

Chapter 1

INTRODUCTION

The work presented in this thesis sits at the intersection of microbiology: studying life too small to see (bacteria, archaea, viruses, fungi, prions, protozoa, algae etc.); and bioengineering: using engineering tools to understand and build biological systems. The approach taken in this work is sometimes called *synthetic* biology for its emphasis on "rebuilding" biology into easily analyzed, well-understood systems with new functions.

Bacterial synthetic biology is particularly exciting due to a powerful synergy between the current states of microbiology, biotechnology and engineering. A long history of microbiology research has taught us the kinds of things bacteria can do and what tools they use to do them; a growing repository of biotechnology allows us to edit, shuffle or otherwise play with the DNA that encodes such tools; and engineering disciplines contain tried-and-true analytical frameworks to understand changes we make to complex systems, allowing us to design new biological systems or better understand ones we study.

Specifically, we are interested in using DNA editing technology and the mathematical framework of dynamical systems to design new systems to control heterogeneous *communities* of bacteria.

The social life of bacteria

Even before the microscope made microbes visible, scientists speculated that life of some unseen sort was responsible for things like food spoilage, disease, or the transformation of grain juice into alcohol. As technology marched forward, so too did our understanding of the very real microbial world. We now know a great deal about where microbes live, what they look like, what they eat, how they grow, etc. These discoveries have provoked revolutionary advancements in medicine, chemistry, manufacturing, environmental science—it is hard to overstate the impact microbiology has had on human society and our understanding of our planet.

Current microbiology research drills into the molecular sub-cellular workings of microorganisms, but also examines the larger roles they play in environments like human bodies, soils, or the planet as a whole. While the concept is not new to the

field, the last few decades have seen an explosion in studies acknowledging microbial populations as pseudo-multicellular organisms comprised of highly interconnected, interdependent—*social*—communities.

These social communities rarely contain only one member species; rather, tens to hundreds to thousands of species coexist alongside each other in most environments [1]. We can describe a community by its *composition*, the relative abundances of each species in the population. A community's composition may be termed *rich* when many different species are present or *even* when the relative abundance of the member species is more equal. Many metrics of diversity attempt to combine richness and evenness into a single numerical quantity, usually increasing with both.

Among communities of macro-scale organisms, higher community diversity is associated with greater community stability and productivity (referring to the community's ability to stably maintain size, diversity, and important ecosystem functions) [2]. On the micro-scale, the same principle holds. Decomposition of organic matter—a critical process for whole ecosystems—is reduced as soil microbial diversity falls [3]. Human gut microbiome composition is correlated with, or even causative of significant medical conditions like diabetes and autoimmune disease [4–6]. The relationship between commensal microbes and the physiology of their hosts raises interesting new ideas about community "productivity" or "function" (does digestion count as "productivity"? Is human health a "function" of microbes?). Whatever role microbes play in their environment, it is clear that **microbial community composition is important** and deserves continued attention from biologists and engineers who wish to understand how it is and why.

Because microbes are such powerful forces in their environments, when environmental change disrupts a community's composition, compromising its sometimes critical function, environmental degradation and microbial community collapse feed back on each other out of control. An engineer thinking about this problem will recognize that stable, diverse community compositions are clearly desirable states and wonder "how can I cause a microbial community to establish a target composition? And how do I make it stay there?".

Genetic engineering

We are only able to attempt microbial community engineering thanks to the proliferation of reliable, easy to use genetic engineering technology. Biochemists discovered, studied and optimized enzymes that act on DNA, transforming them

into the common hammers and nails that allow us to build DNA essentially at will. This transformative technology allows researchers to configure the small pieces of biology we know a lot about into large systems for desired purposes.

The DNA technology used for genetic engineering was itself mined from the DNA of bacteria, viruses, and other microbes. Engineers adapted microbial technology with appropriate DNA-related functions to new roles as human technologies. Polymerase chain reaction (PCR) [7], Gibson assembly [8], GoldenGate assembly [9], Sanger sequencing [10] and bacterial transformation [11] are the basic techniques that enable all the work we present (see 3G assembly for the full protocol [12]). They are clever, powerful things that deserve their own attention, but are considered basic technology in the field and go without detailed explanation.

To build systems that enable control of bacterial communities, the process is the same; we need to know what tools bacteria use to regulate themselves and their communities...then steal those tools for our own purposes. The classic process in synthetic biology is "part-ification" of such tools, the full characterization of a complex biological tool and subsequent transformation into a well-understood "part" (like a capacitor or an I-beam) that can be used to build new biological systems.

