\frac{\partial_t \langle \phi(t) \rangle}{\sigma_\phi(t)} \right)^2 + 2 \left(\frac{\partial_t \sigma_\phi(t)}{\sigma_\phi(t)} \right)^2.$$

The phase averaged over the set of cells $\langle \phi(t) \rangle$ and the corresponding standard deviation σ_ϕ was calculated directly from the phase data (see above); the derivatives were smoothed (Savitzky-Golay filter).

The standard error of the mean for values that correspond to peak locations (ACF or histogram) was estimated from a bootstrapped sample distribution (random resampling of the data with replacement, 1000 times).

The source code used for data analysis and simulation is available on GitHub:
<https://github.com/JohnHenningsen/CRISPRi-Oscillator>

5.2 Lithography for Microfluidics

Mother Machine Fabrication Using EBL A master mold for the ‘mother machine’ microfluidic device was fabricated in a cleanroom by lithography. Four layers of SU-8-2000 negative photoresist (Microchem) with varying thickness (corresponding to the desired channel height) were applied and exposed consecutively on a 2 inch silicon wafer (Siegert Wafer): Firstly, a base layer (roughly 1 μm height) helps the adhesion of fine structures in subsequent layers. A second layer with alignment markers is used as position reference for the subsequent two layers. The growth channels (20-30 μm length) are patterned in thin third layer (1 μm height). Feed channels (150 μm width) are patterned in a fourth layer (20 μm height). In the case of the third layer, the resist was crosslinked using electron-beam lithography (Raith eLINE, 30 kV acceleration voltage, 10 μm aperture). The other three layers were UV-crosslinked in a mask aligner (Süss MicroTec, 200W lamp) with photomasks (Zitzmann). The lithography process parameters (dilutions, spin, bake, development and exposure parameters) that were found to produce these

specifications are given in a detailed protocol in the Appendix.

Maskless Lithography of Sub-Micron Channels The processing steps for maskless exposure based lithography match the steps described in the previous paragraph for the most part, except: Layers were exposed using a MLA 100 maskless aligner (Heidelberg Instruments). The dose and defocus parameters were adjusted for every structure based on the results of an automated parameter screen. Alignment markers were exposed together with the thin layer containing all the small features.

Analysis Results of the fabrication process were analysed with a standard upright microscope (Leica) and a Dektak 150 profilometer (Veeco).

PDMS stress A 2.4 mm high layer of PDMS with varying crosslinker concentration casted in a petri dish and baked at 80 °C for 80 min. 4 cm × 1 cm strips were cut out of the solidified layer. The strips were fixed vertically in a clamp on one side and loaded with variable weight attached to a clamp on the other side. The resulting extension of the strip was measured with a calliper.

5.3 Auxotrophy Experiments

Cloning For this section, the pdCas9-bacteria (Addgene #44249) and pgRNA-bacteria (Addgene #44251)[9] were used with modified 20-bp sgRNA target regions (overhang PCR [120]). The promoter of dCas9 was modified to reduce leaky expression. Theophylline aptamers were adapted from Kundert (2019) [134].

Growth Medium For auxotrophy growth experiments, two different minimal media were used: M9 medium (1× M9 salts, 2 mM MgSO₄, 0.1 mM CaCl₂) with glucose as carbon source (default 10 mM) or Neidhardt Rich Defined Medium (RDM) [149] using a modified ‘EZ recipe’ [150] (all components unchanged except the amino acid supplement, which was prepared from purified powder stocks (Sigma, Roth) leaving out individual amino acids as needed).

Bioreactor Chemostat For all experiments the *E. coli* strain MG1655 was used, containing the AHL receiver or sender plasmids (Chapter 3.4). Unless noted otherwise, cultures were inoculated with 1 mL overnight culture into LB medium with standard antibiotics concentrations. Cultures were grown in a Minifors 2 (Infors HT) bioreactor with 600 mL

vessel and pH (Easy Ferm Plus Arc 325) and pO_2 (VisiFerm DO Arc 325). Standard settings were a 500 min^{-1} stirring rate, 2 L min^{-1} air influx and 37°C temperature. If active, the pH was adjusted by automatic addition of 1 M NaOH or $3 \text{ N H}_3\text{PO}_4$. Anti foam agent SE 15 was added as needed. For a chemostat setup, a pump integrated into the bioreactor was set according to the dilution rate. An outflow tube was connected to a faster pump and regulated by setting the inlet of the tube at the desired fill height. Fluorescence intensity was measured in a CLARIOstar plate reader (BMG Labtech) [186].

Pipetting Robot Chemostat A pipetting robot (Opentrons OT-2) was set up for continuous culture with a 1 mL pipette, a magnetic stirrer hot plate (IKA), a 50 mL flask and 500 mL beakers as reservoirs. The hot plate was adjusted to achieve a temperature of 37°C in the flask filled with medium. 20 mL culture was agitated with a magnetic stir bar. Cultures were inoculated from overnight cultures into LB medium with standard antibiotics concentrations. The pipetting robot was scripted to remove samples, add medium and remove waste with the integrated pipette. Pipette tips were changed only after waste removal and sampling steps to minimize tip consumption (Figure 3.10 B). Samples were fixated as detailed in Figure 3.11.

Flow Cytometry Flow cytometry measurements were conducted with a FACS Melody device (BD Biosciences) with a 488 nm Laser. The turbidity of samples was measured in a photometer (IMPLEN) using 1 mL cuvettes, accordingly samples were diluted into filtered PBS to achieve an event rate of 1000 s^{-1} . The side scatter (SSC) channel was gated to filter out background events. Data was analysed using FloJo V10 software with a ellipsoid gate to isolate single cells based on forward and side scatter channels.

5.4 QS experiments

Cloning The AHL receiver and sender plasmids used in this section were constructed ‘de novo’ from BioBrick parts using standard molecular cloning techniques with synthesized gene fragments (IDT gBlock).

Agar Plate Production Agar plates containing AHL reservoirs (100 nM) were produced in petri dishes (two compartment or standard), all containing LB medium and 1.5% agar-agar. For the agar plates equally divided into two halves, agar with or without AHL was poured into each halve respectively. Once hardened, a top layer covering the whole

dish, containing no AHL, was poured to form a smooth surface. To form an agar plate with a linear AHL gradient, agar containing AHL was left to solidify with the petri dish set at an incline, resulting in a straight diagonal profile in the cross section view (wedge shape). Once solidified, agar without AHL was poured on the now level dish to fill out the dish to a smooth flat surface.

Random Plating Bacterial culture (pre-cultured over night) was plated on a agar plates in a suitable concentration for order of magnitude 100 colony forming units (CFU). The commonly used rule of thumb that liquid culture with an OD600 of 1 corresponds to 8×10^8 cells was used as starting point, adjusted with an empirical factor of 0.25 CFU per cell. The liquid culture in exponential phase was thus serial diluted accordingly before plating with beads.

