Synthetic Biology for Knockouts, Knockdowns, and Phospho-Sensing in Diverse Bacteria

A dissertation presented

by

Isaac Nathan Plant

to

The John A. Paulson School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

January, 2019

Abstract

Biology is defined by experimental stories. These stories, however, are limited by the technologies that are used to tell them. In microbiology, there are limits on the types of organisms we can use, the therapeutic outcomes of anti-microbial treatments, and the enzymatic activities that can be sensed *in vivo*. My Introduction describes these limitations, and the rest of my dissertation presents my work developing engineering-solutions to them. In Chapter 2, I present research towards a Potentially Organism-Agnostic Knockout (POAK) system. Our ability to knockout genes limits the number of bacteria that are tractable, and I demonstrate that POAK can expand the number of species that we can work with. In Chapter 3, I discuss efforts towards DEcreasing the Selective Pressure Of phage Therapy (DESPOT). That chapter deals with a novel approach towards bacterial infections that aims to have therapeutic benefits without selecting for resistant bacteria. In Chapter 4, I detail a set of potentially Host Organism-Agnostic Kinase Sensors (HOAKS). Current technologies do not allow single cell measurement of serine/threonine kinase activity in bacterial cells. In that chapter, I present two tools that can perform these measurements. These three sets of tools, POAK, DESPOT, and HOAKS, expand the microbes we can work with, the infections we can treat, and the biology we can sense.

Contents

Acknowledgments

A Ph.D. is too large an undertaking to be considered simply as the product of a single person. The people who have played a part in mine are too numerous to thank here, so a subset will have to suffice. My family has been an essential source of support. My parents have, perhaps to their regret, always told me I could be anything I desired. My mother encouraged my love of science, my father encouraged my desire for reason and order, and my step-mother made sure I always remembered that the goal is to make the world a better place. My siblings, Ethan, Alex, Ahna, and Matty, are the reason I am always so happy to return home. Doug, Val, Ben, and Sophie made leaving Boston a difficult choice. To the rest of the family: I love you, but there are too many of you to list your names. My friends from college and my first attempt at grad school, Pat, Leah, Jordan, Aaron, and Kyösti, have continued to welcome me back despite my monotonically increasing pedantry. The Silver Lab at Harvard was a perfect place to do a Ph.D., in large part thanks to Pam, who allowed me freedom I did not deserve. The Silver Lab provided me with Marika, Charlotte, and Christine, who have proven to be great climbing buddies and friends. Tobi and Yu Heng were wonderful bay mates and Bryan has been a terrific collaborator (and read my entire thesis). Sasha was with me from the beginning, and I could not have asked for a better companion. Steph was my "mentor" in the Silver Lab and has become one of my dearest friends. Thank you for reading this. Twice. Without Kathy, or the Sys Bio admin, the Silver lab would not function, and this thesis would not exist. Hayley, Sarah, and Wade are wonderful friends who provided a needed dose of non-science reality. Chuck, Oana, Michael, and Hank, were

amazing roommates. They made the whole process bearable.

To all of you, I am better for knowing you. Thank you.

Finally, it is important to emphasize that my earning a Ph.D. from Harvard was largely determined by my birth. I was born into a rich family in the United States and was raised as a white man. These features have made everything I have worked for easier to achieve, and everything I have desired more likely to occur. There are individuals who have come from less privileged backgrounds to achieve what I have, and they have intelligence, work ethic, and perseverance that I do not. I am confident that if I were in their shoes, I would not have a Ph.D., while if they were in my shoes, my Ph.D. would be better for it.

Chapter 1

Introduction

The single story creates stereotypes, and the problem with stereotypes is not that they are untrue, but that they are incomplete. They make one story become the only story.

C. N. Adichie

You get what you select for. Always.

S. G. Hays

Telling Biological Stories

Biology is a science of stories. We seek patterns for the complexity of life, and because we seek them, we find them. From these patterns, we create stories that describe a link between cause and effect. In our stories, we have a protagonist, whether it is Anna Karenina or a kinase, and we ask what drives it, what it does, what its purpose is. Stories allow us to make sense of biological causality[1], and when describing experiments, we ask how they fit into the story.

But as stories allow us to frame reality, they also force reality into our frame[2]. Stories represent a series of choices about what to include and what not to include. These selections are a narrowing of perception onto a set of actors and conditions, and the entirety of a story is dependent on how we make these decisions. Who is the driving force behind the story? Which of their actions are important? How will the story end? This selection process is filled with pitfalls, as the two opening quotes of this chapter suggest. In my dissertation, I have sought to address some of these decisions by developing new tools for describing biology and performing biological engineering.

This dissertation focuses on three particular types of decisions. One is the setting of the biological story we tell, the next is what constitutes an ending, and the last is the type of tool we use to tell the story. We routinely make these decisions based on convenience. Yet, convenience disguises the fact the we made a decision at all, and makes us view the default as the only possible outcome. In this chapter, I locate these three decisions in specific biological contexts and detail some of the technological challenges that arise. I then suggest why we should think about these decisions more explicitly and I briefly summarize the content of the rest of my dissertation.

For laboratory biologists, our stories are set within organisms, or parts thereof. There are almost always multiple possibilities for the type of organism that could be used, but the type of organism is routinely taken for granted. Now, there are often legitimate reasons to choose the default. Some questions can only be reasonably asked in a given organism. Un-

derstanding what makes a specific *Enterococcus* strain vancomycin resistant requires working in that specific *Enterococcus* strain[3]. Yet, often, when we choose the default it is not out of necessity but because the other options are prohibitively difficult. We write about our home but not the rest of the world because we do not have enough money to travel. Or, we write accurately about our home but inaccurately about the world. The latter is perhaps more common in biology, and molecular biology in particular. We study a convenient organism (*Escherichia coli* or *Bacillus subtilis*) and then extrapolate to everything else. There are a number of reasons we settle for using *E. coli* or *B. subtilis*, but two determinative ones are that we can grow them and that we can knockout genes in them. There is important work being done in the former category, but I will not discuss it here. Rather, I have focused on the latter. There are a vast number of bacteria that we can culture, yet for many we cannot make sequence specific knockouts. As a result, we are extremely limited in the range of questions we can ask. My second chapter deals with my biological engineering research on a novel bacterial knockout tool, while in the *Bacterial Knockout Technologies* section below, I review current knockout technologies.

Defining an ending for a story, or an endpoint for a project, is essential. Yet, it is easy to misidentify the most straightforward, or most common, potential endpoint as the only acceptable one. The protagonist slaying the dragon is one endpoint, but so is the dragon being put in a zoo. If all of our stories were the former, we would have a very limited view of the world. Yet we often mistake the goal of treating bacterial infections with the goal of killing infecting bacteria. As such, we have devoted numerous resources towards anti-microbial therapies that by their very nature select for resistance against themselves. However, if we remember the actual goal, that we are trying to make infections less impactful, we can work on treatments that may achieve therapeutic benefit without killing the causative agent. My third chapter describes our progress towards one such therapy, and in the *Anti-Microbial Therapies* section below, I review current anti-microbial and anti-infection strategies.

Stories can be defined by how they are captured, whether that be through sound, vision,

or the written language. They are so often defined this way because the tools that we use to describe the world put specific limits on what parts of the world we can observe. The sound of a fire crackling and a picture of Paradise burning tell us two very different things. Significant efforts have been devoted to expanding what parts of biology we can see (no less than three Nobel prizes have been awarded for microscopy), and advances in synthetic biology have recently expanded the types of phenomena we can observe *in vivo*. Each of these advances expands the stories we can tell, and in some small way resets how we understand biological systems. One of the biological activities that has been historically difficult to observe *in vivo* is post-translational modification, namely protein phosphorylation. In my fourth chapter, I present two tools that can be used to detect protein kinase activity in bacteria, and in the *Serine/Threonine Kinases and Measuring Phosphorylation* section below, I describe bacterial Serine/Threonine Kinases (STK) and current technologies for sensing protein phosphorylation.

Bacterial Knockout Technologies

The first step in understanding what a part in a machine does is to remove the part. In machines that we build, it is straightforward to simply not include the part of interest. In the living machines that eons of evolution built, however, we need to find ways to remove the parts of interest. While some parts of life, such as Ca^{2+}, can be removed by altering the local context of the organism (by addition of a Ca^{2+} chelator such as EDTA), the parts we tend to be interested are DNA, RNA, and proteins. To remove these parts, we must physically alter the organism's DNA. In this section, I will discuss four technologies for creating knockouts: (Classical) Mutagenesis Screens, Transposon Mutagenesis, Homologous Recombination, and Lambda Red, as well as one knockdown approach, CRISPRi. There is some overlap between these technologies, but each has distinct advantages and disadvantages.

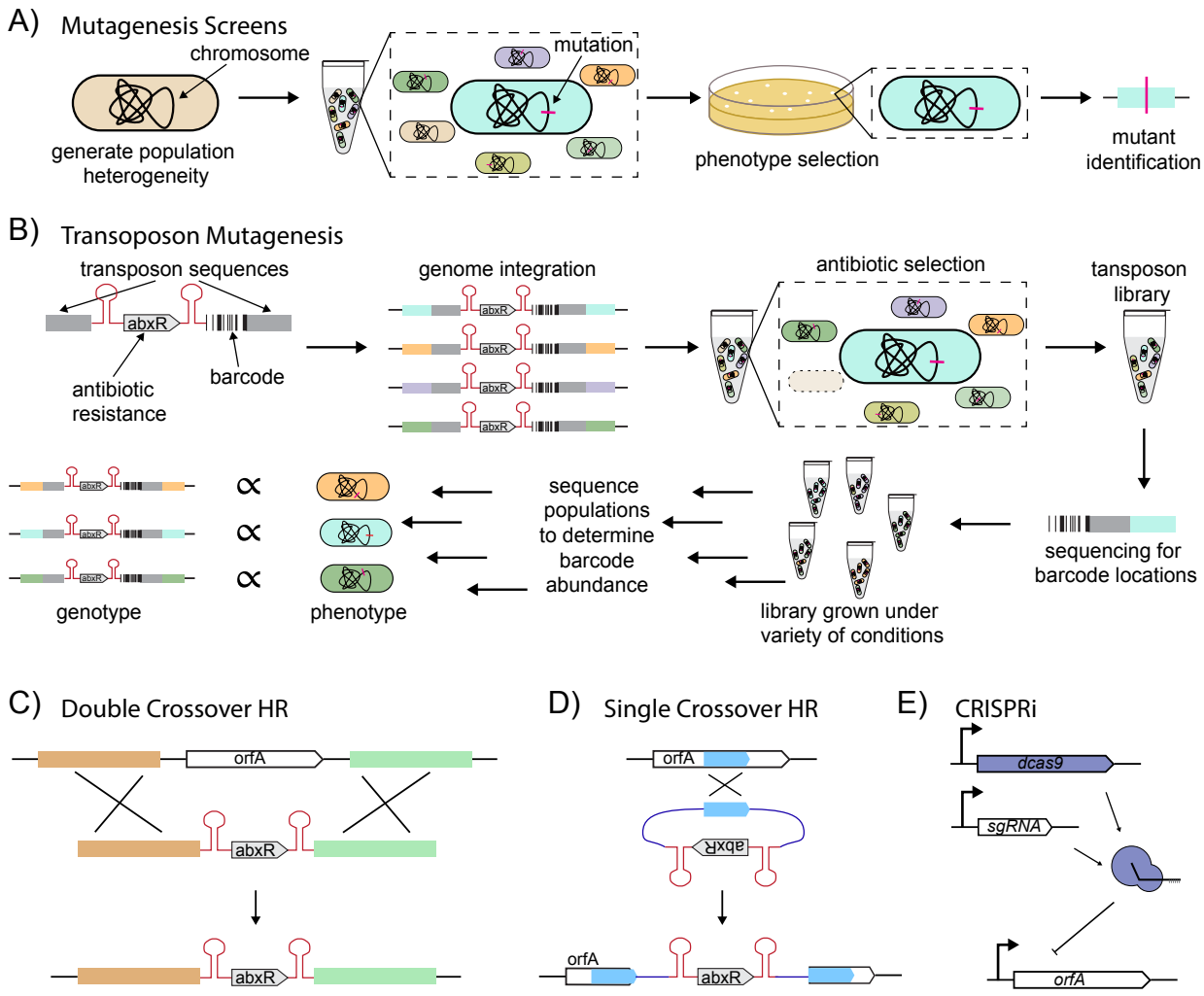


Figure 1.1: **Bacterial knockout technologies A)** Mutagenesis Screens are performed by mutating a population of bacteria, selecting for the desired phenotype, and sequencing the resultant strains. **B)** Barcoded Transposon Mutagenesis involves inserting a modified transposon across the bacterial genome. Modified bacteria are selected using antibiotics, and then the whole library is deep sequenced to locate where each barcode is inserted. Following future experiments, the abundance of the barcodes can be determined cheaply by sequencing and then used to link phenotypes to genotypes. **C)** Double Crossover Homologous Recombination (HR) uses two homology regions to replace a native genetic locus. **D)** Single Crossover HR uses a single homology region that, as a result of integration, is duplicated. The process is reversible, due to the duplicate homology arms. **E)** CRISPRi uses dCas9 and a gRNA to repress transcription of a target gene.

Mutagenesis Screens select for a specific phenotype from a population containing either natural or induced mutations (Fig. 1.1A)[4]. As such, it requires three steps. First, a population with a large number of mutations is created. Population heterogeneity can be achieved a

number of ways, including through natural mutations, mutator strains, chemicals, radiation, transposons, and phages[5]. Second, a selection for the desired phenotype is developed. Finally, the mutation that led to the phenotype of interest is identified through whole genome sequencing. When developing a mutagenesis screen, the selection step is typically the most difficult. If you want to select for an *E. coli* strain that is deficient in galactose metabolism, you require a selection that causes the sugar metabolism deficient strains to have a substantial fitness advantage over the strains that can metabolize the sugar. In the case of galactose, that can be accomplished with the sugar derivative 2-Deoxy-D-Galactose, which is toxic to cells that have *galK* (a gene essential to galactose metabolism)[6,7]. As a result of this selection step, and modern sequencing technologies, these screens can provide a powerful way to link unknown genotypes to known phenotypes. However, they cannot do the reverse and link a known genotype to an unknown phenotype, which is what targeted genetic knockouts do. Further, that mutagenesis screens must link a phenotype to a selection or screening method means that they require either strong selections, or the ability to screen many thousands of colonies. This is not feasible for many experiments, which is why it is often preferable to decouple the selection method from the desired phenotype. Transposon mutagenesis is one method to do this.

Transposons are mobile genetic elements that are found across all domains of life[8,9]. While many transposons will only insert themselves at specific sequences, some transposons have limited-to-no sequence requirements. These promiscuous transposons have been co-opted by biologists as tools for randomized knockouts[10,11]. By adding an antibiotic resistance cassette and terminators to a transposon, insertion of the transposon predictably halts transcription at the insertion site and the insertions can be selected for (Fig. 1.1B). As such, libraries can be created that have transposons inserted across an entire genome. If barcodes are added to the transposons, specific barcodes can be linked to an individual transposon's insertion site[12]. This makes it simple to track the fitness of specific genotypes of interest when the entire library is used for experiments. However, transposon libraries are almost exclusively used

as whole libraries. While there has been some work towards de-convoluting libraries into individually accessible strains that have known gene knockouts, the difficulty and expense of such approaches has limited their adoption[13]. For making targeted knockouts, homologous recombination based techniques hold sway.

Homologous recombination (HR) is a DNA repair mechanism found in all organisms. HR recombines two homologous sequences either at one point, single crossover, or at two points, double crossover. DNA repair requires a double crossover event and functions to copy a region of DNA from between the two homology region of an undamaged template over the damaged section of DNA. If the original genetic locus is undamaged, this process can also replace one genetic locus with another (Fig. 1.1C). Single crossover, on the other hand, functions to exchange genetic loci, and cannot be used to repair circular chromosomes. It can be used to insert an entire circular piece of DNA into another (e.g. integrate a plasmid into a chromosome), although this results in duplication of the homology region (Fig. 1.1D). Single crossover HR is reversible, and recombination of a homology region can result in excision of the original circular piece of DNA. While HR exists as a mechanism of error-free DNA repair, it can also be used to generate knockouts.

Generally, double crossover HR is used to generate knockouts (Fig. 1.1C)[14]. To do so, the cell is provided a DNA template that has a knockout cassette (terminators on either side of an antibiotic resistance cassette) flanked by homology regions from upstream and downstream of the gene to be knocked out. The cell then replaces the gene with the knockout cassette. This general approach will work in any organism, but its difficulty varies considerably. For some naturally competent organisms (such as *B. subtilis* and *Synechococcus elongatus*) HR is so efficient that it is preferred over plasmids for any genetic work. In most organisms, however, HR is very inefficient, requiring extremely long homology arms (in the thousands of base pairs) and suicide vectors[15]. This has led to the development of more efficient recombination systems, such as Lambda Red.

The Lambda Red system allows for highly efficient double crossover recombination into

E. coli with as little as 30 bp of homology[15]. The system uses a single strand recombination system from the lambda phage to stabilize short homology regions and increase the rate of HR integration. The technique is so effective that the Datsenko & Wanner implementation has become the standard for any engineering that requires integration into the *E. coli* chromosome[16]. The Lambda Red system, however, is limited to *E. coli* and *Salmonella typhimurium LT2*. A related type of recombineering that uses single strand oligos to make small mutations in an organism's chromosome has been expanded to a handful of other organisms. However, its engineering uses have been limited and it has not proven to be effective for making targeted knockouts[17]. As a result, for most prokaryotes, there remains no effective way to knockout specific genes. One workaround has been to knockdown gene expression instead of knocking out the target gene, as this does not necessarily require modifying the target DNA. CRISPR Interference (CRISPRi) is a particularly appealing knockdown method for bacteria.

CRISPRi represses transcription of a gene by targeting it with a deactivated Cas9 (dCas9) (Fig. 1.1E)[18]. The dCas9, which binds the DNA but does not cut it, binds tightly enough to prevent transcription. Since all that is needed to target the dCas9 is a gRNA, this approach is easily modified for the target(s) of choice. Further, dCas9 can be modified to work in essentially any organism, which has allowed CRISPRi to be developed for a number of previously less-than-tractable bacteria.[19] Unfortunately, CRISPRi cannot be used as a substitute for knockouts. While repression is highly efficient, it may not be efficient enough to completely prevent expression. Further, mutations in the targeted sequence, gRNA, or dCas9, can relieve repression. Potentially more problematic, though, is that CRISPRi represses transcription of everything downstream of a promoter, resulting in knockdowns of entire operons. This can be a feature, but can also limit attempts to isolate the affects of a single gene. As such, CRISPRi is not a solution for making targeted knockouts and our ability to generate knockouts has remained a limiting factor for studying a variety of bacteria.

In Chapter 2 I will present work combining Cas9 and an error-prone DNA repair system, Non-Homologous End Joining (NHEJ), into a tool to create targeted knockouts in an organism-agnostic manner.