The process of characterizing engineered parts is difficult in any discipline without equally well-understood, standardized testbeds in which a part under study can be precisely manipulated and measured. Biology is an especially interesting discipline for how little we understand it—despite all we know today, there is still so much to learn before any biologist can claim what they study is "fully understood". This sets up a unique challenge for bioengineers trying to make precise, predictable systems in a platform that is inherently complex and unpredictable. To get the best of both worlds, our work is carried out using the world's most well-studied microbe (maybe the world's most well-studied *organism*): the bacterium *Eschericia coli*. *E. coli* is easy to grow, has a completely sequenced genome and has been the testbed for nearly every piece of DNA technology ever developed. Laboratory-optimized strains of *E. coli* provide an ideal balance of a controlled biological system (as "under control" as any biology can be) and relevance to the "real-life" biology that patterns our bodies and our planet.

Synthetic biology and microbial communities

In the early 2000's, "synthetic biology" was born. Bioengineers used DNA technologies to create cell-powered "genetic circuits" that imparted complex behaviors

to the cellular host. These early investigators paired an intuitive dynamical systems modeling approach, based in ordinary differential equations, with precise laboratory work to design circuits that bestowed new behaviors on bacteria [13, 14]. They demonstrated that human-designed genetic programs could be reliably coded into populations of bacteria, whose resulting behavior recapitulated the predictions made by the designer's model.

Synthetic biology has grown to encompass a great variety of topics, from our work in microbial community control to the effort to build a functioning bacterial cell from the ground up, supported by advanced quantitative disciplines like engineering, physics, statistics and control theory. In general, synthetic biologists work like engineers, using standardized tools and techniques to create new things with biological parts.

In the case of microbial community control the synthetic biology framework uses mathematical tools from control and dynamical systems engineering to transform our microbiology knowledge and DNA editing technology into biological designs. However, we want to build systems to control microbes—living things—rather than more standard control targets like temperature in a refrigerator or liquid level in a water tank. To do this, we need biological equivalents of the parts used to make a thermostat or level controller: a sensor to detect the state of the system, a controller to turn system state into a decision about what to do next, and an actuator to affect the system and keep it under control.

Control systems are often modeled as sets of differential equations whose analysis with mathematical and computational tools give lots of insight into the properties of the system. Analytically solving the equations reveals system steady states and the stability of those steady states. Computer simulations can numerically compute the expected behavior of the system from different start states, or identify parameters in the system critical to its performance. Data collected from experiments help uncover the likely values of parameters in the equations, which makes model predictions even more accurate. We rely on a number of these techniques to aid our design process and streamline the process of building our population control systems in the laboratory.

Unlike a refrigerator, an engineered microbial community does not have a central computer that handles all sensing, processing and actuation. The control functions are distributed among all the individual bacteria in the community, meaning the goal is to design a DNA-based system for each bacterium that allows the community to synthesize and sense the community state, then coordinate decision making and

actuation to maintain control. The genetic parts required must:

1. Send information between cells in the community
2. Process and respond to information signals from the community (or environment/experimenter)
3. Regulate the number of cells of any type in the community

These functions knit a group of individual bacteria into a true *community* that acts together to actively change its own characteristics (e.g. total size or composition) in support of a desired ecological/biochemical/medical goal.

Bacteria have a staggering variety of molecular tools to do exactly these things, the great majority of which are likely unknown and unused—except by the bacteria that invented them. Using DNA editing technology, we can take microbial tools for communication, information processing and community regulation and configure them into a population control system whose behavior we can predict using control system modeling tools.

1. Communication is key

Much of bacterial communication is done by exchanging metabolites, enabling different members to grow at different rates, determining composition. This communication is convoluted, arising as a consequence of metabolism—intimately tied to just about every process in a cell. Engineering community control by manipulating metabolism is certainly powerful, but runs the risk of destabilizing the entire biological system if not planned very carefully. This is not meant to undermine the importance of metabolic interactions; they underlie much of native community structure and can be useful tools in microbial community engineering [15].

Some bacteria are more explicit about communication, decoupling it from the basic mechanisms of growth. Acyl-homoserine lactones (AHLs) are chemicals that allow bacteria to communicate specific messages beyond "grow at this rate due to the availability of nutrient x", giving them a more active, precise role in community determination. AHLs are used in nature to ensure community coordination and increase fitness [16, 17]. They can also be parasitized by invaders to destroy coordination and establish new community compositions [18, 19].