Analysis After overnight growth, colonies on the respective agar plates were imaged with a Typhoon FLA 9500 laser scanner (GE Healthcare) with channels suitable for green and red fluorescent proteins. In the GFP channel, background fluorescence allows for the location of colonies that produce no fluorescent protein. The fluorescence intensity and location of individual colonies was extracted with a custom image processing script (Python).

6 Appendix

6.1 Appendix Figures

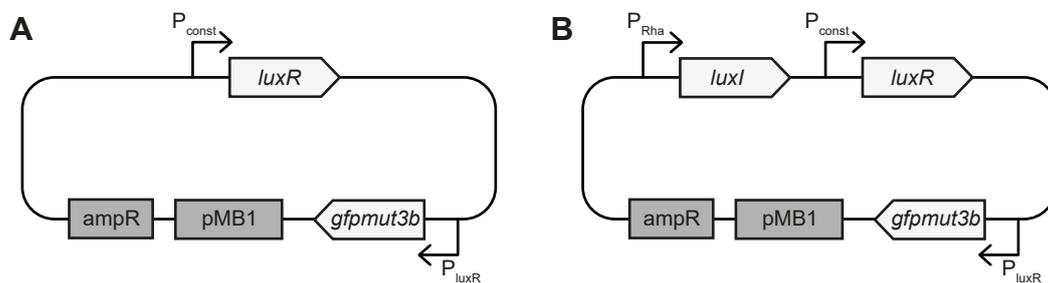


Figure 6.1: Plasmid maps of receiver and sender plasmids.

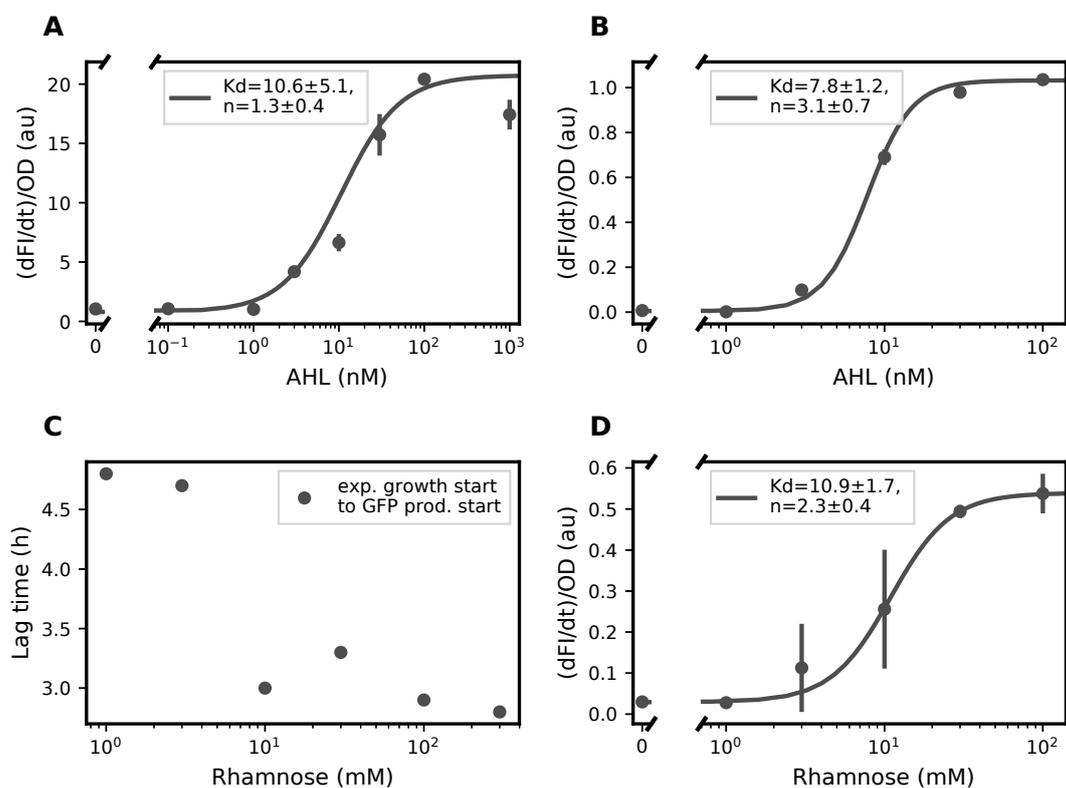


Figure 6.2: Bulk characterization of AHL receiver and sender systems. (A) GFP production rate by the receiver plasmid in response to a titration of AHL inducer, fitted with a Hill function. (B) GFP production rate by the sender plasmid in response to a titration of AHL inducer, fitted with a Hill function. (C) Lag time between the start of exponential growth and the beginning of GFP production by the sender plasmid, depending on the concentration of rhamnose (which controls the synthesis of AHL). (D) GFP production rate (after a lag time) by the sender plasmid in response to a titration of rhamnose inducer, fitted with a Hill function.

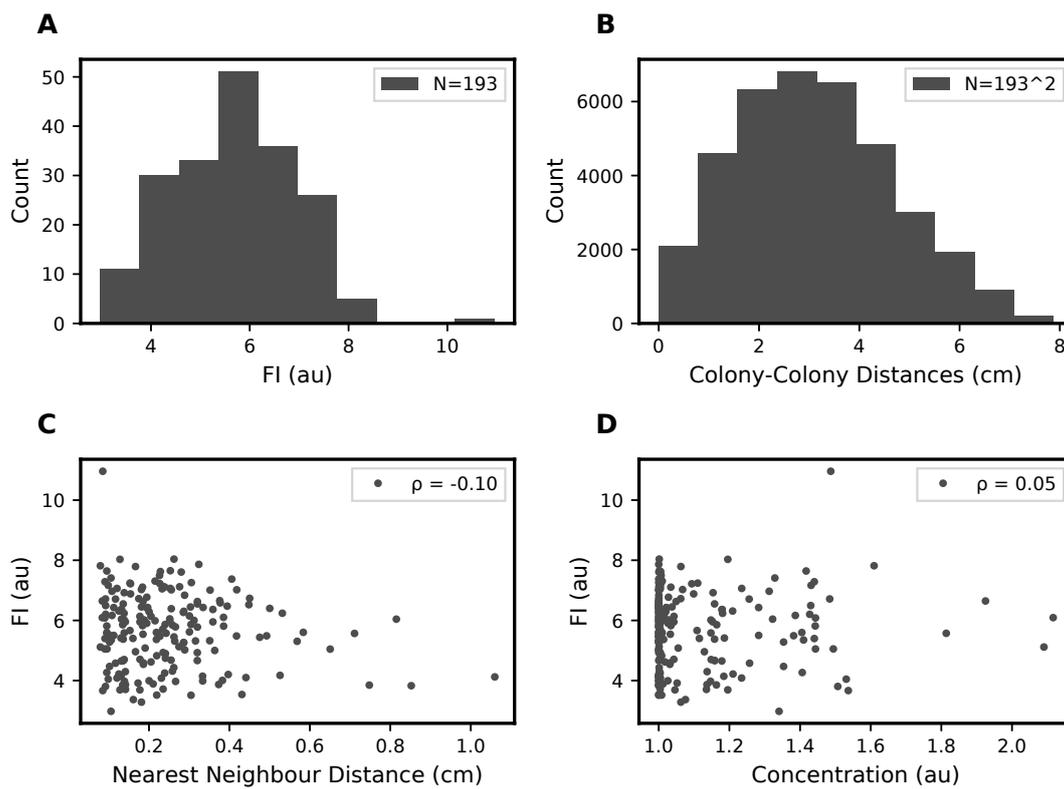


Figure 6.3: Replicate of Figure 3.29, whose caption likewise describes this figure.