Anti-Microbial Therapies

The discovery of antibiotics marked a turning point in how humans dealt with infectious disease. While previously limited to preventing an infection (through vaccines, adequate sanitation, and cleanliness) or treating the symptoms, antibiotics were able to deal with the infection itself. From the first synthetic antibiotic in 1909 (Salvarsan) to the discovery of penicillin in 1928, through the large screening programs of the 1950s and 1960s, antibiotics were one of the defining discoveries of the 20th century[20]. With the close of the 20th century, however, came the rise of antibiotic resistance[21]. Dealing with resistant infections is an increasingly pressing problem, and remains inadequately dealt with. In this section I review what leads to antibiotic resistance, other approaches to killing pathogenic bacteria, and non-bactericidal approaches for dealing with bacterial infections.

It is not difficult to kill bacteria. Heat, desiccation, acids, bases, reductants, and oxidants, among other things, can be used to killing pathogens. Unfortunately, all of these can also kill humans. The difficulty for antibiotics is finding a way to kill bacteria without harming the patient. As such, antibiotics must target biological motifs that are essential but also unique to bacteria[22,23]. The fact that antibiotics must target unique features of bacterial life, and not other types of life, means that there is always a way to for an organism to be alive and also be resistant to an antibiotic. While this does not imply that a given bacteria will be able to become resistant to a given antibiotic, we know empirically that if an antibiotic is used for a sufficient period of time in a sufficient population, resistance to the antibiotic will develop[24]. There are four ways in which resistance arises: alteration of the antibiotic, modification of the antibiotic target, decreased intracellular concentration, and widespread

cellular modifications (Fig. 1.2A)[25,26].

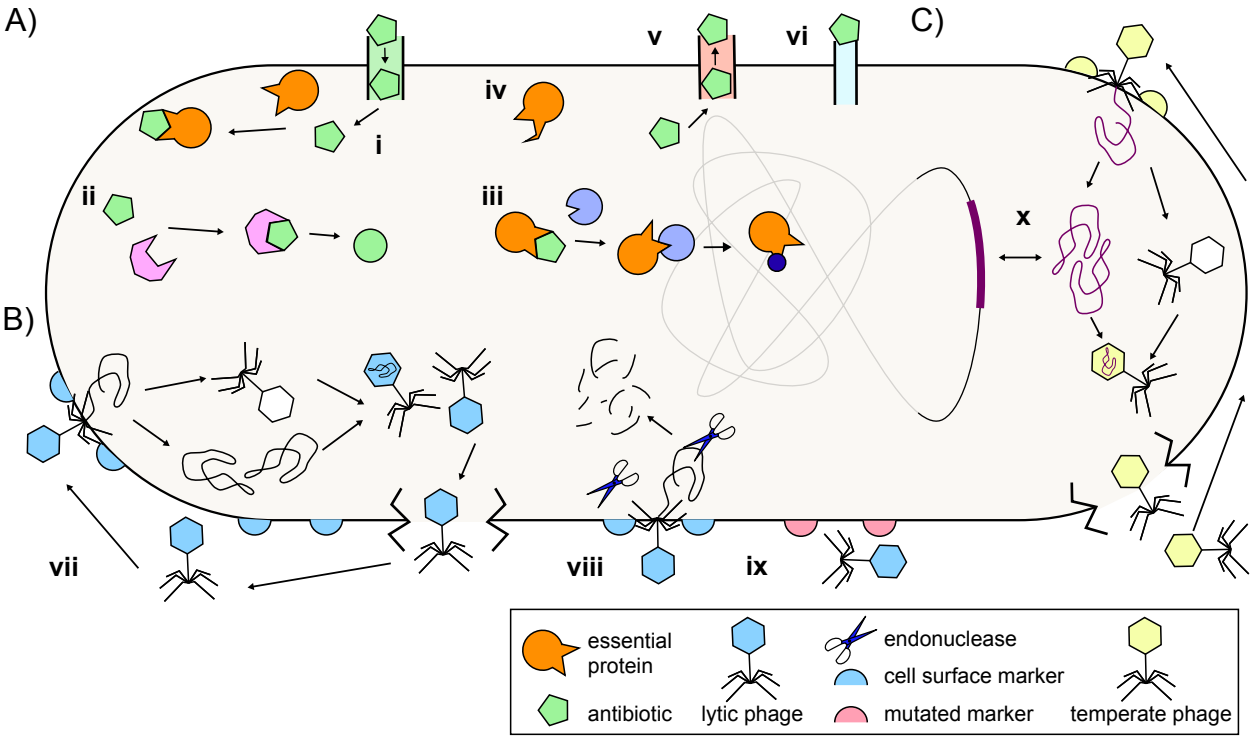


Figure 1.2: **Modes of antibiotic resistance, lytic phage infection, and temperate phage life cycle A)** Normally, (i) an antibiotic passes into the cell and inhibits an essential protein or function. A bacterium can become resistant if (ii) an enzyme modifies or destroys the antibiotic, (iii) an enzyme modifies the antibiotics target, (iv) the essential protein is mutated, (v) an efflux pump removes the antibiotic, or (vi) the antibiotic cannot penetrate into the cell. **B)** A lytic phage (blue) normally (vii) infects a cell by binding to cell surface proteins, injects its genome, replicates its genome, produces its capsid (white), packages its genome into the capsid to form a complete phage particle (blue), and then it lyses the cell and escapes. Some of the ways a cell can become resistant include (viii) endonuclease activity (blue scissors) against the phage genome and (ix) absent or modified surface proteins that the phage cannot bind to. **C)** A temperate phage (yellow) can proceed through a lytic life cycle, but can also (x) integrate its genome into the bacterium's genome and remain there as a prophage.

Cells affect antibiotics directly either by modifying the antibiotic (e.g. chloramphenicol acetyltransferase which modifies chloramphenicol) or be destroying it (e.g. beta-lactamases which degrade beta-lactams, such as penicillin) (Fig. 1.2Aii). Since many antibiotics are natural products, the organisms that produce them, or ones that live in close proximity to the producers, often have these types of resistance genes. These genes can spread to other

organisms through horizontal gene transfer (HGT) and they are usually specific for certain antibiotics (e.g. chloramphenicol) or a set of closely related molecules (e.g. beta-lactams)[27]. Modification of the antibiotic target is also antibiotic specific and can occur by acquisition of modification enzymes through HGT and by mutations. Examples of enzymatic modification (Fig. 1.2Aiii) include erythromycin resistance, which modifies the ribosome (the target of erythromycin). Streptomycin resistance, on the other hand, occurs though a point mutation in *rpoS*, and is an example of a mutation in an antibiotic's target leading to resistance (Fig. 1.2Aiv)[25,26]. While both of these types of antibiotic resistance will only confer resistance to a particular molecule (or class of molecules), mechanisms that decrease the intracellular concentration of antibiotics, as well complex physiological changes that decrease antibiotic sensitivity, may confer resistance to a broader spectrum.

Changes in intracellular antibiotic concentration may be the result of resistance genes acquired through HGT or through mutations. Efflux pumps reduce the intracellular concentration by pumping out the antibiotic, such as *tetA* which pumps out tetracycline, and are a common mechanism of resistance due to HGT (Fig. 1.2Av). Alternatively, alterations to the membrane, cell wall, and porins can all reduce the penetration of antibiotics into the cell (Fig. 1.2Avi), thereby reducing intracellular concentrations. The most extreme example of this is antibiotic susceptibility of Gram-negatives (which have an outer lipid membrane that reduces penetration of many antibiotics) versus Gram-positives (which lack an outer membrane and therefore are not as protected). However, more subtle changes, such as which specific porins are expressed, can also reduce intracellular antibiotic concentration and therefore lead to resistance[28]. Finally, a combination of seemingly minor metabolic and physiological changes, such as mutations in transcription factors or changes in kinase signal cascades, can result in resistance to antibiotics. With numerous paths for a bacterium to gain resistance to any given antibiotic, attention has turned towards non-antibiotic methods of treating infections.

Phage therapy is one method that has seen increased interest. Bacteriophages (henceforth

referred to as phages) are viruses that infect bacteria, and phage therapy refers to the use of phages to treat bacterial infections. Generally, this is done with lytic phages, as opposed to temperate phages. Lytic phages proceed through a lytic cycle, where they infect a cell, replicate, and then lyse the cell (Fig. 1.2Bvii). Temperate phages have a lytic cycle and an additional lysogenic cycle. During the lysogenic cycle, the phage integrates into the host genome and becomes inactive (Fig. 1.2Cx). As the host cell replicates its genome, the integrated phage (called a prophage) is replicated as well. The prophage remains stably integrated until a trigger (often DNA damage) induces the phage to excise itself and start the lytic cycle.

Phage therapy has primarily focused on lytic phages, as they achieve the same end as antibiotics: bacterial death[29,30]. Small-molecule antibiotics and phages differ in a number of ways, but host range is perhaps the most important. Whereas even the most specific antibiotics will affect multiple species of bacteria, phages have very limited host ranges, potentially only affecting one strain of one species of bacteria. When an infection is treated with antibiotics, non-pathogenic bacteria are killed as well. Phages offer the possibility of only affecting the infective agent, and reducing antibiotic side effects (e.g. *Clostridioides difficile* infections that arise after antibiotic treatment depletes the gut microbiota). Lytic phages, however, give rise to resistance as well. Resistance to lytic phages, like that of antibiotics, occurs through a number of mechanisms and can arise from HGT or natural genetic drift. Some common types of phage resistance include modifications of the bacterial cell surface that inhbit phage attachment (Fig. 1.2Bix)[31], and host cell endonuclease activity that selectively degrades phage, but not host, DNA (Fig. 1.2Bviii)[32]. Resistance can arise rapidly, and significantly impact a phage's effect on a population. While it is possible to evolve phages to overcome resistance, the ability of phage therapy to change the paradigm of anti-bacterial treatment is limited. Rather, non-resistance inducing treatments are needed.

To treat bacterial infections without inducing resistance, non-bactericidal approaches are required[33,34]. Up until now, the most successful of these "non-traditional" therapies has

been fecal matter transplants[35,36]. *C. difficile* infections often occur after initial antibiotic treatment depletes the natural gut microbiota to a point where *C. difficile* is able to colonize. Once *C. difficile* has colonized, it can be highly recalcitrant to subsequent anti-microbial treatment. Fecal microbiota transplants work by restoring a sufficiently large amount of native gut-microbiota to prevent continued *C. difficile* colonization. Other approaches that have shown theoretical, if not practical, promise include filtering out bacteria from the blood to reduce the infectious load to the point where the patient's immune system can deal with it, and interrupting virulence genes[37]. Disruption of virulence genes, such as inhibiting toxin secretion, may be possible to implement without decreasing the fitness of the pathogen. This would result in the presence of attenuated pathogens in the same niche as the non-attenuated pathogen without selecting for resistant strains. Small molecules have been the primary approach towards disrupting virulence genes. However, they have thus far failed to yield effective therapies[38]. As such, novel anti-virulence strategies are needed.

In Chapter 3, I will discuss our work developing a temperate phage system to reduce the virulence of bacterial pathogens.

Serine/Threonine Kinases and Measuring Phosphorylation

Hanks-type kinases are the pre-dominant type of Serine/Threonine Kinases (STKs) found in eukaryotes. As their description suggests, they phosphorylate Serines and Threonines, generally in a sequence specific manner[39]. Hanks-type kinases share a conserved active site located within a 300 amino acid region of high identity[40]. These kinases are found in all domains of life, and are believed to be derived from a shared ancestor. Historically, these kinases were studied in eukaryotes, where they first discovered. However, after the first characterization of a Hanks-type STK from a bacterium in 1991[41], and the rise of whole genome sequencing through the 2000s, it was recognized that Hanks-type STKs likely play important signaling roles in bacteria.

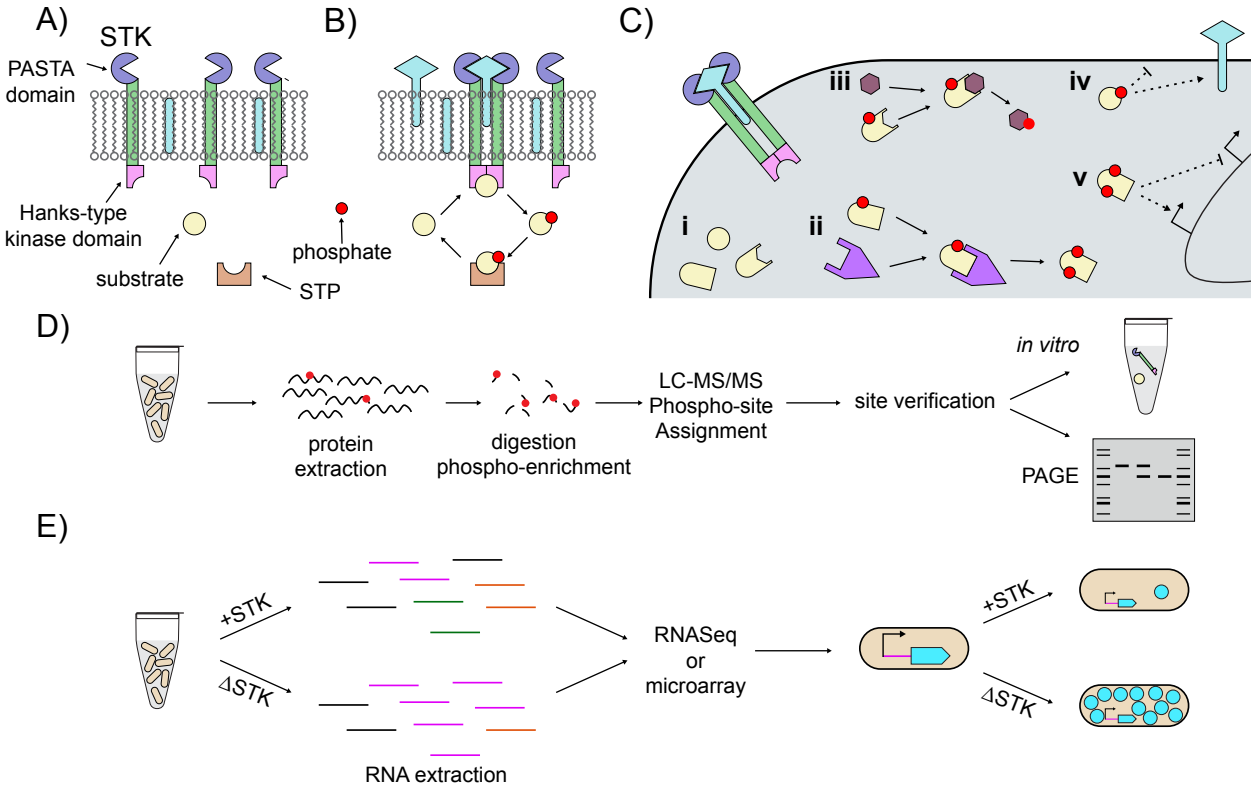


Figure 1.3: **Serine/Threonine Kinase (STK) biology and measurement techniques** **A)** The most well studied bacterial STKs are composed of a Hanks-type kinase domain (pink), a transmembrane domain (green) and a PASTA domain (blue). **B)** When the PASTA domains are brought together by a ligand, the STK can phosphorylate substrates. The cognate STP can remove the phosphate group. **C)** The STKs phosphorylate multiple (i) substrates. Some of these substrates may (ii) also be phosphorylated by other kinases or (iii) may be kinases themselves. Phosphorylated substrates may act (iv) directly on the cell's physiology or they may (v) act as transcriptional regulators, either by repressing transcription, activating transcription, or doing both for different promoters. **D)** Full proteome measurements are the standard method to identify novel phosphorylated residues. All the protein from a culture is extracted, digested, the phosphorylated proteins are enriched, and LC-MS/MS is performed. When phosphorylation sites are identified, they can be verified by *in vitro* assays, or by PAGE+antibody based techniques. **E)** To do single cell measurements, the transcriptional impact of STKs is examined by analyzing the transcriptome of +STK and ΔSTK strains. STK regulated promoters are fused to reporters (cyan) and then used in experiments.

In Gram-positive bacteria, STKs have been implicated in a number of cell wall and cell differentiation pathways, and are thought to be activated by a cell wall intermediate (Fig. 1.3A-C). The first STK characterized in bacteria, Pkn1 in *Myxococcus xanthus*, plays a key role in regulating spore formation[41]. The same is true for the model organism *B. subtilis*,

where deletion of its most abundant STK, *prkC*, results in increased stationary phase lysis and slow outgrowth from spores[42,43]. In *Mycobacterium tuberculosis*[44], *Listeria monocytogenes*[45], and *Staphylococcus aureus*[46], Hanks-type STKs play roles in cell wall regulation (Fig. 1.3Civ). Deletion of the respective STKs often results in increased sensitivity to cell-wall targeting antibiotics, such as beta-lactams[47]. Yet, the evidence for interactions between STKs and their targets is often from *in vitro* experiments, or inferences from complex transcriptional changes upon deletion of the STKs. There is significant interest in being able to measure STK activity and substrate phosphorylation *in vivo*, but the current technologies have significant limitations.

The standard for phosphorylation site discovery is full proteome-mass spectrometry (Fig. 1.3D). All proteins from a cell are isolated, trypsin digested, enriched for phosphorylated residues, and then run through a tandem LC-MS/MS system[48]. Historically, bacterial phospho-proteomics has been limited by the relatively low level of global phosphorylation, and the instability of certain phosphorylated residues (such as histidines)[49]. This requires significant enrichment of phosphorylated proteins to observe even commonly phosphorylated substrates. Progress has been made towards improved enrichment protocols, but the techniques remain less effective than those for mammalian cells[50]. Full cell phospho-proteomics allows for the identification of phosphorylation sites in an unbiased way, although it cannot be used to identify targets of specific kinases unless knockout strains are available. *In vitro* follow up experiments, or phospho-proteomes of both the Wild Type (WT) and kinase knockout strains, are necessary for kinase specific target identification[51].

For a known phosphorylation site, there are a few methods for detecting substrate phosphorylation. If the phosphorylated substrate can be isolated in large quantities, *in vitro* experiments can be performed or an antibody can be raised against the substrate and used for Western Blots, ELISA's, and other immuno-chemistry techniques (Fig. 1.3D)[52]. While to my knowledge it has not been reported in the literature, it is conceivable that an antibody that recognizes only the phosphorylated version of a protein could be used for immunolabel-

ing single bacterial cells[53]. Unfortunately, isolating sufficient amounts of phosphorylated protein from bacteria is difficult, so phosphorylation specific antibodies are not routinely used. Phos-Tag polyacrylamide gels, however, have proven popular. Phos-Tag is a small molecule that complexes with phosphate groups[54]. When Phos-Tag is added to a polyacrylamide gel, phosphorylated proteins have increased retention times relative to their non-phosphorylated counterparts. By using non-phosphorylation specific antibodies, which are much easier to obtain, differences in retention time can be visualized[54]. Unfortunately, like other common antibody techniques, Phos-Tag gels are bulk assays, and cannot detect single cell dynamics.