Metabolism-independent cell-to-cell communications modules are extremely valuable to microbial community engineers; these AHL systems have already been adapted into the synthetic biology toolbox to great effect [20–23].

In AHL systems we find simple, genetically encoded modules for creating communication among our engineered community members. There are many different AHL systems available for use in synthetic biology [24–26]; we primarily use the Lux (3-oxo-C6-HSL) [27] and Cin (3-hydroxy-C14-HSL) [28] systems.

2. Control of gene expression

Already in the toolbox of synthetic biologists are various transcription factors, protein tools bacteria natively use to ensure efficient regulation of their repertoire of genes. Combining a transcription factor (TF), a DNA element that regulates gene expression in response to that TF, and a gene of choice, a synthetic biologist can set up experimenter-controlled expression of that chosen gene in a bacterium. Controlling expression of genes via TFs is the foundational technique in most synthetic biology work; regulating the expression of transcription factors with other transcription factors allows gene expression networks to perform complex, useful functions [29–32].

Most engineered DNA is input into bacteria as a plasmid, a circular piece of DNA that replicates independently from the main genome. However, there is a limit on the total amount of plasmid DNA a bacterium can accept. When a synthetic biologist wants to use a TF to regulate a gene in their system, they need to include it on a plasmid, using up space that may be required for other critical components. To keep design space open, TF expression is usually integrated into the main bacterial genome. Easy to use integration technologies enable researchers to create their own genome-integrated *E. coli* strains [33]; even more elaborate integrated strains are made as chassis for synthetic biology, enabling plug-and-play gene regulation with known parameters [34]. We both create our own custom genome-integrated *E. coli* and use the versatile Marionette *E. coli* strains to create our engineered bacteria.

3. Regulating life and death

In native bacterial communities, there are 4 ways for the population of a member to change: emigration, immigration, reproduction and death. A human gut microbiome regularly receives new input that can contain additional bacteria, likewise it is purged by gut motility. Nutrients are regularly available, but growth is far slower than in optimal conditions. Death is, of course, inevitable at some rate. Controlling a

bacterial community means controlling one or more of these activities. Here again, we can turn to bacteria themselves to find useful tools to meet our needs.

Some bacteria have evolved tools to actively regulate life and death within their communities. These are frequently genetically encoded protein toxins, produced by a cell to kill itself on command. In some cases, the toxin is paired with a specifically matched anti-toxin that spares a producer from the toxin. Toxins and anti-toxins (TA systems) play roles we are only still discovering in modifying bacterial populations. A "selfish gene" explanation might contend that the genes encoding TA systems use their functions to ensure their survival or dominance in the collective genome of a bacterial population (many TA systems cause bacteria to become "addicted" to their presence and create fitness repercussions if the TA system is jettisoned). On the other hand, bacterial populations may retain TA systems as tools to improve their fitness in particular situations [35, 36]. Whether the TA systems selfishly manipulate bacteria or the bacteria have found utility in the TA systems, TA systems absolutely regulate life and death to modify population composition.

With our understanding of bacterial physiology and our ability to manipulate it, there are endless opportunities to create bacteria whose growth or death are controlled by an experimenter—set up outside control of ribosome synthesis [37], control expression of a critical metabolic protein [38], even simply regulating the metabolic load on a cell can alter growth rate [39]. Many toxins and antitoxins affect the rate of *death* in a community; others, like the T7 phage *gp2* protein modify the *growth rate* of cells by interfering with critical growth processes—in this case RNA synthesis by RNA polymerase [40]. Every growth (or death) regulatory tool has a mode of action, potency, strength and weakness that makes it unique and may suggest a particular use case in population regulation. A particularly interesting set of TA systems, the bacteriocins, work *between* cells rather than inside individual cells. Their structure and multi-functionality make them attractive for microbial community engineering and will be discussed in Chapter 3.

Toxins, antitoxins, growth inhibitors, engineered physiology—growth regulators in general—are fascinating, both for their ecological roles and their potential in bioengineering. Many of these tools are receiving renewed interest due to the expansion of synthetic biology; hopefully their "part-ification" can be useful to the field and reveal new complexity in microbial ecology.

Putting it together

With the explosion of interest in microbial communities in soil, the ocean, human guts, scientists are applying research approaches from every discipline of science to shed light on the complex, influential, social lives of microbes. Using insights from microbiome researchers and experience from pioneering microbial community engineers, we have designed genetic circuits we hope will be useful in building and manipulating the powerful microbial life that shapes our planet.