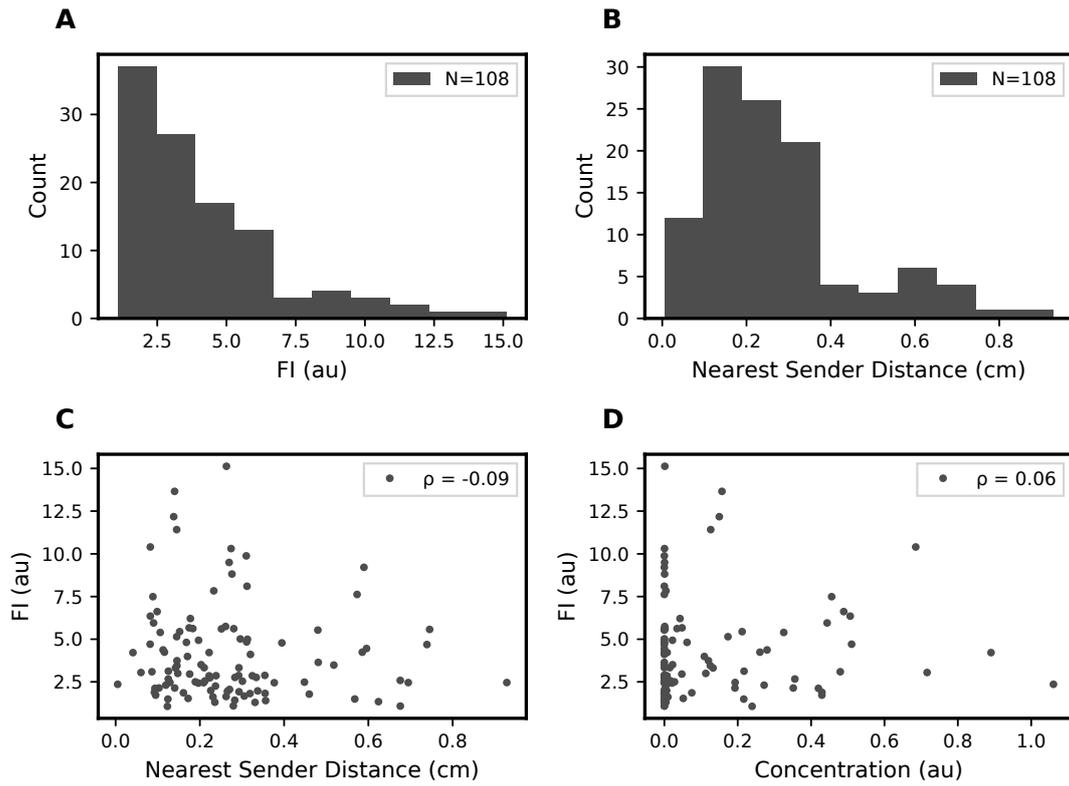


Figure 6.4: Replicate of Figure 3.30, whose caption likewise describes this figure.

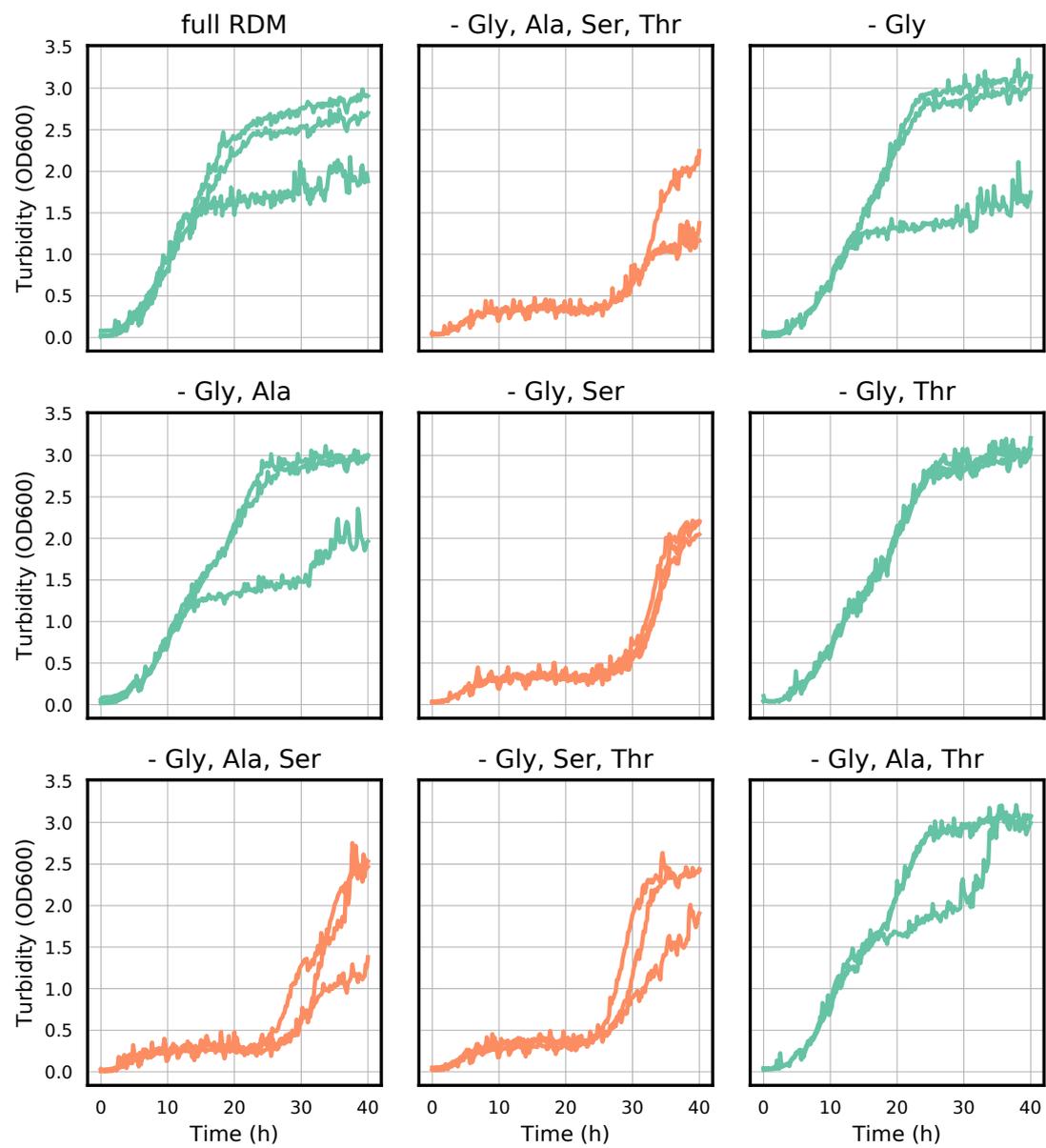


Figure 6.5: See Figure 3.16.