Currently, detection of single cell kinase behavior requires transcriptional signals of kinase activity (Fig. 1.3E)[55]. This has been a preferred method for studying kinase activity in bacteria, due to the relative ease of bacterial transcriptome profiling, but, like phosphorylation-target identification, requires a kinase knockout strain to compare to WT. When transcriptional markers of kinase activity and, therefore, promoters regulated by the kinase, have been identified, reporter-promoter fusions can be made. If fluorescent proteins are used, then promoter activity can be measured on the single-cell level (Fig. 1.3E). However, promoter activity does not necessarily directly correlate to kinase activity. STKs are one step in a complex network of cell-signaling that is designed to integrate and transform signals into coherent cellular responses. It is not uncommon for specific transcriptional reporters to respond to kinase activity under some conditions but not others[42]. As a result, it can be experimentally intractable to determine when a single cell has high or low kinase activity.

While transcriptional fusions are the state of the art in bacteria, progress has been made in mammalian cells towards developing single cell, kinase specific reporters that are otherwise host physiology independent. A number of strategies have been developed, such as fusing a Förster Resonance Energy Transfer (FRET) pair to a substantial portion of a kinase of interest, and measuring the kinase's conformational change[56]. The most interesting of these approaches, however, is a sensor that uses host orthogonal, modular parts, which Fuller

et al. developed in 2008[57]. Their approach was to link a FRET pair by a phosphorylation-dependent substrate binding domain and a kinase substrate. Using this sensor, they were able to measure dynamic phosphorylation by the AuroraB kinase within mammalian cells. Bacterial cells, however, differ significantly from eukaryotic organisms, and that is perhaps why no one has attempted to use these modular parts to make single cell, physiology independent sensors of STK activity for bacteria.

In Chapter 4 I will present our work on adapting the AuroraB sensor for bacteria, and use of its modular parts to design a novel phosphorylation sensor.

Summary

This chapter began with a discussion of stories, and how stories shape biology. I discussed three outstanding problems that I will be addressing in my dissertation. In Chapter 2, I present a Potentially Organism-Agnostic Knockout (POAK) system, my work towards making it easier to make targeted knockouts in non-model organisms. In Chapter 3, I discuss our work towards attenuating bacterial infections with temperate phages and dCas9. In Chapter 4, I present our work on bacterial STKs and the development of Host Organism-Agnostic Kinase Sensors (HOAKS). Finally, in the Conclusion, I summarize some of my thoughts on the importance of these tools, and how we should think about biological projects.

Chapter 2

Potentially Organism-Agnostic Knockout (POAK) System

Preface

The work detailed in this chapter is my independent research. Tobias Giessen, Bryan Hsu, and Stephanie Hays provided thoughtful comments.

2.1 Abstract

Making targeted gene deletions is essential for studying organisms, but is difficult in many prokaryotes due to the inefficiency of homologous recombination based methods. Here, I describe an easily modifiable, single-plasmid system that can be used to make rapid, sequence targeted, markerless knockouts in both a Gram-negative and a Gram-positive organism. The system is comprised of targeted DNA cleavage by Cas9 and error-prone repair by Non-Homologous End Joining (NHEJ) proteins. I confirm previous results showing that Cas9 and NHEJ can make knockouts when NHEJ is expressed before Cas9. Then, I show that Cas9 and NHEJ can be used to make knockouts when expressed simultaneously. I term the new method Potentially Organism-Agnostic Knockout (POAK) system and characterize its function in *Escherichia coli* and *Weissella confusa*. First, I develop a novel transformation protocol for *W. confusa*. Next, I show that, as in *E. coli*, POAK can create knockouts in *W. confusa*. Characterization of knockout efficiency across *galK* in both *E. coli* and *W. confusa* showed that while all gRNAs are effective in *E. coli*, only some gRNAs are effective in *W. confusa*, and cut site position within a gene does not determine knockout efficiency for either organism. I examine the sequences of knockouts in both organisms and show that POAK produces similar edits in both *E. coli* and *W. confusa*. Finally, as an example of the importance of being able to make knockouts quickly, I target *W. confusa* sugar metabolism genes to show that two sugar importers are not necessary for metabolism of their respective sugars. Having demonstrated that simultaneous expression of Cas9 and NHEJ is sufficient for making knockouts in two minimally related bacteria, POAK represents a hopeful avenue for making knockouts in other under-utilized bacteria.

2.2 Introduction

Targeted knockouts are the cornerstone of genetics, and the ease of obtaining knockouts can determine how well an organism is studied[58,59]. This has been illustrated in mammalian systems by the messianic reception given to CRISPR/Cas9[60]. However, while it used to be extremely difficult to generate knockouts for many eukaryotic systems, knockout techniques are still limiting in the majority of known bacteria (and prokaryotes more generally)[61]. Despite the fact that CRISPR/Cas9 and host DNA repair can be used to make knockouts in eukaryotic systems without any additional pieces, this is not the case in most bacteria.

Double Stranded Breaks (DSB) in DNA can be repaired by two mechanisms: Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ)[62,63]. HR repairs DSBs using an unbroken template. It duplicates the unbroken template exactly, resulting in error-free repair. HR can be used to generate knockouts or make modifications to the chromosome, but this requires supplying a modified DNA sequence for the cell to use as a template (Fig. 1.1). NHEJ, on the other hand, takes two ends of DNA and glues them together. NHEJ systems are generally composed of two proteins, Ku and Ligase D (LigD). Ku binds to free DNA ends as a hexamer and recruits LigD, which ligates two DNA ends together (Fig. 2.1A). If this occurs immediately after a DSB, the repair may be error-free. However, if there is any damage to the ends of the DNA, such as exonuclease chew back, NHEJ repair will be error prone. As such, a DSB plus NHEJ can result in knockouts. Up until recently, this method was not practical in any organism as it was prohibitively difficult to make targeted DSBs. CRISPR/Cas9 has alleviated the problem of making target DSBs. Unfortunately, unlike eukaryotes, most bacterial species do not have NHEJ[61].

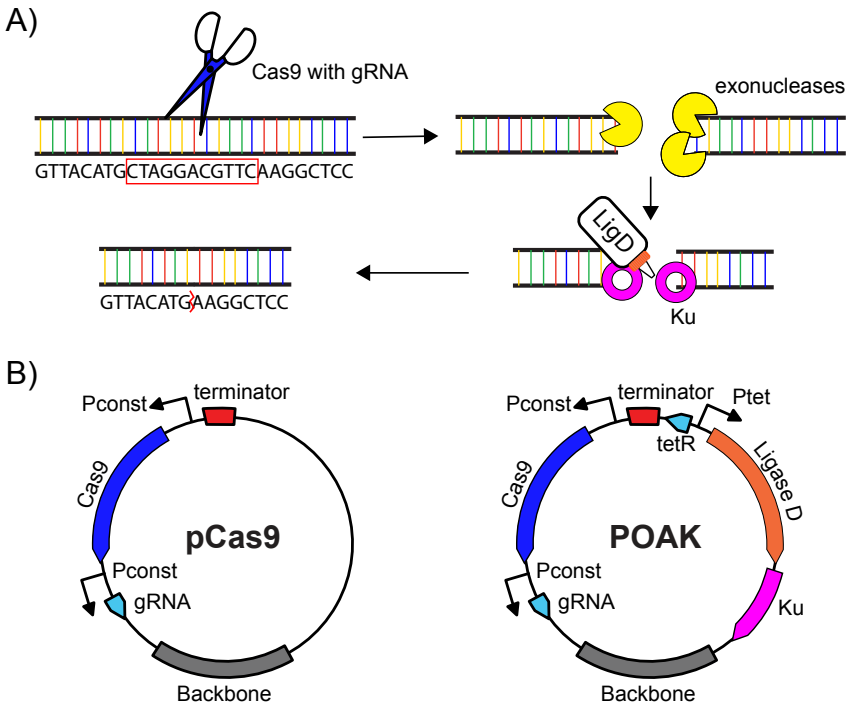


Figure 2.1: **Molecular biology and plasmid maps of pCas9 and POAK A)** The molecular biology of NHEJ. A double strand DNA break (DSB) occurs, potentially caused by Cas9 (with gRNA) cutting at a target site. Exonucleases may chew back the ends of the DNA before Ku binds to the ends of the DNA and recruits Ligase D (ligD), which ligates the ends of the DNA. If chew back occurs, the process results in a deletion. **B)** Plasmid maps of pCas9 and POAK.

As a result, HR mediated approaches have been the only approach for targeted knockouts in most bacteria, with other strategies being even more limited (Table 2.1). In most organisms, however, HR is inefficient and requires long homology arms. In permissive species, a minimum of 500 base pairs may be used, but in many species thousands of base pairs may be necessary[15]. In general, this means making knockouts is very difficult. Certain systems, such as the Lambda Red system in *Escherichia coli*, have greatly decreased the length of homology needed and increased the efficiency of recombination[16]. However, these single strand recombination technologies are derived from phages and are species specific. As a result, they are time consuming to develop and not portable to new organisms.

Table 2.1: **Comparison of knockout technologies**

	Sequence Specific	Broad Host Range	Efficient	Markerless	Multiple Knockouts
Mutagenesis Screens[5]		✓	✓	✓	✓
Transposon Mutagenesis[12,13]		✓	✓		
Homologous Recombination[15]	✓	✓		-	-
Lambda Red[16]	✓		✓	-	-
Su *et al.*[64]	✓		-	✓	✓
POAK	✓	✓	✓	✓	✓

A ✓indicates the technology has the property, a - that it has the property either under specific conditions or only with additional components, and an empty box indicates that the technology does not possess the property.

These difficulties are what have made the use of CRISPR/Cas9 so appealing for use in eukaryotic cells (which have NHEJ). When Cas9 and a gRNA are introduced into the cell, they create a targeted DSB. If repaired through an error free repair mechanism, such as HR, the target sequence can simply be re-cleaved and the cell either dies (if Cas9 cleavage is significantly more efficient than HR) or survives with a wild type sequence (if HR is more efficient). If, instead, the repair mechanism is error prone, such as with NHEJ, the DSB can be repaired with errors that mutate the target sequence and therefore prevent further cleavage (Fig. 2.1A). These errors are generally deletions (due to exonuclease activity), which makes Cas9 with NHEJ an effective knockout generation system.

For this to work in most bacteria, NHEJ must be provided *in trans*. Significant progress has been made towards making large deletions in *E. coli* by expressing both Cas9 and NHEJ heterologously[64]. That system, however, depends on *E. coli* specific technologies, such as multiple characterized plasmids, that are not available for many bacteria of interest. As such, I sought to interrogate Cas9 paired with NHEJ and determine whether they could be

combined into a system that could make knockouts in arbitrary bacteria. Minimally, such a system would have the following features.

1. Both Cas9 and the NHEJ proteins should function in a wide variety of organisms
2. The system should function when combined on a single plasmid
3. The system should be able to create knockouts when expressed simultaneously
4. The system should be easily modifiable to work in any desired organism

Cas9 has been shown to be functional in a wide variety of organisms, both prokaryotic and otherwise[65]. Bacterial NHEJ systems have been less extensively studied[66–68]. Fortunately, NHEJ systems are present in a diverse group of bacteria, from *Mycobacterium tuberculosis* to *Pseudomonas aeruginosa*, and some systems have been used for engineering in their native context[69,70]. Relatively little work has been done to express these systems heterologously, but one study showed NHEJ was able to circularize transformed linear DNA in *E. coli*[71] and another *E. coli* study showed NHEJ can repair DSBs caused by Cas9[64].

In this latter study, Su *et al.* demonstrated that Cas9 and NHEJ could be used in *E.coli* to create knockouts, particularly large deletions (Table 2.1). However, their system depends on NHEJ being expressed in cells prior to expression of the Cas9 and gRNA. In turn, the system also requires the use of multiple plasmids. These features are problematic for use in other bacteria, which often have a minimal number of characterized plasmids[72–74].

After confirming the dual plasmid results from Su *et al.*, I sought to understand whether Cas9 and NHEJ were functional in *E. coli* when expressed from a single plasmid. I also tested Cpf1, a Cas9 like protein that creates a different type of DSB, to see if the type of DSB affects the efficiency of NHEJ. I used a constitutively expressed Cas9, and this had the practical result of also testing whether Cas9 and NHEJ could create knockouts when expressed simultaneously. I then attempted to make knockouts in an organism that had been minimally characterized. For this task, I chose *Weissella confusa*.

W. confusa is a Lactic Acid Bacteria (LAB) in the *Leuconostocaceae* family[75,76]. Species in the *Weissella* genus have been studied for biotechnological applications, but these investigations have been limited by the lack of genetic tools. *W. confusa*, in particular, has been transformed, but has had minimal further characterization[77]. I used *W. confusa* as a test case for the organism agnosticism of my single plasmid knockout system.

In both *E. coli* and *W. confusa*, I characterized the effectiveness of POAK, the nature of the knockouts POAK produced, as well as whether the POAK plasmid was stable. As an example of the value of exploring new organisms, I used the knockouts to examine sugar metabolism in *W. confusa*. I hope that POAK will help lower the barrier to making sequence specific, markerless knockouts in arbitrary organisms.

2.3 Results

2.3.1 Cas9 and NHEJ in *E. coli*

Cas9 and NHEJ were confirmed to be functional in *E. coli*. Due to the limited work that had been done to examine targeted cutting by Cas9 and repair by NHEJ in *E. coli*, I first sought to confirm previous results showing Cas9 could create knockouts in *E. coli* that already expressed the NHEJ proteins. I additionally tested Cpf1. Whereas Cas9 creates blunt end DSBs, Cpf1 creates DSBs with sticky ends, which could conceivably impact the behavior of the NHEJ system. I transformed *E. coli* with either an empty vehicle plasmid (pVeh) or a plasmid encoding constitutively expressed NHEJ (pNHEJ), and subsequently transformed in Cas9 (pCas9sp) or Cpf1 (pCpf1sp) plasmids that contained zero, one, or two gRNAs. Relative survival and knockout efficiency were measured (Fig. 2.2). In the absence of NHEJ, Cas9 and Cpf1 are highly lethal (Supp. Fig. A.1A and B). When NHEJ is present, it significantly increases the survival of *E. coli* transformed with the nucleases plus gRNA (Supp. Fig. A.1A and B). To assay for knockout efficiency, gRNAs were targeted to *galK*, a gene essential for galactose metabolism. When plated on MacConkey-galactose

agar, colonies with *galk* turn red, while colonies without *galK* turn white. Cas9 or Cpf1 transformed with gRNAs creates a significant percentage of knockouts in an NHEJ strain (Supp. Fig. A.1A and B), confirming the work of Su *et al.*.

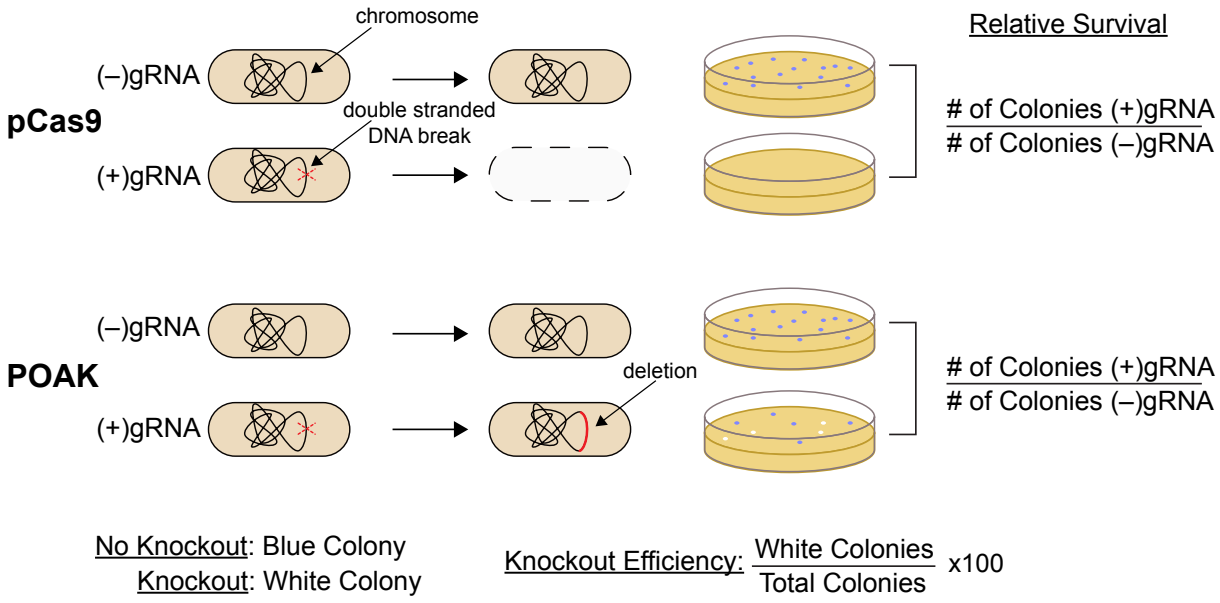


Figure 2.2: **POAK experimental design** Schematic of the experimental design for testing pCas9 and POAK constructs. Transformation efficiency of pCas9/POAK with gRNA is compared to without gRNA to determine relative survival. Colony phenotype, in this illustration blue vs white colonies, is used to determine knockout efficiency.

POAK was made by combining Cas9 and NHEJ on a single plasmid (Fig. 2.1B). To make Cas9 and NHEJ into a Potentially Organism-Agnostic Knockout (POAK) system, I first combined Cas9, a CRISPR array (for gRNA production), and NHEJ onto a single plasmid. Two versions were created-one for Gram-negative and one for Gram-positive bacteria. To construct the Gram-negative plasmids, a CRISPR array with BsaI sites for gRNA and Cas9 from *Streptococcus pyogenes* (or Cpf1 with its CRISPR array) were combined onto plasmids, and then that plasmids was combined with Gram-negative broad host range origin of replication (Bbr1)[78] and kanamycin resistance (KnR) to form pCas9 and pCpf1. NHEJ was then placed into the plasmid under tetR control (from the Tn10 transposon)[79] using two codon optimized gBlocks to create gnPOAK and gnPOAK_Cpf1 (Fig. 2.1B). To create plasmids for use in Gram-positive organisms, the Bbr1 and KnR on pCas9 and gnPOAK

were exchanged for a backbone that contained a broad host range Gram-positive origin of replication[80], the colE1 origin of replication, and erythromycin resistance[81], to create the intermediate plasmids pCas9temp and POAKtemp. Finally, the promoter for Cas9 in pCas9temp and POAKtemp was replaced with the 200 bp upstream of the *W. confusa* enolase gene ($P_{wc\text{-}eno}$), resulting in plasmids pWcCas9 and gpPOAK. These final two plasmids, as well pCas9, pCpf1, gnPOAK and gnPOAK_Cpf1, were used in further experiments. Table 2.2 lists these plasmid backbones as well as what they are used for in this work. gRNAs were added to these plasmids as needed, and are referred to by the gene name they target followed by a letter representing the organism as well as a number (e.g. chbCe1, xylAw1). gRNA names are not italicized.