6.2 Appendix Notes

Linear Stability Analysis This is an overview of linear stability analysis of a 2 dimensional system [187]. By analysing a linear approximation of the nonlinear system (can be extended to n dimensions), the stability of the fixed points can be determined with the following procedure. First, we write down a general form of the non-linear equations:

$$\begin{aligned}\dot{X} &= f(X, Y) \\ \dot{Y} &= g(X, Y).\end{aligned}$$

Let's consider a linear approximation around the fixed point (X^*, Y^*) and substitute variables X' and Y' :

$$\begin{aligned}f(X^*, Y^*) &= g(X^*, Y^*) = 0 \\ U &= X - X^*, \quad V = Y - Y^*.\end{aligned}$$

Now, we use the Taylor expansion of functions f and g to the first order. Written in vector form the differential equations become:

$$\begin{aligned}\begin{pmatrix} \dot{U} \\ \dot{V} \end{pmatrix} &= \begin{pmatrix} \partial_X f & \partial_Y f \\ \partial_X g & \partial_Y g \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} \\ J &= \begin{pmatrix} \partial_X f & \partial_Y f \\ \partial_X g & \partial_Y g \end{pmatrix},\end{aligned}$$

with ∂_X denoting the partial derivative with respect to the subscript variable evaluated at the fixed point and the Jacobian matrix J . To find the solutions to this general linear ODE system, we need to diagonalize the Jacobian. Then, the solution is given by exponential functions with the (complex) eigenvalues λ as exponents. For our 2 dimensional system, this becomes:

$$\begin{pmatrix} U \\ V \end{pmatrix}(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t},$$

with integration constants C_1 and C_2 . Using Euler's formula to rearrange the exponential $e^{\lambda t} = e^{\text{Re}(\lambda)t} (\cos(\text{Im}(\lambda)t) + i \sin(\text{Im}(\lambda)t))$ it becomes clear that the signs of the real and imaginary part of the eigenvalue determine the qualitative dynamics of the solution. If

the imaginary part is non zero, the solution describes periodic oscillations around the fixed point. This is the case when the discriminant of the characteristic polynomial is negative:

$$\delta = (\partial_X f - \partial_Y g)^2 + 4\partial_X g \partial_Y f < 0. \quad (6.1)$$

The real part describes whether the solution collapses into the fixed point or diverges out of the regime where the linear approximation holds. So a negative real part describes a spiral that exponentially decays and eventually stabilizes. Only with a positive real part sustained oscillations are possible. This is the case when the trace of the Jacobian is positive:

$$\text{tr}(J) = \partial_X f + \partial_Y g > 0.$$

Master Fabrication Protocol

- 4 layer lithography
- SU-8 2000 resist
- 2 inch wafer

Base Layer:

- Do not clean wafer
- Spin coat 3ml SU-8 2000.5
 - 500/100/10 (rpm speed/(rpm/sec) acceleration/sec duration)
 - 3000/300/60
 - Observe regular circular interference pattern, otherwise discard wafer
- Soft bake 1 min @ 95 °C (hotplate)
- UV exposure 15 sec (200 W mask aligner, soft contact) with no photo mask
- Post exposure bake 1 min @ 95 °C

Alignment Marker:

- Spin coat 3ml SU-8 2002
 - 500/100/10

– 3000/300/60

- Soft bake 3 min @ 95 C
- UV exposure 15 sec 30 sec (200 W mask aligner, soft contact) with alignment marker photo mask
- Post exposure bake 3 min @ 95 °C
- Develop for 30 sec, gentle agitation in SU-8 developer
- Rinse wafer for 10 sec with fresh SU-8 developer, 10 sec with isopropyl alcohol. Dry with pressured air.

Growth Channels:

- Spin coat 3ml SU-8 2000.5
 - 500/100/10
 - 3000/300/60
- Soft bake
 - 1 min @ 65 °C (65 °C steps optional)
 - 3 min @ 95 °C
 - 1 min @ 65 °C
- e-beam exposure with eLINE, take care about the following:
 - loadlock procedure
 - dose (exposure screen necessary)
 - alignment and rotation
 - focus (use scratch on wafer), stigmatation, aperture parameters
 - writefield alignment
 - structure design
- Post exposure bake
 - 1 min @ 65 °C
 - 3 min @ 95 °C

– 1 min @ 65 °C

- Develop for 30 sec, gentle agitation in SU-8 developer
- Rinse wafer for 10 sec with fresh SU-8 developer, 10 sec with isopropyl alcohol. Dry with pressured air.
- Hard bake 15 min @ 150 °C optional

Feed Channels:

- Spin coat 3ml SU-8 2025
 - 500/100/10
 - 5000/300/60
- Soft bake
 - 1 min @ 65 °C
 - 4 min @ 95 °C
 - 1 min @ 65 °C
- Locally develop alignment markers by cleaning resist with swab soaked with SU-8 developer. Mark the locations of alignment markers before spin coating last layer if alignment markers are not visible.
- UV exposure 45 sec (200 W mask aligner, hard contact) with feed channel photo mask aligned to the alignment markers on the wafer. Alignment critical for correct growth channel length.
- Post exposure bake
 - 1 min @ 65 °C
 - 4 min @ 95 °C
 - 1 min @ 65 °C
- Develop for 90 sec, gentle agitation in SU-8 developer
- Rinse wafer for 10 sec with fresh SU-8 developer, 10 sec with isopropyl alcohol. Dry with pressured air.
- Hard bake 15 min @ 150 °C

Bibliography

1. Kwok, R. Five hard truths for synthetic biology. *Nature News* **463**, 288–290 (2010).
2. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
3. Elowitz, M. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
4. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nature Reviews Microbiology* **12**, 381–390 (2014).
5. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010–2020. *Nature Communications* **11**, 1–4 (2020).
6. Brophy, J. A. & Voigt, C. A. Principles of genetic circuit design. *Nature Methods* **11**, 508–520 (2014).
7. Nielsen, A. A., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross, D., Densmore, D. & Voigt, C. A. Genetic circuit design automation. *Science* **352** (2016).
8. Voigt, C. A. Synthetic biology 2020–2030: Six commercially-available products that are changing our world. *Nature Communications* **11**, 1–6 (2020).
9. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. & Lim, W. A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
10. Nielsen, A. A. & Voigt, C. A. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Molecular Systems Biology* **10**, 763 (2014).
11. Santos-Moreno, J. & Schaerli, Y. CRISPR-based gene expression control for synthetic gene circuits. *Biochemical Society Transactions* **48**, 1979–1993 (2020).
12. Gallup, O., Ming, H. & Ellis, T. Ten future challenges for synthetic biology. *Engineering Biology* **5**, 51–59 (2021).

13. Tsoi, R., Wu, F., Zhang, C., Bewick, S., Karig, D. & You, L. Metabolic division of labor in microbial systems. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 2526–2531 (2018).
14. El Karoui, M., Hoyos-Flight, M. & Fletcher, L. Future trends in synthetic biology—a report. *Frontiers in Bioengineering and Biotechnology* **7**, 175 (2019).
15. Lee, J. W., Chan, C. T., Slomovic, S. & Collins, J. J. Next-generation biocontainment systems for engineered organisms. *Nature Chemical Biology* **14**, 530–537 (2018).
16. Monod, J. The growth of bacterial cultures. *Annual Review of Microbiology* **3**, 371–394 (1949).
17. Murray, J. D. *Mathematical biology I. An introduction* 3rd ed. (Springer, New York, 2002).
18. Monod, J. La technique de culture continue: Theorie et applications. *Annales de l'Institut Pasteur* **79**, 390–410 (1950).
19. Novick, A. & Szilard, L. Description of the chemostat. *Science* **112**, 715–716 (1950).
20. Herbert, D., Elsworth, R. & Telling, R. The continuous culture of bacteria; a theoretical and experimental study. *Microbiology* **14**, 601–622 (1956).
21. Smith, H. L. & Waltman, P. *The theory of the chemostat: Dynamics of microbial competition* (Cambridge University Press, 1995).
22. Miller, A. W., Befort, C., Kerr, E. O. & Dunham, M. J. Design and use of multiplexed chemostat arrays. *Journal of Visualized Experiments: JoVE* (2013).
23. Takahashi, C. N., Miller, A. W., Ekness, F., Dunham, M. J. & Klavins, E. A low cost, customizable turbidostat for use in synthetic circuit characterization. *ACS Synthetic Biology* **4**, 32–38 (2015).
24. Wong, B. G., Mancuso, C. P., Kiriakov, S., Bashor, C. J. & Khalil, A. S. Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nature Biotechnology* **36**, 614–623 (2018).
25. Steel, H., Habgood, R., Kelly, C. L. & Papachristodoulou, A. In situ characterisation and manipulation of biological systems with Chi. Bio. *PLOS Biology* **18**, e3000794 (2020).
26. Hol, F. J. & Dekker, C. Zooming in to see the bigger picture: microfluidic and nanofabrication tools to study bacteria. *Science* **346** (2014).

27. Rusconi, R., Garren, M. & Stocker, R. Microfluidics expanding the frontiers of microbial ecology. *Annual Review of Biophysics* **43**, 65–91 (2014).
28. Whitesides, G. M., Ostuni, E., Takayama, S., Jiang, X. & Ingber, D. E. Soft lithography in biology and biochemistry. *Annual Review of Biomedical Engineering* **3**, 335–373 (2001).
29. Squires, T. M. & Quake, S. R. Microfluidics: Fluid physics at the nanoliter scale. *Reviews of Modern Physics* **77**, 977 (2005).
30. Weibel, D. B., DiLuzio, W. R. & Whitesides, G. M. Microfabrication meets microbiology. *Nature Reviews Microbiology* **5**, 209–218 (2007).
31. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625 (2004).
32. Cookson, S., Ostroff, N., Pang, W. L., Volfson, D. & Hasty, J. Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Molecular Systems Biology* **1**, 2005.0024 (2005).
33. Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A. & Jun, S. Robust growth of *Escherichia coli*. *Current Biology* **20**, 1099–1103 (2010).
34. Long, Z., Nugent, E., Javer, A., Cicuta, P., Sclavi, B., Lagomarsino, M. C. & Dorfman, K. D. Microfluidic chemostat for measuring single cell dynamics in bacteria. *Lab on a Chip* **13**, 947–954 (2013).
35. Balagaddé, F. K., You, L., Hansen, C. L., Arnold, F. H. & Quake, S. R. Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science* **309**, 137–140 (2005).
36. Topiwala, H. & Hamer, G. Effect of wall growth in steady-state continuous cultures. *Biotechnology and Bioengineering* **13**, 919–922 (1971).
37. Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R. & Lappin-Scott, H. M. Microbial biofilms. *Annual Review of Microbiology* **49**, 711–745 (1995).
38. Maxwell, J. C. On governors. *Proceedings of the Royal Society of London*, 270–283 (1868).
39. Wiener, N. *Cybernetics or control and communication in the animal and the machine* 2nd ed. (MIT Press, Cambridge, MA, 1948).
40. Del Vecchio, D., Dy, A. J. & Qian, Y. Control theory meets synthetic biology. *Journal of The Royal Society Interface* **13**, 20160380 (2016).

41. Hsiao, V., Swaminathan, A. & Murray, R. M. Control theory for synthetic biology: recent advances in system characterization, control design, and controller implementation for synthetic biology. *IEEE Control Systems Magazine* **38**, 32–62 (2018).
42. Del Vecchio, D., Qian, Y., Murray, R. M. & Sontag, E. D. Future systems and control research in synthetic biology. *Annual Reviews in Control* **45**, 5–17 (2018).
43. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
44. Austin, D., Allen, M., McCollum, J., Dar, R., Wilgus, J., Sayler, G., Samatova, N., Cox, C. & Simpson, M. Gene network shaping of inherent noise spectra. *Nature* **439**, 608–611 (2006).
45. Bassler, B. L. & Losick, R. Bacterially speaking. *Cell* **125**, 237–246 (2006).
46. Boo, A., Amaro, R. L. & Stan, G.-B. Quorum sensing in synthetic biology: a review. *Current Opinion in Systems Biology*, 100378 (2021).
47. Grant, P. K., Dalchau, N., Brown, J. R., Federici, F., Rudge, T. J., Yordanov, B., Patange, O., Phillips, A. & Haseloff, J. Orthogonal intercellular signaling for programmed spatial behavior. *Molecular Systems Biology* **12**, 849 (2016).
48. Scott, S. R. & Hasty, J. Quorum sensing communication modules for microbial consortia. *ACS Synthetic Biology* **5**, 969–977 (2016).
49. You, L., Cox, R. S., Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
50. Miano, A., Liao, M. J. & Hasty, J. Inducible cell-to-cell signaling for tunable dynamics in microbial communities. *Nature Communications* **11**, 1–8 (2020).
51. Balagaddé, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake, S. R. & You, L. A synthetic *Escherichia coli* predator–prey ecosystem. *Molecular Systems Biology* **4**, 187 (2008).
52. Scott, S. R., Din, M. O., Bittihn, P., Xiong, L., Tsimring, L. S. & Hasty, J. A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. *Nature Microbiology* **2** (2017).
53. Din, M. O., Martin, A., Razinkov, I., Csicsery, N. & Hasty, J. Interfacing gene circuits with microelectronics through engineered population dynamics. *Science Advances* **6**, 1–8 (2020).