Table 2.2: **Plasmid backbones**

Plasmid	Use	Description
pCas9	Cas9 vector with gRNA cloning site for use in *E. coli*	PproC(Cas9), pBBR1 origin, KnR
gnPOAK	POAK vector with gRNA cloning site for use in *E. coli*	PproC(Cas9), Pteta(ligd, ku), pBBR1 origin, KnR
pCpf1	Cpf1 vector with gRNA cloning site for use in *E. coli*	PproC(Cpf1), pBBR1 origin, KnR
gnPOAK_Cpf1	POAK_Cpf1 vector with gRNA cloning site for use in *E. coli*	PproC(Cpf1), Pteta(ligd, ku), pBBR1 origin, KnR
pCas9temp	Precursor to pWcCas9	PproC(Cas9), pBAV1K backbone
gpPOAKtemp	Precursor to gpPOAK	PproC(Cas9), Pteta(ligd, ku), pBAV1K backbone
pWcCas9	Cas9 vector with gRNA cloning site for use in *W. confusa*	Pwc-eno(Cas9), pBAV1K backbone
gpPOAK	POAK vector with gRNA cloning site for use in *W. confusa*	Pwc-eno(Cas9), Pteta(ligd, ku), pBAV1K backbone

The plasmid backbones used in the main text of the study. gRNAs are appended when they are added to the backbone. So, the pWcCas9 backbone with the galKw1 gRNA would be pWcCas9galKw1. Full lists of the plasmids, strains, and gRNAs used in this study can be found in Supplemental Material and Methods.

NHEJ expressed simultaneously with Cas9 rescues *E. coli* and results in knockouts. To test whether NHEJ expressed at the same time as Cas9 was sufficient to reduce Cas9 lethality and create knockouts, I transformed gnPOAK and gnPOAK_Cpf1 into *E. coli* either with or without a gRNA targeting *galk*. NHEJ is capable of rescuing *E. coli* from Cas9 targeting, but not from Cpf1 targeting (Fig. 2.3A). Likewise, gnPOAK creates significantly more knockouts than pCas9, while gnPOAK_Cpf1 does not create significantly more knockouts than pCpf1 (Fig. 2.3B). The relative survival and knockouts efficiency are significantly lower for the gnPOAK system than the dual plasmid system. Further, there is high variability in both survival as well as knockout efficiency in the single plasmid system.

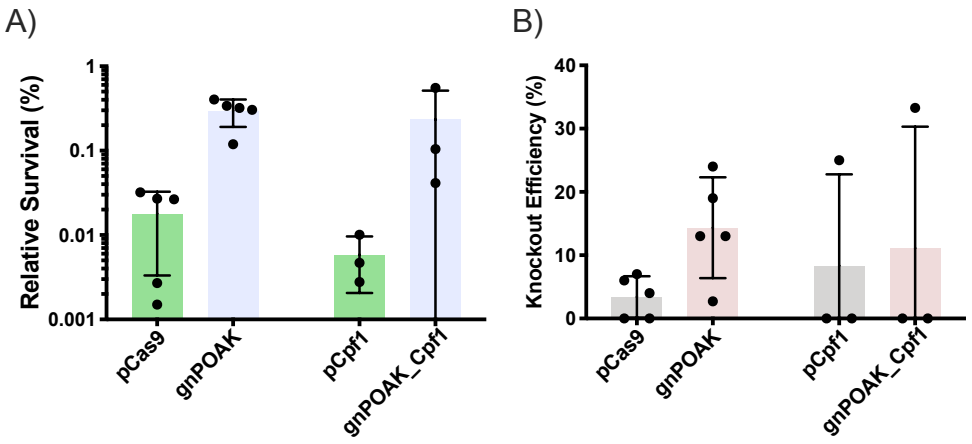


Figure 2.3: **NHEJ increases survival and knockouts from DSBs caused by Cas9 but not Cpf1 in *E. coli*** pCas9, pCpf1, gnPOAK, and gnPOAK_Cpf1 were transformed into *E. coli* with galKe9 (for pCas9 and gnPOAK) or galKe1F (for pCpf1 and gnPOAK_Cpf1). **A)** Relative survival of transformations of pCas9galKe9 and gnPOAKgalKe9 (n=6) and of pCpf1galKe1F and gnPOAK_Cpf1galKe1F (n=3). Bars are the mean, and errors bars represent the standard deviation. **B)** Knockout efficiency of transformations of pCas9galKe9 and gnPOAKgalKe9 (n=6) and of pCpf1galKe1F and gnPOAK_Cpf1galKe1F (n=3). Bars are the mean, and errors bars represent the standard deviation.

2.3.2 Cas9 and NHEJ in *W. confusa*

I found that *W confusa* can be transformed with high efficiency using a simple electroporation protocol. The only previously published transformation protocol for *W. confusa* involves making protoplasts through enzymatic digestion of the cell wall[77]. Here I developed a *W.*

confusa electrocompetent preparation as a room temperature (RT) protocol, although certain steps showed sensitivity to overheating (Fig. 2.4A). During the wash steps, if cells were spun down in a non-temperature-controlled centrifuge that was being heavily used, transformations routinely failed (data not shown). I suspect this is due to overheating, and in these cases the cell pellets were looser and significant debris remained in the media, possibly indicating cell lysis. This is supported by *W. confusa's* rapid death upon heat shock (Supp. Fig. A.2). The electroporation itself was also sensitive to overheating. Cells electroporated in ice cold (0 °C) cuvettes showed significantly higher transformation rates than cells electroporated in RT cuvettes (Fig. 2.4B). Samples electroporated in RT cuvettes had released significant amounts of genomic DNA (which could be observed by eye during the rescue step of the transformations), suggesting lysis occurred during electroporation.

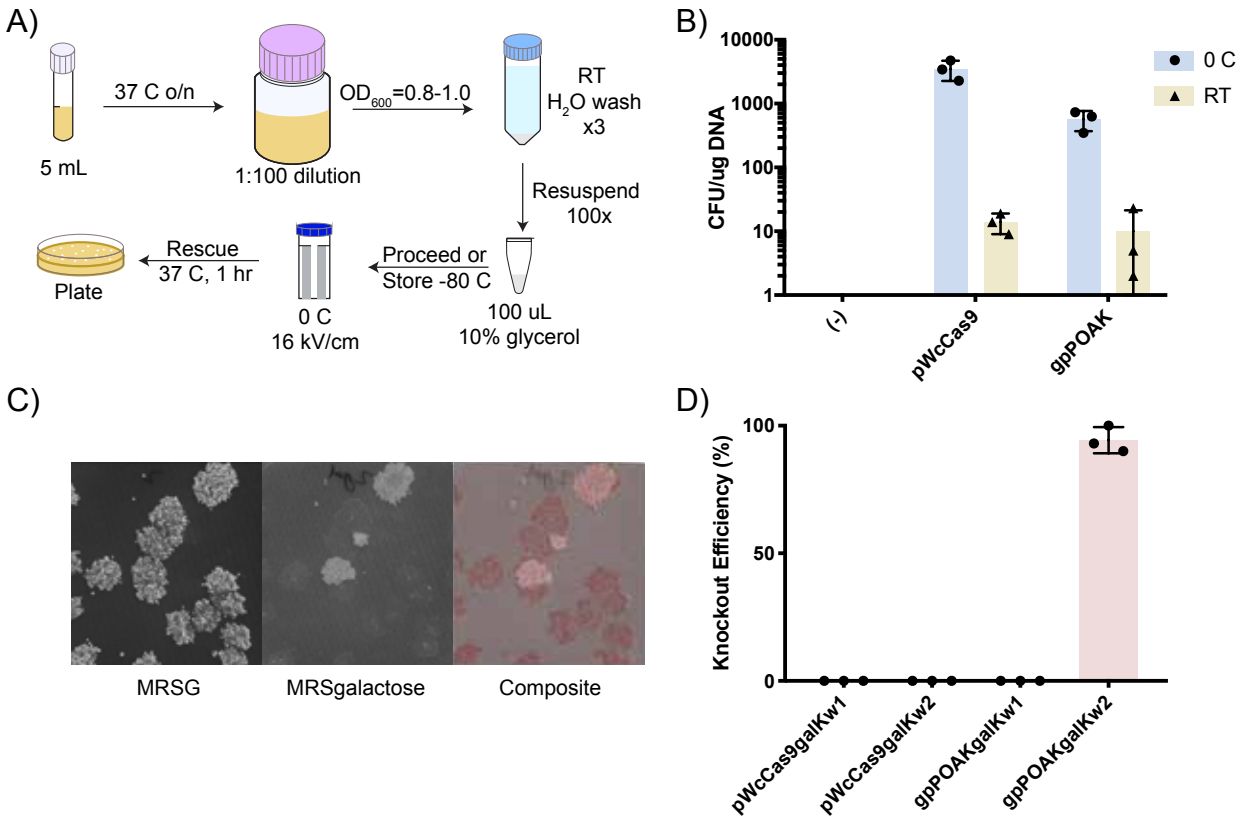


Figure 2.4: **Transformation and knockouts using POAK in *W. confusa* A)** Schematic of transformation protocol for *W. confusa*. An overnight culture is back-diluted 1:100 and grown until OD_{600}≈0.8-1. The culture is washed three times with RT water, resuspended at 100x concentration, then electroporated at 16 kV/cm. **B)** Transformation efficiency of electroporation in *W. confusa* with two different plasmids, pWcCas9 and gpPOAK, and at two different cuvette temperatures, 0 °C and RT. Bars show the mean of three transformations and error bars represent the standard deviation. **C)** Knockouts after transformation of *W. confusa* with gpPOAKgalKw2. Replica plating of a representative transformation onto MRSG, MRSgalactose, and a composite of the two is shown. **D)** Knockout efficiency after transformation with pWcCas9galkw1, gpPOAKgalkw2, pWcas9galkw2, and gpPOAKgalkw2. Bars are the mean of three transformations and error bars show standard deviation.

POAK can make knockouts in *W. confusa*. To test whether gpPOAK could make knockouts in *W. confusa*, I designed gRNAs targeted either 100 bp (galKw1) or 200 bp (galKw2) downstream of the start codon in the *W. confusa galK* gene. A knockout in *galK* should result in a loss of galactose metabolism. To determine if transformants could metabolize galactose, I developed a replica plating technique for *W. confusa*. In short, this involved replica plating transformation plates onto a second set of selective plates (to mimic re-

streaking of colonies) then replica plating the second set of selective plates onto plates selective for galactose metabolism (MRS media with galactose, MRSgalactose) and non-selective for galactose metabolism (MRS media with glucose, MRSG). This allows for assaying of sugar metabolism knockouts (Fig. 2.4C). gpPOAK's ability to make knockouts is gRNA dependent and NHEJ dependent in *W. confusa* (Fig. 2.4D). To test this, I transformed pWcCas9galKw1, pWcCas9galKw2, gpPOAKgalKw1, and gpPOAKgalKw2 into *W. confusa*. Of these, only gpPOAKgalKw2 created observable knockouts, suggesting that both NHEJ and specific gRNAs are necessary.

pWcCas9 does not efficiently cut the *W. confusa* genome. To test the efficiency of Cas9 in *W. confusa*, the relative survival of pWcCas9 with galKw1 or galKw2 was assayed (Supp. Fig. A.3). Despite leading to knockouts when used with gpPOAK, neither gRNA decreased the survival of *W. confusa* below 10% whether used in pWcCas9 or gpPOAK. In contrast, Cas9 with gRNA efficiently reduces survival in *E. coli* (Fig. 2.3A).

2.3.3 Comparison of POAK Behavior in *E. coli* and *W. confusa*

POAK knockouts can be made by targeting any part of a gene in *E. coli* and *W. confusa*. To test the effect of cut location on knockout efficiency, I transformed *E. coli* and *W. confusa* with POAK plasmids that targeted regions approximately 100 bp apart along each organism's respective *galK* sequence. The percentage of colonies that had a $\Delta galK$ phenotype was assayed by MacConkey-Galactose plates or replica plating on MRSgalactose for *E. coli* and *W. confusa* respectively. Knockouts occurred for gRNAs targeted across the gene in both organisms, although all *E. coli* gRNAs produced knockouts (Fig. 2.5A) while only ∼50% of *W. confusa* gRNAs produced knockouts (Fig. 2.5B). For gRNAs that produced knockouts, there is significant variability in the knockout efficiency, ranging from 10% to 40% in *E. coli* and from 4% to 95% in *W. confusa*. The highest knockout efficiencies do not cluster around known GalK actives sites.

POAK knockout efficiency is correlated to Cas9 efficiency in *W. confusa* but not in *E. coli*.

To determine why all of the *E. coli* gRNAs produced knockouts but not all of the *W. confusa* gRNAs did, I compared the relative survival to knockout efficiency for all transformations. The relative survival of *E. coli* when transformed with gnPOAK bares no relationship to the number of knockouts that will be produced (Fig. 2.5C). *W. confusa*, however, has a clear relationship between relative survival and knockout efficiency, with no transformation that had above 50% survival producing knockouts. Survival rates were not correlated with the location of the cut site within the gene for either *E. coli* or *W. confusa*, but individual gRNAs do show similar levels of relative survival across the three transformations (Supp. Fig. A.4).

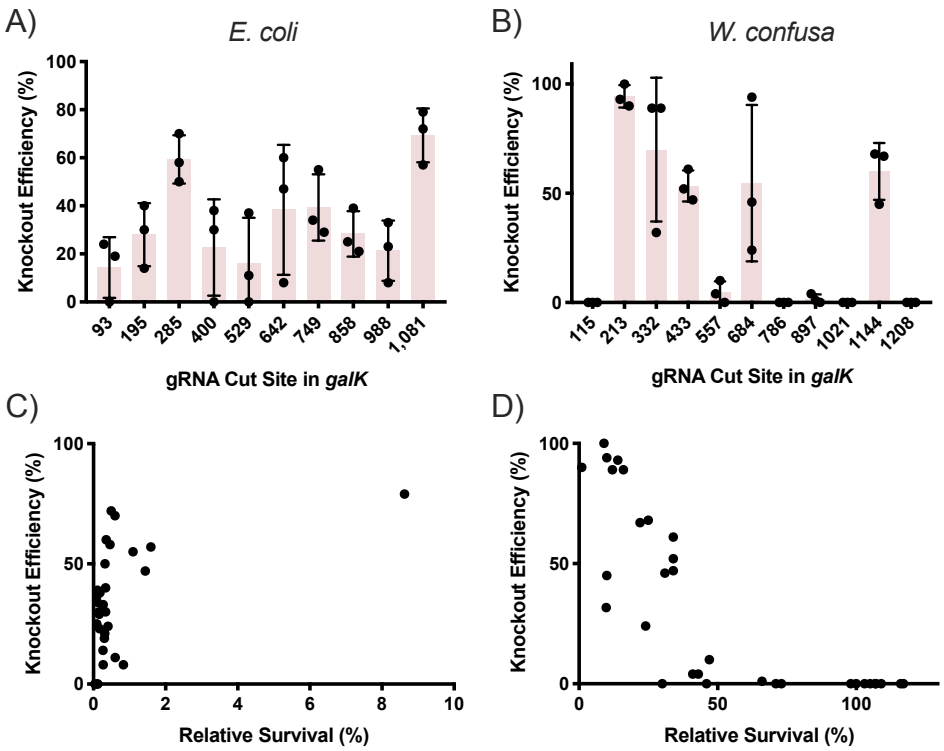


Figure 2.5: **Effect of cut position and relative survival on POAK knockouts in *E. coli* and *W. confusa*** **A)** Knockout efficiency observed on MacConkey+galactose plates when *E. coli* is transformed with gnPOAK containing a gRNA that cuts at the indicated position of *galK*. Bars are the mean of three transformations, error bars represent the standard deviation. **B)** Knockout efficiency observed by replica plating when *W. confusa* is transformed with gpPOAK containing a gRNA that cuts at the indicated position of *galK*. Cut sites 115 bp and 213 bp are gRNAs galKw1 and galKw2 respectively. Bars are the mean of three transformations, error bars represent the standard deviation. **C)** Knockout efficiency vs. relative survival of all transformation in **A**. **D)** Knockout efficiency vs. relative survival of all transformation in **B**.

To understand the type of knockouts that POAK creates I targeted five different sugar metabolism genes from the *E. coli* and *W. confusa* genomes (Table 2.3). Briefly, all surviving colonies from each transformation were pooled and genomic DNA collected. The 4 kbp region surrounding the gRNA cut site was amplified and sequenced using an NGS pipeline. Sequence deletions were observed for all targeted *E. coli* genes. In *W. confusa*, sequence deletions were observed in *celbPTSIIC*, *manPTSIIC*, and *xylA*. No deletions were observed in *maltP*, and *galK* deletions were observed only from the galKw2 gRNA, and not from galKw1.

Table 2.3: **Sugar metabolism genes targeted for NGS experiment**

	Gene	Expected Function	gRNA
E. coli			
	chbC	cellobios importer	chbCe1
	galk	galactose kinase	galKe1
	lacZ	beta-galactosidase (lactose metabolism)	lacZe1
	manY	mannose importer	manYe1
	xylA	xylose isomerase	xylAe1
W. confusa			
	celbPTSIIC1	cellobios importer	celbPTSIICw1
	galk	galactose kinase	galKw1
			galKw2
	maltP	maltose phosphorylase	lacZw1
	manPTSIIC	mannose importer	manPTSIICw1
	xylA	xylose isomerase	xylAw1

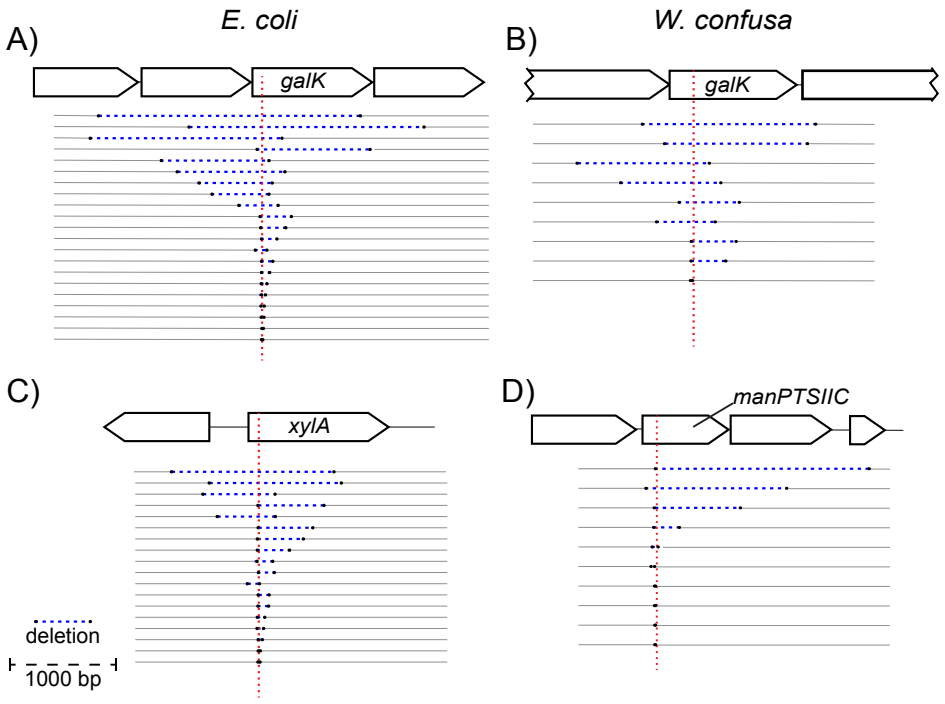


Figure 2.6: **Representative editing by gnPOAK in *E. coli* and of gpPOAK in *W. confusa*** Each horizontal line represents an observed deletion supported by more than five reads. All deletions from three different transformations are shown in each panel. Deletions are indicated by black dots connected by a dotted blue lines, and the Cas9-gRNA cut site is represented by a dotted red line. 1000 bp shown for scale. **A)** Transformations of gnPOAKgalKe1 in *E. coli*. **B)** Transformations of gpPOAKgalKw2 into *W. confusa* **C)** Transformations of gnPOAKxylAe1 in *E. coli* **D)** Transformations of gpPOAKmanPTSIICw1 in *W. confusa*

POAK produces a range of deletions in *E. coli* and *W. confusa*. Fig. 2.6 shows all observed deletions from three replicates of *galK* and *xylA* genes in *E. coli* and *galK* and *manPTSIIC* in *W. confusa*. Each targeted gene shows a range of deletions from as small as 7 bp to up to 2500 bp (sample preparation may have precluded observation of significantly larger deletions). Most deletions occur bi-directionally around the cut site. However, in some cases, such as gpPOAKmanPTSIICw1 (Fig. 2.6D), observed deletions do occur primarily in one direction. This was also observed in gpPOAKxylA1 for individual replicates (Supp. File A.4). There are more observed deletions for *E. coli*, which is consistent with the larger number of recovered transformants. In *W. confusa*, deletions were observed in *galK* from the gRNA galKw2, but not from galKw1, which is consistent with the knockout data (Fig.