54. Miliadis-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I., Pincus, D., El-Samad, H., Khammash, M. & Lygeros, J. In silico feedback for in vivo regulation of a gene expression circuit. *Nature Biotechnology* **29**, 1114–1116 (2011).
55. Uhlendorf, J., Miermont, A., Delaveau, T., Charvin, G., Fages, F., Bottani, S., Batt, G. & Hersen, P. Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 14271–14276 (2012).
56. Del Vecchio, D., Ninfa, A. J. & Sontag, E. D. Modular cell biology: Retroactivity and insulation. *Molecular Systems Biology* **4**, 161 (2008).
57. Dobzhansky, T. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* **35**, 125–129 (1973).
58. Castle, S. D., Grierson, C. S. & Gorochoowski, T. E. Towards an engineering theory of evolution. *Nature Communications* **12**, 1–12 (2021).
59. Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E. & Kim, J. F. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
60. Liu, W., Cremer, J., Li, D., Hwa, T. & Liu, C. An evolutionarily stable strategy to colonize spatially extended habitats. *Nature* **575**, 664–668 (2019).
61. Henningsen, J. *Synthetic control of bacteria chemotaxis* MA thesis (Technische Universität München, 2017).
62. Arnold, F. H. Design by directed evolution. *Accounts of Chemical Research* **31**, 125–131 (1998).
63. Lezia, A., Csicsery, N. & Hasty, J. Design, Mutate, Screen: High-throughput creation of genetic clocks with different period-amplitude characteristics. *bioRxiv* (2021).
64. Luro, S., Potvin-Trottier, L., Okumus, B. & Paulsson, J. Isolating live cells after high-throughput, long-term, time-lapse microscopy. *Nature Methods* **17**, 93–100 (2020).
65. Potvin-Trottier, L., Lord, N. D., Vinnicombe, G. & Paulsson, J. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**, 514–517 (2016).

66. Sleight, S. C., Bartley, B. A., Lieviant, J. A. & Sauro, H. M. Designing and engineering evolutionary robust genetic circuits. *Journal of Biological Engineering* **4**, 1–20 (2010).
67. Gouda, M. K., Manhart, M. & Balázsi, G. Evolutionary regain of lost gene circuit function. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 25162–25171 (2019).
68. Klumpp, S., Zhang, Z. & Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**, 1366–1375 (2009).
69. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
70. Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Current Opinion in Biotechnology* **22**, 559–565 (2011).
71. Riglar, D. T. et al. Bacterial variability in the mammalian gut captured by a single-cell synthetic oscillator. *Nature Communications* **10**, 1–12 (2019).
72. Ceroni, F. et al. Burden-driven feedback control of gene expression. *Nature Methods* **15**, 387–393 (2018).
73. Qian, Y., Huang, H.-H., Jiménez, J. I. & Del Vecchio, D. Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology* **6**, 1263–1272 (2017).
74. McBride, C. D. & Del Vecchio, D. Predicting composition of genetic circuits with resource competition: demand and sensitivity. *bioRxiv* (2021).
75. Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D. & Voigt, C. A. A ‘resource allocator’ for transcription based on a highly fragmented T7 RNA polymerase. *Molecular Systems Biology* **10**, 742 (2014).
76. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
77. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics* **6**, 451–464 (2005).
78. Limpert, E., Stahel, W. A. & Abbt, M. Log-normal distributions across the sciences. *BioScience* **51**, 341–352 (2001).

79. Beal, J. Biochemical complexity drives log-normal variation in genetic expression. *Engineering Biology* **1**, 55–60 (2017).
80. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology—identifying causes of failure of synthetic biological systems. *Biotechnology Journal* **7**, 856–866 (2012).
81. Ben-Jacob, E. in *Systems Biology* 25–35 (Springer, 2009).
82. Alon, U. *An introduction to systems biology: design principles of biological circuits* (CRC Press, 2019).
83. Hill, A. V. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *The Journal of Physiology* **40**, i–vii (1910).
84. Forger, D. *Biological clocks, rhythms, and oscillations: the theory of biological timekeeping* (MIT Press, Cambridge, Massachusetts, 2017).
85. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
86. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics* **31**, 64–68 (2002).
87. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11980–11985 (2003).
88. Mangan, S., Zaslaver, A. & Alon, U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *Journal of Molecular Biology* **334**, 197–204 (2003).
89. Novák, B. & Tyson, J. J. Design principles of biochemical oscillators. *Nature Reviews. Molecular Cell Biology* **9**, 981–991 (2008).
90. Geva-Zatorsky, N., Dekel, E., Batchelor, E., Lahav, G. & Alon, U. Fourier analysis and systems identification of the p53 feedback loop. *Proceedings of the National Academy of Sciences* **107**, 13550–13555 (2010).
91. Gallego, M. & Virshup, D. M. Post-translational modifications regulate the ticking of the circadian clock. *Nature Reviews. Molecular Cell Biology* **8**, 139–148 (2007).

92. Goodwin, B. C. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation* **3**, 425–437 (1965).
93. Fraser, A. & Tiwari, J. Genetical feedback-repression: II. Cyclic genetic systems. *Journal of Theoretical Biology* **47**, 397–412 (1974).
94. Müller, S., Hofbauer, J., Endler, L., Flamm, C., Widder, S. & Schuster, P. A generalized model of the repressilator. *Journal of Mathematical Biology* **53**, 905–937 (2006).
95. MacKay, D. J. C. *Information theory, inference and learning algorithms* (Cambridge University Press, 2003).
96. Bialek, W. *Biophysics: searching for principles* (Princeton University Press, 2012).
97. Tkačik, G. & Walczak, A. M. Information transmission in genetic regulatory networks: a review. *Journal of Physics: Condensed Matter* **23**, 153102 (2011).
98. Rivoire, O. & Leibler, S. The value of information for populations in varying environments. *Journal of Statistical Physics* **142**, 1124–1166 (2011).
99. Tkačik, G. & Bialek, W. Information processing in living systems. *Annual Review of Condensed Matter Physics* **7**, 89–117 (2016).
100. Schrödinger, E. *What is life? The physical aspect of the living Cell* (Cambridge University Press, 1944).
101. Tribus, M. & McIrvine, E. C. Energy and information. *Scientific American* **225**, 179–190 (1971).
102. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
103. Slonim, N., Atwal, G. S., Tkacik, G. & Bialek, W. Estimating mutual information and multi-information in large networks. *arXiv preprint cs/0502017* (2005).
104. Tkačik, G., Dubuis, J. O., Petkova, M. D. & Gregor, T. Positional information, positional error, and readout precision in morphogenesis: a mathematical framework. *Genetics* **199**, 39–59 (2015).
105. Cover, T. M. *Elements of information theory* (John Wiley & Sons, 1999).
106. Jaynes, E. T. Information theory and statistical mechanics. *Physical Review* **106**, 620 (1957).