2.4B).

Knockouts of the putative cellobios and mannose importers do not prevent cellobios or mannose metabolism in *W. confusa*. In *E. coli* all gRNAs created sequence deletions and also resulted in loss of sugar metabolism (except for *chbC*, which could not be assayed) (Supp. Fig. A.5B). In *W. confusa*, of the four gRNAs that produced observable sequence deletions, only galKw2 resulted in colonies that were completely deficient for metabolism of the relevant sugar (Fig. 2.4C). xylAwI produced segmented colonies at low rates (Fig. 2.7B), which is consistent with the gRNA's high rates of survival (Fig. 2.7A). To determine whether or not the deletions observed through NGS in *celBPTSIIC* and *manPTSIIC* were preventing *W. confusa* from metabolizing those sugars, knockouts for both genes were re-isolated. Knockouts were verified by colony PCR and Sanger sequencing (Supp. Fig. A.6). The knockouts both grew when their respective sugar was the sole carbon source (Fig. 2.7C).

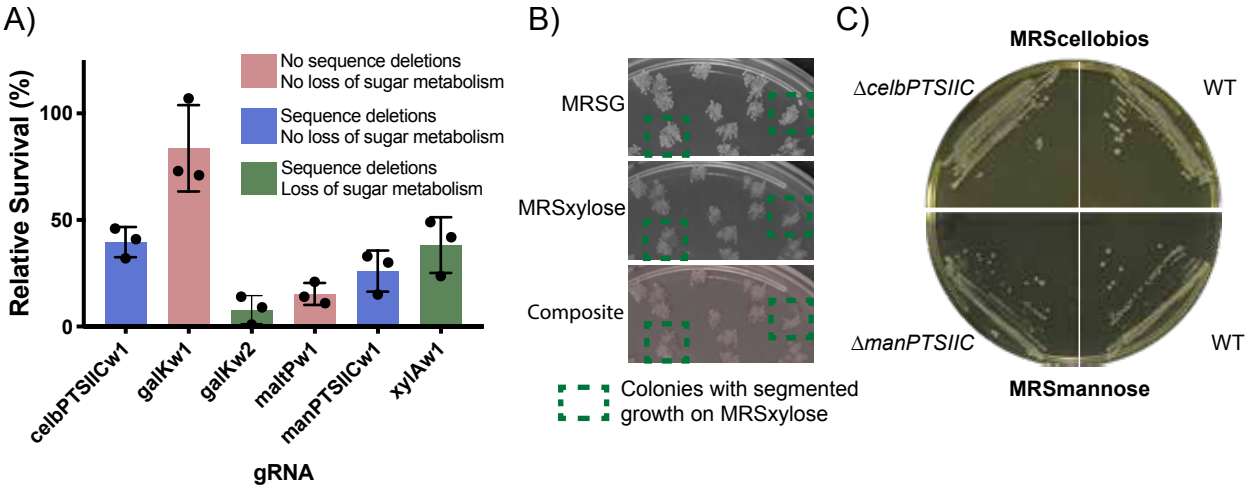


Figure 2.7: **Two genetics knockouts from the *W. confusa* NGS screen do not result in loss of sugar metabolism A)** Relative survival for five targeted *W. confusa* genes. Blue bars indicate genes for which deletion mutants were observed in sequencing, but no loss of sugar metabolism was observed after replica plating. Green bars are for genes that showed both sequence deletions and loss of sugar metabolism. Red bars indicate that neither sequence deletions nor loss of sugar metabolism was observed. Bars are the mean of three transformations, error bars represent the standard deviation. **B)** Replica plating of a xylAw1 transformation. Colonies that showed segmented xylose-dependent survival are highlighted. **C)** Knockouts in *celBPTSIIC* and *manPTSIIC* genes grow on plates that require cellobios metabolism (MRScellobios) and mannose metabolism (MRSmannose), respectively. WT *W. confusa* shown on each type of plate.

POAK plasmids rapidly cure from both *E. coli* and *W. confusa*. The components of NHEJ can be toxic to *E. coli* when over-expressed. This appears to due to the effects of Ku, which cause an induction dependent growth defect (Supp. Fig. A.7B and D), whereas induction of LigD shows much more subtle effects (Supp. Fig. A.7A and C). To test whether this would lead to curing of the plasmid, POAK and Cas9 plasmids were grown in *E. coli* and *W. confusa* in the presence of aTc (which induces the NHEJ proteins) but without selection. Plasmids with NHEJ were cured from over 99% of *E.coli* cells and greater than 90% of *W. confusa* cells (Supp. Fig. A.8). In *E. coli*, the Cas9 only plasmid was stable during overnight growth, while in *W. confusa* the Cas9 plasmid was lost almost as rapidly as the POAK plasmid. This suggests that while in *E. coli* plasmid instability is caused by NHEJ proteins, in *W. confusa* the plasmid instability is due to the plasmid backbone and not the NHEJ components.

2.4 Discussion

In the introduction, I laid out four conditions that are required for a knockout system to work in arbitrary bacteria. Of these, the outstanding questions were whether Cas9 and NHEJ could create knockouts when expressed simultaneously, and how difficult it would be to modify a Cas9 and NHEJ expression system to work in a chosen organism. In this work, I have shown that Cas9 and NHEJ can create knockouts in both *E. coli* and *W. confusa* when expressed simultaneously. Further, only minimal modifications were required to get the system working in *W. confusa*.

Replacing the promoter for Cas9 was the only species-specific change required for POAK to work in *W. confusa*. To create gpPOAK, the backbone of gnPOAK was changed to an origin of replication (from pBavIK) and antibiotic cassette known to work in Gram-positive baceria. The backbones of both POAK vectors are considered "broad host range", so these should only require occasional adjustment. The only species-specific adjustment I made to

gpPOAK was to add $P_{wc\text{-}eno}$ (the 200 bp upstream of the *W. confusa* enolase gene) in front of Cas9. The choice of the enolase promoter was based on data suggesting enolases are often highly expressed in bacteria, not on *W. confusa* specific data. This change was sufficient to produce levels of Cas9 capable of making knockouts. However, it is worth noting that not all gRNAs work in *W. confusa*, while all gRNAs work in *E. coli*. Further, Cas9 and POAK with gRNAs both show much lower rates of survival in *E. coli* than in *W. confusa*. This is not surprising, considering that both the gRNA promoters and the Cas9 promoter used in pCas9 and gnPOAK were optimized for *E. coli*[82] while neither the gRNA promoters nor the Cas9 promoter in pWcCas9 or gpPOAK were particularly optimized for use in *W. confusa*. As such, it is likely that further optimization would increase the effectiveness of pWcCas9 and gpPOAK in *W. confusa*.

W. confusa is useful for biotechnological and laboratory use. *W. confusa* grows quickly, with a doubling time of 40 minutes at 37 °C (Supp. Fig. A.9), in both aerobic and anaerobic conditions. Unlike many other LAB and Gram-positive bacteria, *W. confusa* does not lyse in stationary phase. I show it can be transformed with a simple electroporation protocol and is robust to freezing and storage at −80 °C. Transformation efficiencies are high enough that I was able to use it as a cloning host for a plasmid that was toxic in *E. coli*. Overall, the organisms ease of use rivals that of *E. coli* and warrants investigation as Gram-positive chassis.

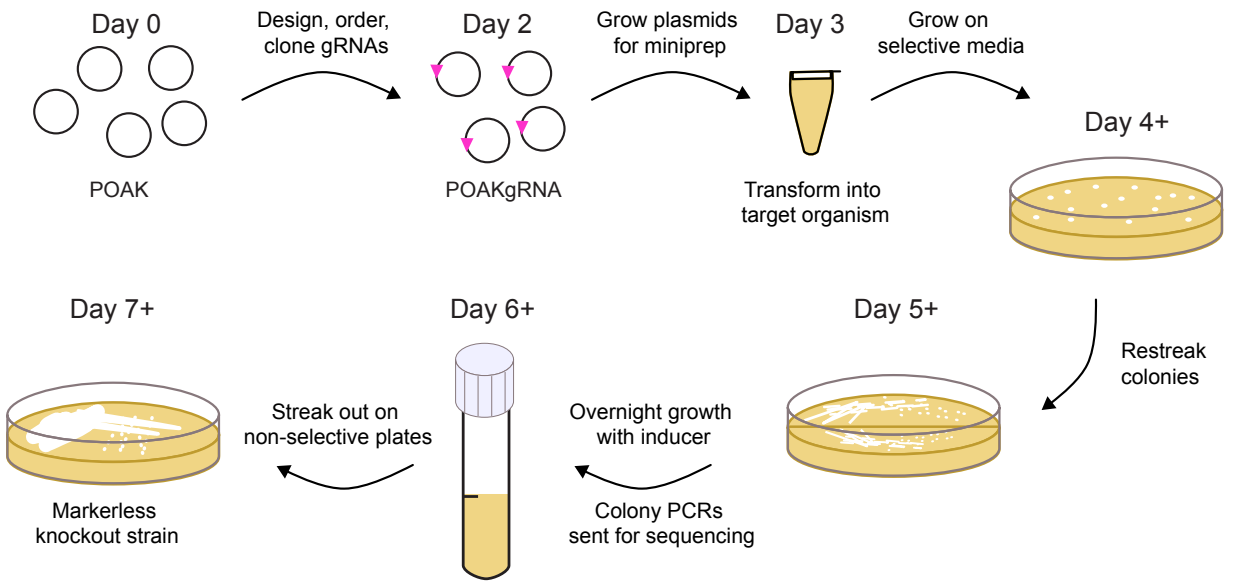


Figure 2.8: **POAKing out genes** From the time gRNAs are designed (**Day 0**) it takes three days until transformation of POAK into desired organism (**Day 3**). After the transformants have grown up (**Day 4**+) it takes three more days before markerless knockouts are obtained (**Day 7**+).

POAK can be used to go from design to sequenced markerless knockout in a week. Once POAK has been validated in an organism, knockouts can be made in under a week, with the marginal time for each additional knockout being about 4 days (Fig. 2.8). First, gRNAs are designed and ordered (Day 0), then cloned into the appropriate POAK vector and transformed into a cloning strain (Day 1). A colony PCR of the gRNA is sent to confirm the sequence of gRNA insertion and colonies are picked for overnight cultures (Day 2). Sequence validated clones are mini-prepped and transformed into the target organisms (Day 3). Surviving colonies are struck out on selective plates (Day 4+, depending on organism's growth rate). Colonies are sent in for sequencing of the target locus and colonies are grown overnight in aTc (to induce the NHEJ proteins) and without antibiotics (Day 5+). Sequence confirmed colonies are struck out on non-selective plates (Day 6+). Colonies can then be used as parent strains for further knockouts, although if used immediately, colonies should be picked into both selective and non-selective media to confirm loss of plasmid (Day 7+).

POAK creates a range of deletions. Unlike recombination-based techniques that create a set of uniform transformants, POAK creates a range of mutations. These mutations range from small mutations (SNPs and <10 bp deletions) that can be used for targeted editing, to large mutations (>2000 bp) that can be used to knockout multiple genes at once. A similar range of deletions is observed in both *E. coli* and *W. confusa*. Across multiple replicates, these deletions generally occur symmetrically around the cut site. In certain *W. confusa* replicates, such as xylAw1, there appeared to be asymmetry in the deletion (i.e. most of the deletions primarily include portions either upstream or downstream of the cut site). However, this effect did not occur consistently, and as such I cannot draw any strong conclusions about whether the directionality is a real effect or the conditions under which it might occur.

The deletions created by POAK are markerless, and therefore can be used in applications where downstream repression (e.g. CRISPRi) or the presence of antibiotic cassettes (e.g. HR) are not desirable. This makes POAK ideal for targeting a single gene in an operon, or in organisms for which only a single antibiotic is routinely used. This contrasts with homologous recombination strategies that require additional technologies (such as recombinases) to be rendered markerless. POAK nicely complements CRISPRi, a technique that is based on dCas9 and has been applied in a range of bacteria[18]. Unlike POAK, CRISPRi can create knockdowns of entire operons. The combination of the two techniques has the potential to allow robust genetic interrogation of heretofore recalcitrant organisms.

In this study, I have expanded on previous work that hinted that Cas9 and NHEJ might be capable of creating knockouts in a wide range of bacteria. I showed that simultaneous expression of Cas9 and NHEJ creates knockouts in *E. coli* and in *W. confusa*. The efficiency of making knockouts appears to depend on the efficiency of Cas9 cutting, a feature that can be tuned based on the organism. Regardless of the efficiency of knockout, the resulting sequence edits are similar between both organisms. POAK plasmids are quickly cured in both organisms as well. Taken together, this suggests that I have made progress towards a markerless, sequence specific, broad host range, knockout system. It is my hope that this

Potentially Organism Agnostic Knockout (POAK) system will be taken by the community and further developed into a true Prokaryotic Organism Agnostic Knockout (POAK) system.

2.5 Materials and Methods

Strains and Growth Conditions

Complete strain information can be found in Supplemental Table A.1. DH10β was used as the default cloning strain, and was routinely grown in Lysogeny Broth (LB) Miller and on LB agar plates. For the Gram-positive shuttle plasmid, DH10β was grown in 2xYT broth and on Blood Heart Infusion (BHI) agar. Single gene knockouts from the Keio[83] collection with the antibiotic resistance removed were used for cloning of plasmids that contained gRNAs targeted to the *E. coli* genome. Experiments were performed in MG1655 (*E. coli K12*) or in DSM 20196 (*W. confusa*). MG1655 was routinely grown in LB and on LB plates. *W. confusa* was grown in Man, Rogosa and Sharpe (MRS)[84] media with glucose (MRSG) broth and on MRSG plates. All MRS plates were prepared by autoclaving the media *without* sugar, and supplementing filter sterilized sugar to two percent final concentration afterwards.

For *E. coli* galactose, mannose, and xylose metabolism experiments, MacConkey[85] agar was used with the appropriate sugar supplemented to one percent. *E. coli* lactose metabolism was assayed using LB plates containing X-gal. MRS plates with the appropriate sugar substituted for glucose were used as selective plates for *W. confusa* sugar metabolism experiments. Initially MRS without yeast extract was used, so as to eliminate residual mannose, but the presence or absence of yeast extract was observed to be inconsequential, and so it was included in the MRS used for later replicates. For selective growth, spectinomycin at 100 µg/mL, kanamycin at 45 µg/mL, and erythromycin at 150 µg/mL were supplemented for *E. coli*, and erythromycin at 10 µg/mL was supplemented for *W. confusa*.

Plasmid Construction

A complete list of plasmids can be found in Supplmental Table A.2 and plasmid maps can be found in Supplemental File A.1. Golden Gate Assembly[86] was used for cloning plasmids and inserting gRNAs. For all cloning besides insertion of gRNAs, Q5 Hot Start polymerase was used for amplification of assembly pieces. DNA for the coding sequences of LigD and Ku were ordered as gBlocks from IDT. For Golden Gate reactions, 10x T4 Ligase Buffer (Promega), T4 Ligase (2,000,000 units/mL, NEB), and BSA (10 mg/mL, NEB) were used in all reactions. The appropriate restriction enzyme, either Eco31I, Esp3I, or SapI (Thermo FastDigest), was added. gRNAs were added to plasmids by Golden Gate after annealing and phosphorylating pairs of oligos. For detailed information, see Supplemental Materials and Methods. Briefly, complementary oligos were incubated with T4 Ligase Buffer (NEB) and T4 Poly Nucleotide Kinase (NEB), heated to boiling, and then slowly cooled to room temperature. Annealed oligos were then added to plasmids using Eco31I.

Golden Gate reactions were desalinated using drop dialysis (for a minimum of 10 minutes) and electroporated in DH10β Electrocompetent Cells (Thermo Fischer).

E. coli Electrocompetent Preparation and Transformations

E. coli was made electrocompetent using a modified standard protocol[87]. *E. coli* was grown in either LB or 2xYT until OD_{600} of between 0.4 and 0.6. Cells were spun down at 4000xG for 10 minutes at room temperature. Cells were then washed twice with 0.5x volume room temperature ddH_2O and then once with 0.1x volume room temperature ddH_2O. Cells were re-suspended in approximately 0.001x volume of room temperature ddH_2O (if competent cells were to be used immediately) or 10% room temperature glycerol (if cells were to be frozen and stored at $-80\,^{\circ}\mathrm{C}$). 25 µl of cells were transformed with 2.5 µl of 20 ng/µl of mini-prepped plasmid. Electroporations were performed using 0.1 cm cuvettes at 1.8 kV (Ec1) in a BioRad MicroPulser$^{\mathrm{TM}}$. Cells were rescued in 972.5 µl SOC supplemented with 200 nM aTc for 1 hour at 37 °C. Cells were then plated on agar plates with appropriate selection.

Occasionally, cells prepared using this method are too concentrated and arc, so a no-DNA control electroporation was always performed. If the pulse time was less than 5 ms, cells were diluted until an appropriate pulse time was achieved.

Agar plates were imaged using a macroscope. Colonies were counted manually, and metabolism of the relevant sugar was indicated by red colored colonies on MacConkey plates or blue colonies on LB+X-gal plates. White colored colonies on either media was indicative of no metabolism of the specific sugar.

Replicates are of at least three separate competent cell preparations, except for Supp. Fig. A.1 which is three transformations from the same electrocompetent preparation.

W. Confusa Electrocompetent Preparation and Transformations

W. confusa was made electrocompetent using a modified version of the above protocol. *W. confsua* was grown, either from a colony or from a 1:100 dilution of an overnight culture, in MRSG until OD_{600} of between 0.8 and 1.0 (often closer to 0.8, due to it being faster). Cells were spun down at 4000xG for 12 minutes at 22 °C.[†] Cells were then washed twice with 0.5x volume room temperature ddH_2O and then once with 0.1x volume room temperature ddH_2O. Cells were re-suspended in 0.01x volume of room temperature ddH_2O (if competent cells were to be used immediately) or 10% room temperature glycerol (if cells were to be frozen and stored at −80 °C). 100 µl of cells were transformed with 10 µL of 10 ng/µl of mini-prepped plasmid. Electroporations were performed using 0.2 cm ice cold cuvettes at 2.5 kV (Ec2) in a BioRad MicroPulser™. Cells were rescued using 900 µl MRSG supplemented with 200 nM aTc for 1.5 hours at 37 °C. Cells were then plated on MRSG agar plates with erythromycin and grown at 37 °C for two to three days (or until colonies had grown to an adequate size). It is important to use MRS media that has the sugar added after autoclaving. Plates that have been made from MRS autoclaved with sugar significantly reduce transformation efficiency.

Agar plates were imaged using a macroscope. Colonies were counted using the Cell

[†]It is important to use a *temperature-controlled centrifuge* for these steps, as a "room temperature" centrifuge will become too hot and cause reduced cell viability.

Colony Edge FIJI macro[88] that had been modified. Results were manually checked for accuracy. Macro is provided as Supp. File A.2. Replicates are of at least three separate competent cell preparations.

Frozen competent cells remain viable for at least nine months (and likely much longer). To use, thaw on ice and proceed as described above.

W. confusa Replica Plating

W. confusa was replica plated using either a sterile felt or two to three paper towels from the inside of an unopened stack. Colonies were first replica plated onto MRSG+erythromycin and then grown at 37 °C overnight. The MRSG+erythromycin plates were then replica plated onto MRS with the sugar being assayed (cellobios, galactose, mannitol, mannose, or xylose) and onto MRSG. After overnight incubation, growth was compared on the selective plate and the glucose plate. Plates were imaged using the macroscope. Colonies that grew on MRSG but not on the selective plate were considered knockouts for the targeted sugar metabolism gene.

Library Construction and Next Generation Sequencing

For sequencing, all the colonies from the initial selection plate (for *E. coli*) or the final MSRG replica plate (for *W. confusa*) were scraped off of the plate and re-suspended using 1 mL of ddH_2O into 1.5 mL micro-centrifuge tubes. The cells were spun down and the supernatant was discarded. Pellet were stored at −20 °C. Genomic DNA was extracted using Promega Genomic DNA Kit. *W. confusa* genomic DNA can be extracted using lysozyme, as per the Promega protocol for Gram-positive bacteria. Once extracted, the DNA was normalized to 10 ng/µL and used as a template for a PCR of the appropriate genomic region (see Supp. File A.3 for amplified regions). PCRs were purified using Zymo Clean and Concentrate and were normalized to 50 ng/µL. Sets of PCRs from different loci (e.g. the first replicates from all of the sugar genes in both *E. coli* and *W. confusa*) were combined. These were sheered on

a M220 Focused-ultrasonicator (Covaris) with a target size of 500 bp. Size distributions were verified on a Bioanalyzer. The final concentration of DNA was generally low (e.g. 20 ng/µL).

A NEBNext Ultra II DNA Library Prep Kit was used for preparation of sheered DNA for sequencing. The initial amplification step was done with 20 cycles due to the low concentration of sheered DNA. Loss of diversity due to the large number of amplification cycles was not a concern due to the library sizes being small. Standard index primers, purchased from NEB, were used.

Samples were run on a MISeq. Before loading, library concentration was measured using qPCR, nanodrop, and qBIT. These gave varying concentrations, so the qPCR concentration was used for loading. Read density indicated the concentration was 1/4 of the concentration indicated by the qPCR (the nanodrop was, in fact, the most accurate).

Reads were aligned using Geneious software. Each replicate was aligned to its reference sequence allowing for discovery of any size deletions. Supplemental File A.4 contains a list of observed deletions supported by at least five reads for each gRNA and replicate.

Confirmation of *W. confusa* knockout phenotypes

For confirmation of knockout genotype and sugar metabolism deficiencies in cellobios and mannose, *W. confusa* was re-transformed with the relevant plasmids. For each transformation, 16 colonies were re-struck onto MSRG+erythromycin plates. One colony from each re-streak was propagated and glycerol stocks were made. At the same time, 100 µL of each culture was spun down, decanted, and stored at −80 °C.

For colony PCRs, frozen cell pellets were first re-suspended in 100 µL of pH 8.0 TE. 1 µL of each re-suspensions was used as template in 25 µL PCR reactions. Standard PCR protocol was used, except the initial 98 °C denaturation step was extended to 5 minutes. Deletions of cellobios and mannose genes were confirmed by "primer-walking" from the PCR primers until the deletions were sequenced.

Sugar metabolism was confirmed by using the minimal media plates described above.

Equipment

OD_{600} was measured using an Ultrospec 10 (Amersham Biosciences) and plastic cuvettes. Biorad thermocyclers were used, as were Eppendorf 5810 and 5810 R centrifuges. HT Multitron and Shell Lab Low Temperature Incubator were used for shaking and stationary incubation, respectively. A custom built macroscope, courtesy of the Kishony Lab and the Harvard Department of Systems biology (see: `https://openwetware.org/wiki/Macroscope`), was used to take pictures of agar plates. For plate reader experiments, a Biotek HT1 Synergy was used.

Chapter 3

DEcreasing the Selective Pressure Of phage Therapy (DESPOT)

Preface

The work detailed in this chapter was a collaboration between Bryan Hsu, Lorena Lyon, and myself. Bryan initially conceived of using temperate phage to reduce virulence without inducing resistance, and performed most of the phage work detailed herein. Lorena designed some of the initial constructs and tested the plasmid based system in *E. coli*. Bryan Hsu and Stephanie Hays gave thoughtful comments on this chapter.

3.1 Abstract

Lytic phage therapy has gained credence as an antimicrobial approach due to the rise of antibiotic resistant microbes. Unfortunately, lytic phages face the same therapeutic problem as antibiotics: they select for resistance to themselves. In this work, we describe a CRISPRi approach for DEcreasing the Selective Pressure Of phage Therapy (DESPOT). DESPOT is comprised of a minimal *Staphalycococus auereus* de-activated Cas9 (sadCas9) and antibiotic resistance cassette. We show that this system functions to repress transcription in both *Escherichia coli* and *Salmonella typhimurium LT2* when expressed from a plasmid. We then demonstrate that DESPOT can be integrated into the lambda phage genome and used to repress specific genes. However, we also show that integration into the B region of lambda phage results in abortive lysogeny in a sadCas9 independent fashion. Ongoing work to integrate DESPOT into the *S. typhimurium LT2* phage P22 is described. DESPOT presents an ability to modulate the activity of virulence genes through engineering of previously unreachable bacteria.

3.2 Introduction

As antibiotic resistance has become an urgent problem[20,26], phage therapy has seen increased interest[29,89]. Phage therapies make use of bacterial predators, bacteriophages (hereafter re-

ferred to as phages), instead of small molecules (antibiotics) to deal with pathogenic bacteria. Historically, dealing with these pathogenic bacteria has meant killing them. As such, using phages to kill bacteria is a natural next step in the current treatment paradigm.

Phages are often divided into lytic and temperate phages[90,91]. Lytic phages (such T4) infect a cell, replicate, and then lyse the cell[92]. As such, for a lytic phage to survive, it has to kill the bacteria[91]. Temperate phages, though, have two options after infection: they can proceed through a lytic life-cycle, which is identical to that of a lytic phage, or they can undergo lysogeny. Lysogeny occurs when the phage integrates into the genome of the target bacteria. The integrated phage is called a prophage (the bacteria that contains it is a lysogen) and the prophage remains stably integrated into the genome until a signal triggers a switch to lytic phase[93,94]. Lysogens are protected from super-infection, which means that the bacteria is immune to re-infection from that specific temperate phage as long as the prophage remains in the genome. The signal to switch life cycles is often DNA damage, which tells the phage that the bacteria is in danger and that the phage should rescue itself by excising itself and proceeding through the lytic cycle.

Between the two types of phages, only lytic phages have been be used as antibacterials[29]. Just like antibiotics, though, lytic phages give rise to resistance (Fig. 3.1A)[30,31,95]. So, while lytic phages have the potential to add to our repertoire of antibiotics, they cannot replace them. Temperate phages, on the other hand, offer a different path forward. Because they can coexist with the target bacteria, they avoid the selection and resistance issue that is inherent in the use of antibiotics and their ilk[37]. But, they will not kill off the pathogenic bacteria. Instead, temperate phages can be used to engineer otherwise un-accessible bacteria.

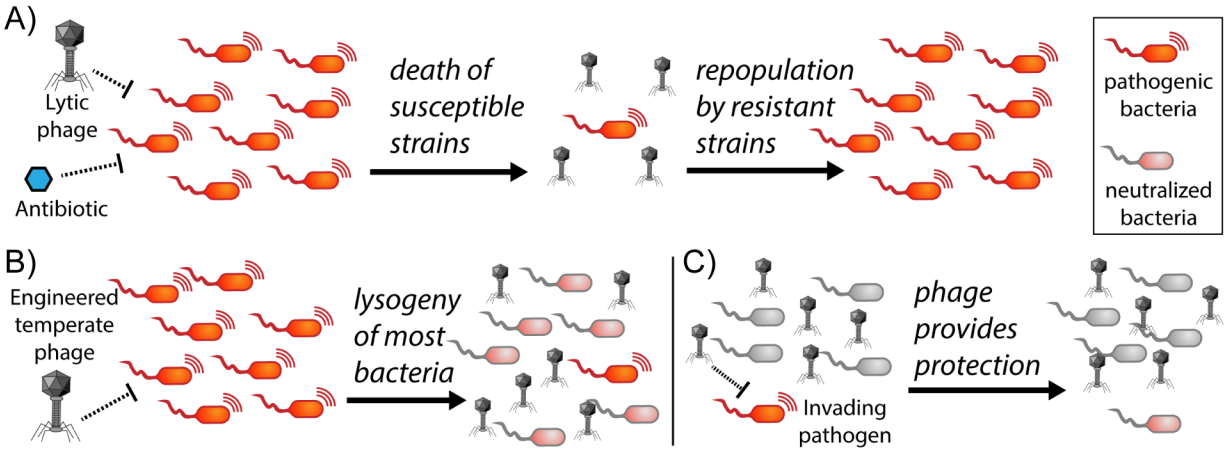


Figure 3.1: **Treating pathogenic bacteria with antibiotics and lytic phage versus engineered temperate phage** Figure courtesy of Bryan Hsu. **A)** When a pathogen is treated used lytic phage or antibiotics, susceptible strains die, but resistant strains survive and repopulate. **B)** Treatment with engineered temperate phage results in lysogeny, and therefore reduced virulence, of most of the bacteria. The engineered bacteria then compete with the non-engineered bacteria, reducing virulence. **C)** The engineered phage can be administered prophylactically and protect the native gut microbiota (gray) from invading pathogens.

Pathogenic bacteria are routinely attenuated before they are used in the laboratory, and it is not hard to take a bacterial sample from an infection and genetically modify it to reduce or eliminate its pathogenicity[96–98]. However, it is impossible to reach the majority of bacteria in an infection and attenuate them. Temperate phages, however, can access and lysogenize these inaccessible bacteria[99,100]. As such, they are able to engineer target bacteria remotely, and have the potential to allow us to attenuate pathogenic bacteria during an infection (Fig. 3.1B) or be prophylactically administered to a healthy microbiota to suppress infectious bacteria (Fig. 3.1C).

In this work, we sought to develop a system that could be delivered by temperate phage to a target bacteria and repress arbitrary virulence genes. Previous work in our lab (unpublished) had shown that treatment of a mouse with a phage combining immunity regions from lambda phages and shiga toxin producing phages could attenuate an *Escherichia coli* O157 infection. This approach, however, is dependent on specific features of *E. coli* O157. To avoid this specificity issue, we sought to show that a dCas9 system could be integrated

into a chosen temperate phage and used to modulate gene expression of a target bacteria.

We designed a system for DEcreasing the Selective Pressure Of phage Therapy (DESPOT). DESPOT was constructed from a dCas9, a gRNA locus, and an antibiotic cassette, and was engineered to contain the minimal number of base pairs needed to remain functional. We tested this minimal systems ability to repress genes in *E. coli* and *Salmonella typhimurium LT2* when expressed from a plasmid, and then whether it could be integrated onto lambda phage (to produce lambda::DESPOT). Next, we examined the stability of lambda::DESPOT lysogens and whether those lysogens could repress transcription in *E. coli*. Finally, we sought to move DESPOT into the *S. typhimurium LT2* phage P22.

3.3 Results

Staphylococcus aureus deactivated Cas9 (sadCas9), tracrRNA, and crRNA were combined onto one plasmid (Fig. 3.2A). Available Cas9 constructs generally keep the Cas9 (or dCas9) and sgRNA on separate plasmids[82]. Further, Cas9 (and related proteins) are very large[101]. To integrate a dCas9 construct into a temperate phage, we needed minimal Cas9, gRNA, and antibiotic resistance constructs. For this, we turned to *S. aureus* deactivated Cas9 (sadCas9)[102,103], which is "only" 3000 bp, and chloramphenicol acetyl transferase (CAT), a chloramphenicol resistance gene that is 750 bp (which is smaller than other commonly used antibiotic resistance such as kanamycin resistance, ∼1000 bp, spectinomycin, ∼1000bp, and ampicillin resistance, ∼1000 bp). These pieces, plus the *S. aureus* tracrRNA and crRNA were combined onto a single plasmid (Fig. 3.2A). The length to be integrated into the phage genomes totaled 4785 bp. In the rest of this work, gRNAs are the name of the gene they target followed by a number (i.e. a gRNA targeting *rfp*, would be rfp1). When a gRNA is added to a construct, it is appended to the constructs name (e.g. pDESPOT with the rfp1 gRNA is pDESPOTrfp1)

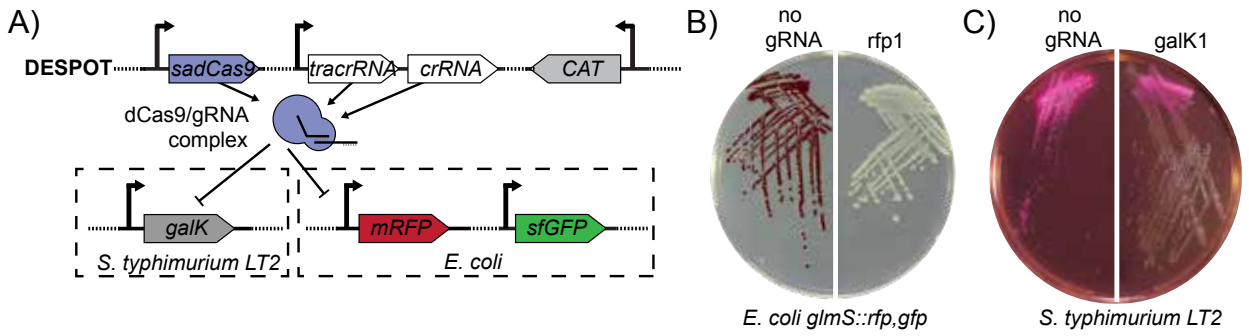


Figure 3.2: **Plasmid delivered DESPOT represses transcription A)** pDESPOT is composed of sadCas9, tracrRNA, crRNA, and chloramphenicol resistance (CAT). When target-specific gRNAs are cloned in, pDESPOT represses the target gene. In this case, the target gene is *rfp* in *E. coli glmS::rfp,gfp* and *galK* in *S. typhimurium LT2*. Panel A courtesy of Bryan Hsu. **B)** Repression of *rfp* in *E. coli glmS::rfp,gfp* by pDESPOT with either no gRNA or the rfp1 gRNA. RFP overlay of brightfield picture is shown. **C)** Repression of *galK* in *S. typhimurium LT2* by pDESPOT with either no gRNA or the galk1 gRNA. Streakouts are on MacConkey-galactose agar. Pink colonies are GalK positive, and white colonies are GalK negative.

The sadCas9 system represses in both *E. coli* and *S. typhimurium LT2*. To confirm that deactivated sadCas9 is capable of repressing transcription in *E. coli* and *S. typhimurium LT2*, we used our plasmid system to express sadCas9 from the promoter P_{proC}, a constitutive promoter previously used to express spCas9 in *E. coli* (see Chapter 2). We call this system pDESPOT. In *E. coli glmS::rfp,gfp*, an *E. coli* strain that constitutively produces both RFP and GFP, pDESPOTrfp1 repressed *rfp*, while pDESPOT alone did not (Fig. 3.2B). Likewise, in *S. typhimurium LT2*, pDESPOTgalk1 repressed *galK* while pDESPOT alone did not (Fig. 3.2C).

DESPOT was integrated into lambda phage. The B region in lambda phage[104,105] and the *gtrA-C* region in P22[106–108] were identified as potential regions for integration of DESPOT, as both of these regions have been reported as non-essential in the literature. pDESPOT was modified for integration by adding two 40 bp homology arms. The initial construct, pDESPOT1, was constructed to integrate across most of the non-essential B region. Lambda phage was introduced to a culture of the permissive host C600 containing pDESPOT1, and allowed to undergo a cycle of infection. The supernatant of the culture, now containing

lambda phage with integrated DESPOT1, was collected, sterile filtered and added to naive C600. After a period of infection, the culture was plated on chloramphenicol to select for resistant colonies. Resistant colonies were restruck to remove residual phage contamination. Restreaks were used to propagate the now modified lambda phage, and the modified lambda phage was used to infect *E. coli glmS::rfp,gfp*.

Lambda::DESPOT1 phage can repress transcription in *E. coli glmS::rfp,gfp*. When grown on selective plates, lambda::DESPOT1rfp1 robustly represses RFP fluorescence, but it does not affect GFP expression (Fig. 3.3Aiv and Biv). In contrast, Lambda-DESPOT1 alone shows no effect on GFP or RFP expression when struck out on selective plates (Fig. 3.3Aiii and Biii).