107. Henningsen, J., Schwarz-Schilling, M., Leibl, A., Gutiérrez, J., Sagredo, S. & Simmel, F. C. Single cell characterization of a synthetic bacterial clock with a hybrid feedback loop containing dCas9-sgRNA. *ACS Synthetic Biology* **9**, 3377–3387 (2020).
108. Cheng, A. A. & Lu, T. K. Synthetic biology: an emerging engineering discipline. *Annual Review of Biomedical Engineering* **14**, 155–178 (2012).
109. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113**, 597–607 (2003).
110. Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S. & Hasty, J. A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 5 (2008).
111. Tigges, M., Marquez-Lago, T. T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312 (2009).
112. Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M. & Maerkl, S. J. Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* **4**, e09771 (2015).
113. Goldbeter, A. & Berridge, M. J. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour* (Cambridge University Press, 1996).
114. Winfree, A. T. *The geometry of biological time* (Springer Verlag New York, New York, 2001).
115. Loose, M., Kruse, K. & Schwille, P. Protein self-organization: Lessons from the min system. *Annual Review of Biomedical Engineering* **40**, 315–336 (2011).
116. Tyson, J. J., Csikasz-Nagy, A. & Novak, B. The dynamics of cell cycle regulation. *BioEssays* **24**, 1095–1109 (2002).
117. Ferrell, J. E., Tsai, T. Y.-C. & Yang, Q. Modeling the cell cycle: Why do certain circuits oscillate? *Cell* **144**, 874–885 (2011).
118. Roenneberg, T. & Mrosovsky, M. Circadian clocks - the fall and rise of physiology. *Nature Reviews. Molecular Cell Biology* **6**, 965–971 (2005).
119. Din, M. O. et al. Synchronized cycles of bacterial lysis for in vivo delivery. *Nature* **536**, 1–12 (2016).

120. Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S. & Qi, L. S. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8**, 2180–2196 (2013).
121. Santos-Moreno, J., Tasiudi, E., Stelling, J. & Schaerli, Y. Multistable and dynamic CRISPRi-based synthetic circuits. *Nature Communications* **11**, 2746 (2020).
122. Kuo, J., Yuan, R., Sánchez, C., Paulsson, J. & Silver, P. A. Toward a translationally independent RNA-based synthetic oscillator using deactivated CRISPR-Cas. *Nucleic Acids Research* **48**, 8165–8177 (2020).
123. Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B. & Cho, B.-K. High-level dCas9 expression induces abnormal cell morphology in *Escherichia coli*. *ACS Synthetic Biology* **7**, 1085–1094 (2018).
124. Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V. & Bikard, D. A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9. *Nature Communications*, 1–10 (2018).
125. Yang, D., Jennings, A. D., Borrego, E., Retterer, S. T. & Männik, J. Analysis of factors limiting bacterial growth in PDMS mother machine devices. *Frontiers in Microbiology* **9**, 1–12 (2018).
126. Whinn, K. S. et al. Nuclease dead Cas9 is a programmable roadblock for DNA replication. *Scientific Reports* **9**, 13292 (2019).
127. Friedman, N., Cai, L. & Xie, X. S. Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. *Physical Review Letters* **97**, 168302 (16 2006).
128. Boyle, E. A., Andreasson, J. O. L., Chircus, L. M., Sternberg, S. H., Wu, M. J., Guegler, C. K., Doudna, J. A. & Greenleaf, W. J. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 5461–5466 (2017).
129. Jones, D. L., Leroy, P., Unoson, C., Fange, D., Curic, V., Lawson, M. J. & Elf, J. Kinetics of dCas9 target search in *Escherichia coli*. *Science* **357**, 1420–1424 (2017).
130. Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: a universal concept in nonlinear sciences* (Cambridge University Press, 2001).

131. Zhang, S. & Voigt, C. A. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Research* **46**, 11115–11125 (2018).
132. Clamons, S. & Murray, R. Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking. *bioRxiv* (2019).
133. Jayanthi, S., Nilgiriwala, K. S. & Del Vecchio, D. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology* **2**, 431–441 (2013).
134. Kundert, K. et al. Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs. *Nature Communications* **10**, 2127 (2019).
135. Liu, Y., Wan, X. & Wang, B. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nature Communications* **10**, 3693 (2019).
136. Ho, H.-I., Fang, J. R., Cheung, J. & Wang, H. H. Programmable CRISPR-Cas transcriptional activation in bacteria. *Molecular Systems Biology* **16**, 7:1–12 (2020).
137. Wu, F., Shim, J., Gong, T. & Tan, C. Orthogonal tuning of gene expression noise using CRISPR–Cas. *Nucleic Acids Research* **13**, 497–10 (2020).
138. Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C. & Xing, X.-H. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nature Communications* **9**, 1–15 (2018).
139. Chan, C. T., Lee, J. W., Cameron, D. E., Bashor, C. J. & Collins, J. J. 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology* **12**, 82–86 (2016).
140. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, 2006–0008 (2006).
141. Vidal, L., Pinsach, J., Striedner, G., Caminal, G. & Ferrer, P. Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in Escherichia coli. *Journal of Biotechnology* **134**, 127–136 (2008).
142. Tepper, N. & Shlomi, T. Computational design of auxotrophy-dependent microbial biosensors for combinatorial metabolic engineering experiments. *PLOS ONE* **6**, 1–11 (2011).

143. Bertels, F., Merker, H. & Kost, C. Design and characterization of auxotrophy-based amino acid biosensors. *PLOS ONE* **7**, 1–8 (2012).
144. Kunjapur, A. M. et al. Synthetic auxotrophy remains stable after continuous evolution and in coculture with mammalian cells. *Science Advances* **7**, eabf5851 (2021).
145. Sandberg, T. E., Salazar, M. J., Weng, L. L., Palsson, B. O. & Feist, A. M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metabolic Engineering* **56**, 1–16 (2019).
146. Tjørve, K. M. & Tjørve, E. The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. *PLOS ONE* **12**, e0178691 (2017).
147. Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A. & Alon, U. Hierarchy of non-glucose sugars in Escherichia coli. *BMC Systems Biology*, 1–12 (2014).
148. Coli Genetic Stock Center. *List of Commonly Requested Auxotrophic Strains* Accessed: 2021-10-18. 2010. <https://cgsc.biology.yale.edu/Auxotrophs.php>.
149. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture medium for enterobacteria. *Journal of Bacteriology* **119**, 736–747 (1974).
150. E. coli Genome Project at UW-Madison. *EZ Rich Defined Medium* Accessed: 2021-10-18. 2003. <https://www.genome.wisc.edu/resources/protocols/ezmedium.htm>.
151. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Research* **46**, D633–D639 (2017).
152. Mondragon-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L. & Hasty, J. Entrainment of a population of synthetic genetic oscillators. *Science* **333**, 1315–1319 (2011).
153. Männik, J., Driessen, R., Galajda, P., Keymer, J. E. & Dekker, C. Bacterial growth and motility in sub-micron constrictions. *Proceedings of the National Academy of Sciences* **106**, 14861–14866 (2009).
154. Phan, T. V., Morris, R., Black, M. E., Do, T. K., Lin, K.-C., Nagy, K., Sturm, J. C., Bos, J. & Austin, R. H. Bacterial route finding and collective escape in mazes and fractals. *Physical Review X* **10**, 31017 (2020).