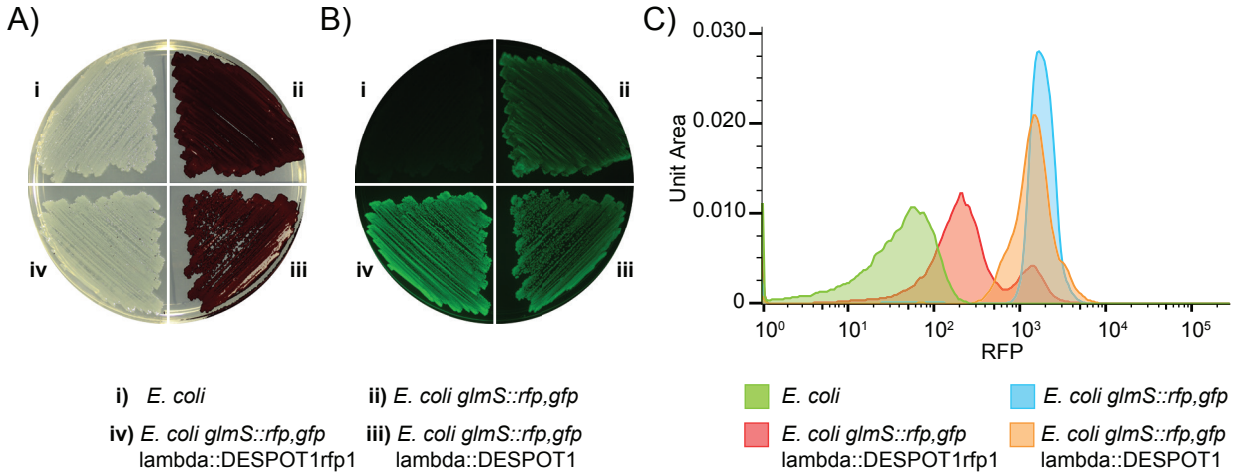


Figure 3.3: **Phage delivered DESPOT1 represses transcription A)** and **B)** Streak outs of *E. coli*, *E. coli glmS::rfp,gfp*, *E. coli glmS::rfp,gfp* lysogenized with lambda:DESPOT1 and *E. coli glmS::rfp,gfp* lysogenized with lambda::DESPOT1rfp1 on LB+chloramphenicol. **A)** RFP channel overlay of brightfield image. **B)** GFP channel image. **C)** Flow cytometry data of strains from **A** and **B** grown in non-selective LB. RFP fluorescence is shown.

Lambda::DESPOT1 is unstable in *E. coli glmS::rfp,gfp*. When *E. coli glmS::rfp,gfp* lambda::DESPOT1rfp1 was grown on non-selective media, we observed that significant portions of the cells regain RFP fluorescence. This was seen when the cells were grown in liquid media and analyzed using flow cytometry (Fig. 3.3C). While the majority of cells show reduced fluorescence relative to lambda:DESPOT1 alone, the lambda::DESPOT1rfp1 popu-

lation is distinctly bi-modal. Restreaks on non-selective agar also regained fluorescence (data not shown). Colonies that had regained RFP expression were not immune to super-infection by virulent lambda phage (i.e. they did not have a functioning lambda prophage anymore) and they where re-sensitized to chloramphenicol. To confirm that the presence of sadCas9 was not causing instability in the phage, we grew a lambda lysogen with sadCas9 produced from a plasmid. However, no loss of phage was observed. When lambda::DESPOT1rfp1 was repeatedly re-struck on selective plates, RFP repression was maintained and cells grew robustly.

DESPOT variants integrated in different parts of the B region also cause instability. After a review of the literature, we determined that the B^{105} region was not sufficiently characterized to guarantee integration would not affect the phage's ability to maintain lysogeny. As such, we constructed two other integrating constructs, DESPOT2 and DESPOT3. These constructs both integrated in the B region, but in slightly different areas and with different effects on the size of the phage genome (Fig. 3.4A). Both of these constructs, when maintained on selective media (Fig. 3.4B), showed repression of RFP activity and robust growth. When the cells were moved to non-selective media, colonies lost repression (Fig. 3.4B). However, different isolates lost repression at different rates, with some isolates requiring multiple re-streaks before they lost repression (such as isolate two in Fig. 3.4B).

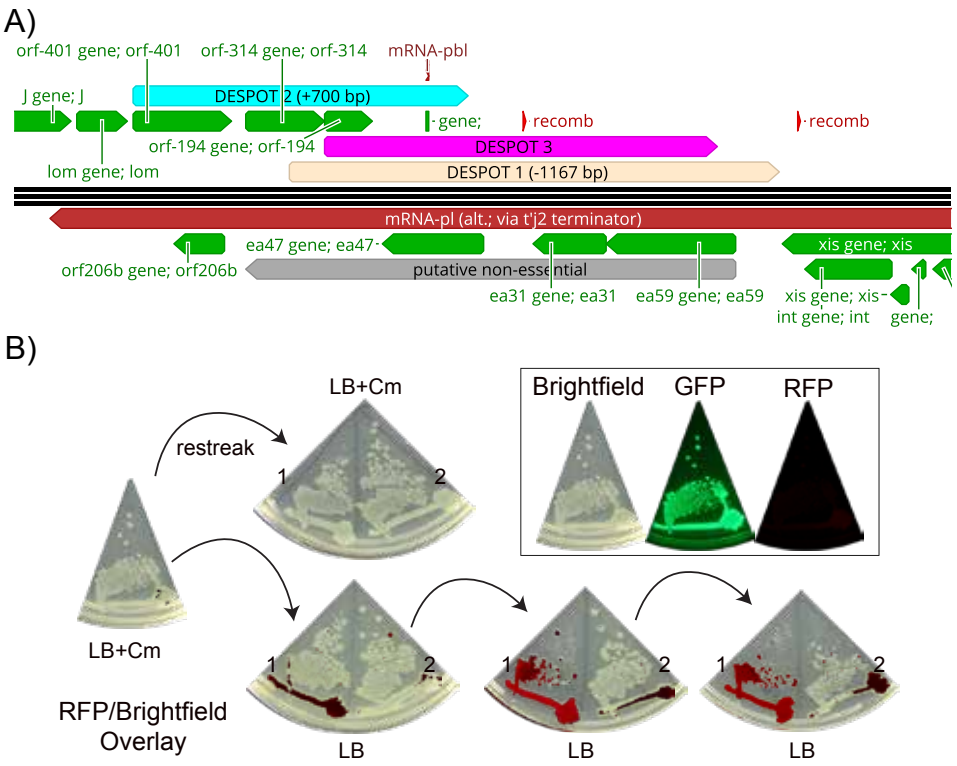


Figure 3.4: **Lambda::DESPOT shows abortive lysogeny under non-selective conditions A)** Genetic schematic of integration of DESPOT constructs into the lambda genome. Change in genome size for DESPOT1 and 2 is noted. **B)** Representative restreak of DESPOT3 colonies from initial streak on LB+chloramphenicol. RFP brightfield overlay is shown for the restreaks. A pair of isolates (1 and 2) are shown.

P22 has proved recalcitrant to our attempts to integrate DESPOT. To engineer P22, a P22 lysogen was established in *S. typhimurium LT2 hsdR::λRED*. In principle, induction of the Lambda Red system should allow for integration of a transformed piece of linear DNA that has sufficient homology arms. Initial attempts to integrate DESPOT, using 40 bp homology arms, failed. Further attempts with 100 bp homology arms also failed. To avoid dealing with PCR products, we constructed a temperature sensitive plasmid that contained 1000 bp of homology (pDESPOTp22). This plasmid was transformed into the target strain, which was then grown under Lambda Red inducing conditions. The plasmid was cured, while integrants were selected for. This strategy, however, also failed to yield integrants. Work on this issue is ongoing.

3.4 Discussion

Antibiotics and lytic phages select for resistance, and therefore select for their future failure. In this work we have sought to develop a strategy for dealing with bacterial infections for which resistance is not correlated with therapeutic success. We approached this problem by using temperate phages to deliver a genetic payload. The genetic payload is DESPOT, which is a sadCas9 system that can be used to repress transcription in targeted cells.

As expected pDESPOT is able to repress transcription in both *E. coli* and *S. typhimurium LT2*. saCas9 had been previously studied in mammalian cell and bacterial systems[102], but sadCas9 had only been applied in mammalian systems[103]. We demonstrate that sadCas9 can be used for CRISPRi in both *E. coli* and *S. typhimurium LT2*. In contrast to previous work that used sgRNAs, here we use a tracrRNA and crRNA system that can easily produce more than one gRNA and therefore repress multiple genes.

DESPOT can be integrated into lambda phage. We showed that replacing the B region of the lambda phage with DESPOT is not lethal, and can be done in a number of different ways. However, when DESPOT is integrated into the B region, lambda phage creates unstable lysogens. When we expressed pDESPOT in the presence of a lambda lysogen, the lysogens remained stable, leading us to conclude that integrants into the B region lead to abortive lysogeny. While not necessarily an issue for proof of concept experiments, seeing abortive lysogeny in arguably the most well studied bacterial phage suggests that better methods are needed to screen for non-essential regions.

Lambda with integrated DESPOT effectively represses targeted transcription. Despite lysogen instability, lambda with integrated DESPOT repressed targeted genes. Repression appeared to be equally effective between all three versions of DESPOT. Further work needs to be done to examine whether DESPOT can be used to repress virulence genes to a level that reduces pathogenicity, and whether more than one gRNA effects repression efficacy.

Attempts to integrate DESPOT into the P22 phage have so far been unsuccessful. While we expect DESPOT to function to repress transcription when integrated into the *S. ty-*

phimurium LT2 genome, integrating DESPOT into P22 is a key next step. Work on solving this problem is ongoing.

3.5 Materials and Methods

Strains and Growth Conditions

Complete strain information can be found in Supplemental Table B.1. *E. coli* and *Salmonella LT2* were routinely grown in Lysogeny Broth (LB) Miller and on LB agar plates. DH10β was used as default cloning strain, C600 was used for propagation and engineering of lambda phage, and *E. coli glmS::rfp,gfp* was used for experiments. Antibiotics were supplemented as follows: 25 µg/mL (for plasmids) or 10 µg/mL (for engineered phages) chloramphenicol.

Plasmid Construction

A complete list of plasmids can be found in Supplmental Table B.2 and plasmid maps can be found in Supplemental File B.1. Golden Gate Assembly was used for cloning plasmids and inserting gRNAs. Q5 Hot Start polymerase was used to amplify DNA for all cloning besides insertion of gRNAs. DNA for the coding sequences of sadCas9 was ordered as gBlocks from IDT. For Golden Gate reactions, 10xT4 Ligase Buffer (Promega), T4 Ligase (2,000,000 units/mL, NEB), and BSA (10 mg/mL, NEB) were used in all reactions. The appropriate restriction enzyme, either Eco31I, Esp3I, or SapI (Thermo FastDigest), was added. gRNAs were added to plasmids by Golden Gate after annealing and phosphorylating pairs of oligos. For detailed information, see Supplemental Materials and Methods. Briefly, complementary oligos were incubated with T4 Ligase Buffer (NEB) and T4 Poly Nucleotide Kinase (NEB), heated to boiling, and then slowly cooled to room temperature. Annealed oligos were then added to plasmids using Eco31I.

Golden Gate reactions were desalinated using drop dialysis (for a minimum of 10 minutes) and electroporated in DH10β Electrocompetent Cells (Thermo Fischer).

E. coli and *Salmonella LT2* Electrocompetent Preparation and Transformations

E. coli and *S. typhimurium LT2* were made electrocompetent using a modified standard protocol[87]. *E. coli* was grown in either LB or 2xYT until OD_{600} of between 0.4 and 0.6. Cells were spun down at 4000xG for 10 minutes at room temperature. Cells were then washed twice with 0.5x volume room temperature ddH_2O and then once with 0.1x volume room temperature ddH_2O. Cells were re-suspended in approximately 0.001x volume of room temperature ddH_2O (if competent cells were to be used immediately) or 10% room temperature glycerol (if cells were to be frozen and stored at −80 °C). 25 µl of cells were transformed with 2.5 µl of 20 ng/µl of mini-prepped plasmid. Electroporations were performed using 0.1 cm cuvettes at 1.8 kV (Ec1) in a BioRad MicroPulser™. Cells were rescued in 972.5 µl SOC for 1 hour at 37 °C. Cells were then plated on agar plates with appropriate selection.

Occasionally, cells prepared using this method are too concentrated and arc, so a no-DNA control electroporation was always performed. If the pulse time was less than 5 ms, cells were diluted until an appropriate pulse time was achieved.

Integration of DESPOT into λ Phage

2 mL of C600 containing a pDESPOT plasmid was grown overnight. The next day, WT Lambda phage was introduced to a culture and was incubated for one hour, to allow for lytic reproduction and recombination of pDESPOT into the phage. The supernatant of the culture, now containing lambda phage with integrated DESPOT1, was collected and passed through a 2 µm filter. The phage lysate was then added to a stationary culture of naive C600 (grown overnight as above) and allowed to incubate for three hours, so as to allow for infection, lysogeny, and production of antibiotic resistance. The cultures were then struck out onto selective plates, grown overnight, and colonies were restruck to remove residual phage contamination. To move the modified phages (lambda::DESPOT) into other *E. coli*

strains, the C600 lysogens were grown overnight and the supernatant was sterile filtered to produce a phage lysate that was used to infect EC001 or other strains as desired.

S. typhimurium LT2 P22 integration attempts

A P22 lysogen of PAS710 (*S. typhimurium LT2 hsdR::λRED*) was created using standard procedures. For attempts at integrating PCR products into P22, the strain was made electrocompent and transformed as described above, except that it was grown during the day of with 10 mM l-arabinose (to induce λRED) and the rescue time was increased up to 3 hours. For plasmid based attempts, the strain was made electrocompetent and transformed as normal, except rescue and growth on selective plates (ampicillin) were done at 30 °C. Colonies were picked into LB with 10 mM L-arabinose and chloramphenicol and grown at 30 °C for 8 to 10 hours. This culture was diluted 100 fold into LB with chloramphenicol and grown at 42 °C overnight. The dilution and growth was repeated the next day, and in the evening the culture was plated on LB+Cm plates and grown at 37 °C. Colonies were checked for ampicillin resistance (which indicates the presence of the plasmid). Unfortunately, colonies consistently showed ampicillin resistance.

Equipment

OD_{600} was measured using an Ultrospec 10 (Amersham Biosciences) and plastic cuvettes. Biorad thermocyclers were used, as were Eppendorf 5810 and 5810 R centrifuges. HT Multitron and Shell Lab Low Temperature Incubator were used for shaking and stationary incubation, respectively. A custom built macroscope, courtesy of the Kishony Lab and the Harvard Department of Systems biology (see: `https://openwetware.org/wiki/Macroscope`), was used to take pictures of agar plates.

Chapter 4

Host Organism-Agnostic Kinase Sensing (HOAKS)

Preface

The work detailed in this chapter was a collaboration between Christine Zheng, Elizabeth Libby, and myself. Elizabeth Libby designed a prototype of the FRET sensor, conceived of the possibility of breaking a repressor with kinase activity, and performed most of the microscopy described herein. Christine Zheng cloned a number of the LORK constructs. Bryan Hsu and Stephanie Hays gave thoughtful comments on this chapter.

4.1 Abstract

Serine/Threonine Kinases (STKs) are widely distributed among extant organisms, and are involved in a number of fundamental cell signaling pathways. In humans, where they were first discovered, STKs have been extensively studied, and there has been some limited tool development for studying their activity *in vivo*. With the recent focus on STKs in bacteria, the lack of tools for studying them on a single cell level has become more apparent. With that in mind, we developed two complementary tools to study STKs in *Bacillus subtilis*. First, we take a Förster Resonance Energy Transfer (FRET) sensor that was developed for mammalian cells and re-design it for use in *B. subtilis*. The FRET sensor is composed of a FRET pair (CFP and YFP) linked by a Fork Head Associated (FHA2) domain and a substrate. Upon phosphorylation of the substrate, the FRET sensor undergoes a conformational change that modifies its fluorescent properties. The FRET sensor senses, and responds through, a post-translational mechanism. In contrast, to integrate kinase activity into transcriptional pathways, we developed a set of Lactose OR Kinase (LORK) sensors by inserting the FHA2 and substrate domains into LacI. The LORK sensors respond to kinase activity by de-repressing a target gene. We show that both of these tools can be used to sense the *B. subtilis* kinase PrkC, and that they allow us to determine the phosphorylation state of our target substrate. Finally, we use the LORK sensor to show that staurosporine does not specifically inhibit PrkC in *B. subtilis*.

4.2 Introduction

Ser/Thr Kinases (STKs) were historically studied in the context of mammalian cell signal transduction[109], while histidine kinase systems predominated bacterial work[110]. Over the last two decades, however, there has been a growing recognition that STKs are widely distributed throughout the tree of life and play a crucial role in the cellular physiology of certain bacteria[40,43,51]. Much of this recent work has been performed in firmicutes and focused on the ways that STKs regulate cell wall synthesis. However, the available tools for measuring STK activity has limited what these studies can discover[50,52].

STKs phosphorylate serines and threonines in a sequence specific manner[111]. STKs often have cognate phosphatases (STPs) that dephosphorylate the serines and threonines that the STK has phosphorylated, although the STK and STP may have different sequence preferences[112,113]. This results in serine and threonine phosphorylations with different half-lives depending on their local protein context. In the absence of an STP, a serine or threonine phosphorylation is stable for the life of the protein[114].

The most well studied STKs in bacteria are comprised of a cytoplasmic Hanks-type kinase domain, a transmembrane domain, and an extracellular Penicillin Binding and STK Associated (PASTA) domain[115]. The kinase domain of these STKs is highly homologous to mammalian Protein Kinase A (PKA) and related STKs. Hanks-type kinases are often inactive as monomers and become active upon dimerization[116,117]. In bacteria, STKs are thought to sense cell wall intermediates through their PASTA domains, dimerize, and then begin to phosphorylate protein targets[39,47,115]. To examine whether this is true, however, it is necessary to assay kinase activity. While certain tools exist for measuring STK activity, they have significant shortcomings.