155. Tokárová, V. et al. Patterns of bacterial motility in microfluidics-confining environments. *Proceedings of the National Academy of Sciences of the United States of America* **118** (2021).
156. Gupta, S., Ross, T. D., Gomez, M. M., Grant, J. L., Romero, P. A. & Venturelli, O. S. Investigating the dynamics of microbial consortia in spatially structured environments. *Nature Communications* **11**, 1–15 (2020).
157. Hol, F. J., Galajda, P., Nagy, K., Woolthuis, R. G., Dekker, C. & Keymer, J. E. Spatial structure facilitates cooperation in a social dilemma: empirical evidence from a bacterial community. *PLOS ONE* **8**, e77042 (2013).
158. Westheimer, G. Hyperacuity. *Scholarpedia* **6**, 9973 (2011).
159. Lorenz, H., Despont, M., Fahrni, N., LaBianca, N., Renaud, P. & Vettiger, P. SU-8: a low-cost negative resist for MEMS. *Journal of Micromechanics and Microengineering* **7**, 121 (1997).
160. Vinje, J., Beckwith, K. S. & Sikorski, P. Electron beam lithography fabrication of SU-8 polymer structures for cell studies. *Journal of Microelectromechanical Systems* **29**, 160–169 (2020).
161. Chollet, F. *SU-8: Thick Photo-Resist for MEMS* Accessed: 2021-10-18. 2020. <http://memscyclopedia.org/su8.html>.
162. Lorenz, H., Laudon, M. & Renaud, P. Mechanical characterization of a new high-aspect-ratio near UV-photoresist. *Microelectronic Engineering* **41**, 371–374 (1998).
163. Feng, R. & Farris, R. J. The characterization of thermal and elastic constants for an epoxy photoresist SU8 coating. *Journal of Materials Science* **37**, 4793–4799 (2002).
164. Becker, P., Scyfried, P. & Siegert, H. The lattice parameter of highly pure silicon single crystals. *Zeitschrift fur Physik B Condensed Matter* **48**, 17–21 (1982).
165. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13** (2012).
166. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* **104**, 148–175 (2016).
167. Burger, B. et al. A mobile robotic chemist. *Nature* **583**, 237–241 (2020).

168. Moolman, M. C., Huang, Z., Krishnan, S. T., Kerssemakers, J. W. & Dekker, N. H. Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**, 1–10 (2013).
169. Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., Voigt, C. A. & Ellington, A. D. A synthetic genetic edge detection program. *Cell* **137**, 1272–1281 (2009).
170. Gilbert, C. & Ellis, T. Biological engineered living materials: growing functional materials with genetically programmable properties. *ACS Synthetic Biology* **8**, 1–15 (2018).
171. Teague, B. P., Guye, P. & Weiss, R. Synthetic morphogenesis. *Cold Spring Harbor Perspectives in Biology* **8**, a023929 (2016).
172. Santos-Moreno, J. & Schaeferli, Y. Using synthetic biology to Engineer spatial patterns. *Advanced Biosystems* **3**, 1800280 (2019).
173. Choi, W. S., Ha, D., Park, S. & Kim, T. Synthetic multicellular cell-to-cell communication in inkjet printed bacterial cell systems. *Biomaterials* **32**, 2500–2507 (2011).
174. Ramalho, T., Meyer, A., Mückl, A., Kapsner, K., Gerland, U. & Simmel, F. C. Single Cell Analysis of a Bacterial Sender-Receiver System. *PLOS ONE* **11**, e0145829 (2016).
175. Hense, B. A., Müller, J., Kuttler, C. & Hartmann, A. Spatial heterogeneity of autoinducer regulation systems. *Sensors* **12**, 4156–4171 (2012).
176. Dilanji, G. E., Langebrake, J. B., De Leenheer, P. & Hagen, S. J. Quorum activation at a distance: spatiotemporal patterns of gene regulation from diffusion of an autoinducer signal. *Journal of the American Chemical Society* **134**, 5618–5626 (2012).
177. Reis, A. C., Halper, S. M., Vezeau, G. E., Cetnar, D. P., Hossain, A., Clauer, P. R. & Salis, H. M. Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nature Biotechnology* **37**, 1294–1301 (2019).
178. McCarty, N. S., Graham, A. E., Studená, L. & Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nature Communications* **11**, 1–13 (2020).
179. Anderson, D. A. & Voigt, C. A. Competitive dCas9 binding as a mechanism for transcriptional control. *Molecular Systems Biology* **17**, e10512 (2021).

180. Huang, H.-H., Bellato, M., Qian, Y., Cárdenas, P., Pasotti, L., Magni, P. & Del Vecchio, D. dCas9 regulator to neutralize competition in CRISPRi circuits. *Nature Communications* **12**, 1–7 (2021).
181. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A. & Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
182. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).
183. Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nature Methods* **16**, 1226–1232 (2019).
184. Gabor, D. Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers Part 3* **93**, 429–441 (1946).
185. Cramér, H. *Mathematical methods of statistics* (Princeton University Press, 1999).
186. Mühlhofer, M.-T. *Establishment of Cultivation Strategies for the Investigation of Synthetic Genetic Circuits in the Bioreactor* Bachelor's Thesis (Technische Universität München, 2018).
187. Strogatz, S. H. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (Westview Press, 1994).

Acknowledgements

I would like to extend my deepest gratitude to all the people that made this thesis possible. I realize that for me, the decision to pursue this work and the ability to follow through with it is was contingent on a strong support network, long term financial stability and the privileged learning environment throughout my life.

Firstly, Fritz gave me the opportunity to work as a scientist and was the source for never ending inspiration and knowledge. Thank you for fostering a culture of intellectual freedom and mutual responsibility that allowed me to develop my ideas and to find my path.

To Helene, Andrea, Tobias, Daniela and Susanne, who create the productive and friendly environment in our lab, thank you for the team work. For their collaboration in all endeavours relating to administration in the lab and representation in the faculty, I want to thank Lukas, Massimo, Albert and Josef. Last but not least, thanks to all of the E14 group for the countless help I received from you in the lab, office and seminars and of course all the fun we had on the way.

The scientific work in this thesis was almost always a collaborative effort. I generally want to thank all of my co-authors and students, with extra thanks to Matthaeus and Andi for their essential part in the successful bacterial oscillator project and to my students Markus and Maria for their help and team work. Thank you to all proofreaders, Lukas, Markus, Louis and Tim.

Especially, I want to say thank you to my family and my friends. Thank you to my parents, Christina and Peter, for making everything possible and for being there with me.