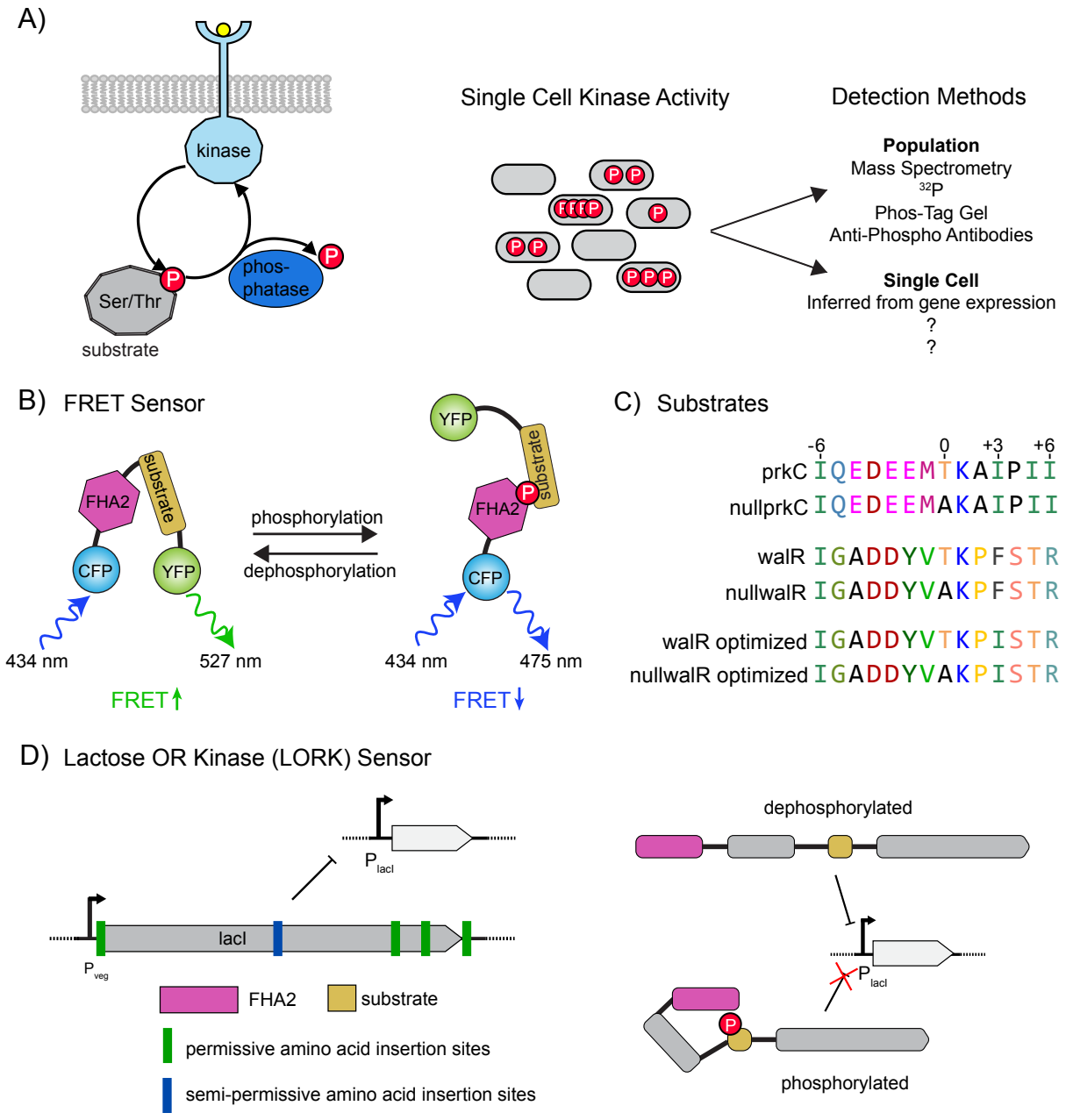


Figure 4.1: **Kinase detection and sensor schematics** Panels A and B of this figure are courtesy of Elizabeth Libby. **A)** Diagram of bacterial STK system with substrate and cognate STP. Single cell kinase activity may be variable, but most current detection methods can only detect bulk population kinase activity. Single cell measurements based on inferred gene expression are driven by kinase activity, but are also effected by general cell physiology. **B)** Diagram of a phosphorylation-state dependent FRET sensor. **C)** The six 15 amino acid substrates tested in this paper. **D)** Diagram of Lactose OR Kinase sensor. Permissive and semi-permissive amino acid insertion sites are labeled.

When nothing is known about an STK of interest, bulk culture phospho-proteomics is

often the first step for determining possible targets (Fig. 4.1A)[49,50,52]. Ideally, the phosphoproteomes of a Wild Type (WT) strain, a kinase knockout strain, and a phosphatase knockout strain are assayed and compared[51]. This approach often has low "on-rate" (i.e. it does not reveal all phosphorylated proteins), but it allows *de novo* identification of the sequences of phosphorylated protein targets. Once these targets have been identified, their phosphorylation can be assayed using ^{32}P, antibodies that recognize only the phosphorylated version of a protein, or Phos-Tag gels[54,118]. However, all of these assays measure kinase activity in bulk culture, and cannot help elucidate cell-by-cell STK variation.

Currently, kinase dependent transcriptional signals must be identified to study single cell dynamics (Fig. 4.1A). Ideally, the complete transcriptomes of a WT strain, a kinase knockout strain, and a phosphatase knockout strain are compared, in an analogous process to that used for identifying phosphorylation sites[119,120]. With an identified transcriptional signal, kinase activity can then be inferred for various conditions. However, it can be difficult to assign single cell behavior to kinase activity even when robust differences are observed because both kinase activity and a host of other cellular machinery can affect transcription, thereby muddying the causal story[121–123]. With this in mind, we sought to develop a suite of tools that reported on kinase activity independent of the host cell physiology.

A cell physiology independent sensor of the STK AuroraB was previously developed for mammalian cells[57,124]. This type of sensor, which we will refer to as a Förster Resonance Energy Transfer (FRET) sensor, is a protein composed of a FRET pair (usually CFP and YFP) linked by a Forkhead Associated Domain (FHA)[125] and a substrate peptide that can be bound by the FHA (Fig. 4.1B). The substrate peptide is designed such that it is a target both of the STK of interest and will be bound by the FHA[126]. The FRET sensor occupies its "normal" conformation when the substrate is not phosphorylated and the FHA is not bound. This conformation puts the CFP and YFP in close proximity, resulting in a high FRET signal (FRET signal is reported in this work as FRET/CFP signal). When the STK phosphorylates the substrate, the FHA binds to the substrate and "breaks" the sensor.

In the broken conformation, the CFP and YFP are far enough away from each other that they cannot undergo FRET. As such, kinase activity reduces FRET signal. This sensor has the advantage of responding purely in a post-translational manner to kinase signal and has a theoretically rapid response time. Despite the advantages of a purely post-translational sensor, FRET signals tend to be quite small and only observable with microscopy[127]. As such, while developing the FRET sensor for the *B. subitilis* STK PrkC, we sought to develop a complementary sensor that responded through a transcriptional mechanism. However, unlike inferred transcriptional activity, we designed a sensor that would be largely independent of the host-cell's physiology and would be portable to other organisms.

Building off of previous FRET sensor work, we took advantage of FHA's ability to modify the conformational state of proteins. We reasoned that a FHA and substrate domain would potentially be able to "break" a repressor upon phosphorylation of the substrate. Upon breaking, the repressor would stop repressing and allow for production of a signal (Fig. 4.1D). This strategy requires being able to place the FHA and substrate in a repressor and have the repressor still function in the absence of kinase activity. As such, we chose to modify LacI, a well-studied repressor. LacI has a number of "permissive" sites where exogenous protein sequences can be introduced without affecting LacI's repressive ability[128]. Because any gene can be controlled by the LacI regulated promoter, LacI is able to produce signals that can be read in a plate reader (luminescence), in a microscope (fluorescence) and potentially even genetically (such as producing a recombinase, although this possibility is not discussed here). Such a sensor can be thought of as a specific example of an Allosteric OR Post-Translational switch (AORPT), and we refer to this broken LacI as a Lactose OR Kinase (LORK) sensor.

In this paper, for the first time, we adapt a kinase responsive FRET sensor for use in bacteria. As a proof of concept, we demonstrate that it can respond to activity of the *B. subtilis* STK PrkC. We test a number of substrates (Fig. 4.1C), and show that for a substrate that responds to kinase activity, the FRET sensor produces consistent absolute signal across

different experiments. We take the substrate that functions in the FRET sensor and use that to develop a small library of LORKs. We show that a number of our LORKs respond to kinase activity, and identify an optimal variant (LORK4). We show that LORK4 responds to kinase activity on a single cell level.

We use both LORK4 and the FRET sensor to comment on some outstanding physiology questions surrounding *B. subtilis* kinase PrkC. As negative controls for kinase activity, we use either the $\Delta prkC$ or the $\Delta(prpC\text{-}prkC)$ genetic backgrounds. PrpC is the cognate phosphatase of PrkC, so as a positive control for kinase activity, we use the $\Delta prpC$ genetic background (which should have very stable PrkC phosphorylations). To test our sensors, we confirm previous bulk culture assays[51] that suggest the T290 autophosphorylation site is not phosphorylated under normal growth conditions. Second, we look at the mammalian STK kinase inhibitor staurosporine, which is often used to inhibit bacterial kinases despite the limited evidence for its efficacy. We use LORK4 to show that staurosporine effects bacterial physiology but that some, if not all, of the effect is kinase independent.

4.3 Results

4.3.1 A Phosphorylation Responsive FRET Sensor

The FRET sensor was designed and optimized for high expression in *B. subtilis* by using a constitutive promoter, a canonical RBS, and codon optimization. The sensor is protein composed of CFP-FHA2-substrate-YFP (FHA2 is a well characterized FHA protein). A similar protein construct had originally been used in mammalian cells to sense AuroraB activity[57], and efforts had been made to improve it for mammalian systems[129]. These efforts generally involved testing different fluoropohores and linker lengths[130]. The most successful of these variants used CPET (a CFP variant) and YPET (a YFP variant) fluorophore pair. However, evidence suggested that when the CPET and YPET were close enough to undergo FRET, they locked in that conformation[127]. Since we want a dynamic sensor, the CPET-

YPET pair was ruled out, and the CFP-YFP pair was used instead.

In *B. subtilis* we wanted to achieve high levels of protein expression. While high expression of the sensor might not be preferable from a sensing perspective, it is generally easier to reduce biological expression in a high expression system, than to raise expression in a low expression system. To produce high levels of the FRET sensor we turned to the P_{veg} promoter, which has high levels of expression under most conditions in *B. subtilis*, followed by the canonical RBS (AGGAGG) optimally spaced from the start codon. Finally, the entire FRET sensor was codon optimized for *B. subtilis*.

Multiple substrate variants of the FRET sensor were constructed. A FRET sensor substrate must meet two condition for it to sense PrkC: it must be phosphorylated by PrkC and the FHA2 domain must bind to the phosphorylated substrate. However, PrkC and FHA2 domains have different sequence specificities, only the latter of which is known. As such, to increase the chance of having a functioning sensor, we tried three different 15 amino acid substrates, two of which had been characterized as substrates of PrkC and one modified substrate to increase the likelihood of FHA2 binding (Fig. 4.1C). The first is a substrate on the PrkC protein itself and is referred to as the prkC substrate. It is derived from the known T290 autophosphorylation site on PrkC, and has a TXXI amino acid motif, which is favored for FHA2 binding[51,131,132]. The second substrate (walR) is from the known PrkC protein substrate WalR. The walR substrate is known to be phosphorylated by PrkC, but has a TXXF amino acid motif instead of TXXI, and therefore is unlikely to be bound by the FHA2 domain[42]. The walRopt substrate is the walR site with an isoleucine substitution at the +3 position. As such, walRopt is likely to be bound by FHA2 if phosphorylated, but it is unknown if PrkC will phosphorylate walRopt in the first place. Finally, as controls, we made nullprkC, nullwalR, and nullwalRopt targets, which have T→A mutations, and therefore are not phosphorylatable by PrkC (Fig. 4.1C).

Activity of the FRET sensor with different substrates was tested in WT, a phosphatase knockout strain, Δ*prpC*, and a dual kinase and phosphatase knockout strain, Δ*(prpC-prkC)*.

Sensors were constructed on plasmids that were propagated in *E. coli*, and then integrated at the *sacA* locus in *B. subtilis* 168 with selection on chloramphenicol. The sensors were then crossed into Δ*prpC*, Δ*(prpC-prkC)* and backcrossed into *B. subtilis 168*. All experiments were performed by streaking out on non-selective agar plates, growing overnight, and then picking colonies into MOPS Minimal media.

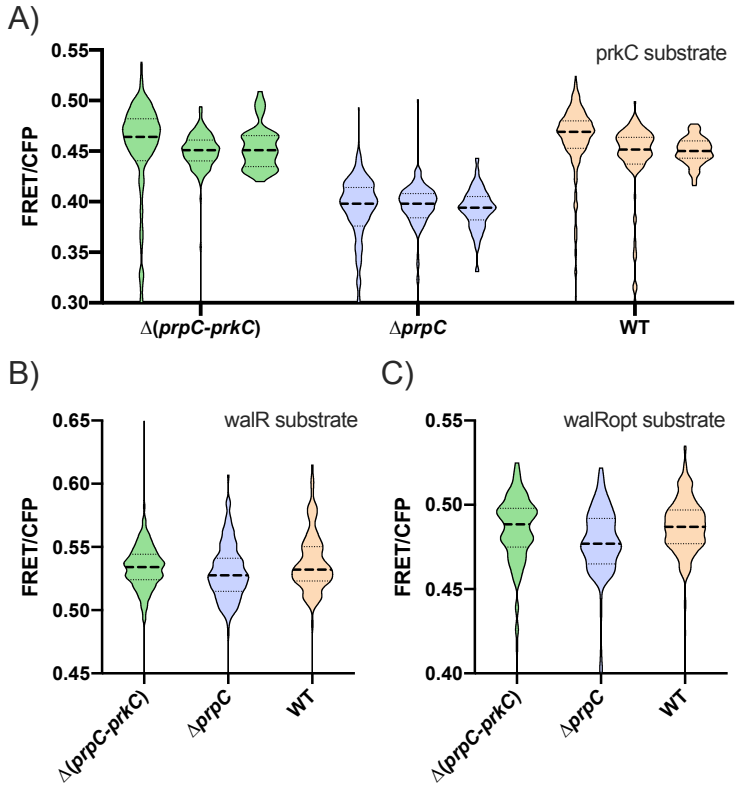


Figure 4.2: **FRET sensor responds to kinase activity in a substrate dependent manner** Each violin is a single microscopy experiment replicate of approximately 90 cells and FRET signal is reported normalized to CFP signal (FRET/CFP). For a functional sensor, kinase activity is expected to decrease FRET/CFP signal. **A)** FRET sensor with the prkC substrate in the WT (IP_352), Δ*prpC* (IP_356), and Δ*(prpC-prkC)* (IP_360) genetic backgrounds. Dashed black lines are the median, and dotted black lines are the interquartile range. Replicates in the three genetic backgrounds were from three individual days. **B)** FRET sensor with the walR substrate in the WT (IP_600), Δ*prpC* (IP_601), and Δ*(prpC-prkC)* (IP_602) genetic backgrounds. **C)** FRET sensor with the walRopt substrate in the WT (IP_354), Δ*prpC* (IP_358), and Δ(*prpC-prkC)* (IP_362) genetic backgrounds.

The FRET sensor with the prkC substrate shows PrkC dependent activity, while the sensor does not show PrkC dependent activity when it has either the walR or the walRopt

substrates. The prkC FRET sensor shows robust signal differences (complete separation of the interquartile ranges) between the Δ*prpC* and Δ*(prpC-prkC)* genotypes, although there is no observable difference in signal between WT and Δ*(prpC-prkC)* (Fig. 4.2A). The signal is robustly replicable across different days. In comparison, neither the walR (Fig. 4.2B) nor walRopt (Fig. 4.2C) sensors show any kinase response. Interestingly, the walR and walRopt sensors have different absolute FRET signals (medians of ∼0.54 and ∼0.48, respectively) which suggests that minor changes to the substrate peptide can have significant conformational implications for the sensor.

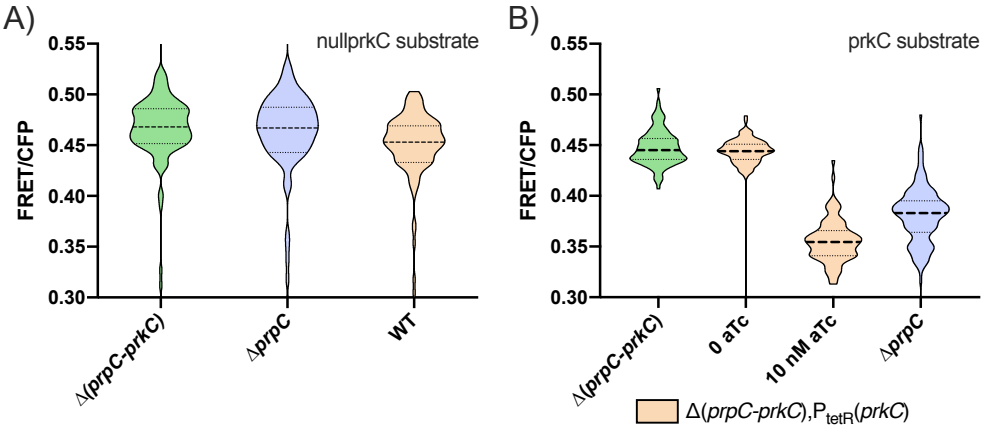


Figure 4.3: **FRET sensor requires phosphorylatable substrate and responds to PrkC complementation** Each violin is a single microscope experiment replicate of approximately 90 cells and FRET signal is reported normalized to CFP signal (FRET/CFP). For a functional sensor, higher kinase activity is expected to decrease FRET/CFP signal. Data is from representative experiments. **A)** FRET sensor with nullprkC substrate (T→A) in the WT (IP_353), Δ*prpC* (IP_357), and Δ(*prpC-prkC*) (IP_361) genetic backgrounds. **B)** FRET sensor with prkC substrate in Δ(*prpC-prkC*)P_{tetR}(*prkC*) (IP_630) background. PrkC is induced at two levels: no induction (0 nM) and full induction (10 nM). FRET sensor in the Δ*prpC* (IP_356), and Δ(*prpC-prkC*) (IP_360) backgrounds also shown.

The FRET sensor requires a phosphorylatable residue to respond to kinase activity. To confirm that the FRET sensor was responding to phosphorylation of its prkC substrate, we tested the FRET sensor with the nullprkC substrate. The nullprkC substrate has a T→A mutation at the phosphorylation site, and should therefore be unresponsive to kinase activity. As expected, the nullprkC FRET sensor does not respond to kinase activity (Fig.

4.3A). This suggests that the active prkC FRET sensor is responding to phosphorylation of its substrate.

The FRET sensor responds to complemented PrkC. To confirm that PrkC, and not some effect of knocking out *prkC*, was responsible for our observed signal, we produced PrkC from the *amyE* locus instead of from the native *prkC* locus. For the FRET sensor, absence of induction results in a FRET signal equal to that of Δ*(prpC-prkC)* (Fig. 4.3B). When the kinase is induced at 10 nM aTc, the FRET signal is measurably stronger than that of Δ*prpC*. This suggests that normal levels of kinase are well within the dynamic range of the FRET sensor.

4.3.2 A Lactose OR Kinase Sensor

A small library of LORK variants was constructed. To construct the library, permissive sites for coding sequence insertion were determined based on a past insertional mutagenesis study. Residues 152, 317, 338 are labeled as semi-permissive (152) and permissive (317 and 338) based on whether a small coding sequence can be inserted without affecting LacI's ability to repress transcription (Fig. 4.1D)[128]. The N- and C-termini were included as permissive sites as many studies have made LacI fusion proteins[133,134]. The library was designed such that the FHA2 and the substrate would be at every possible pair of positions, as well as together at each positions, with one exception. Due to concerns about large insertions at the semi-permissive site, the FHA2 and the FHA2-substrate fusion were not placed there. These were cloned using a combinatorial golden gate approach (see Supp. File C.1 for more details). Initially, all the constructs were cloned under P_{veg} and the canonical RBS. Only a small number of constructs proved possible to clone (LORKs 1-5, 12, and 17), so we moved forward with testing those constructs. It was eventually determined that over-expression of LORK (or LacI) was causing toxicity. A Δ*pcnB E. coli* strain (which reduces the plasmid copy number, and therefore toxicity) was used to clone the two constructs of particular interest (LacI and LORK13, which is LORK4 reversed), while the rest of the un-cloned

constructs were abandoned.

LORK variants were tested for kinase response. To test whether the LORK variants responded to kinase activity, they were integrated into *B. subtilis* 168, and then crossed into Δ*prpC* and Δ*(prpC-prkC)* strains that had $P_{hyperspank}$(mCherry) integrated at the *amyE* locus. $P_{hyperspank}$ is a LacI regulated promoter, making mCherry signal responsive to the LORK variants.

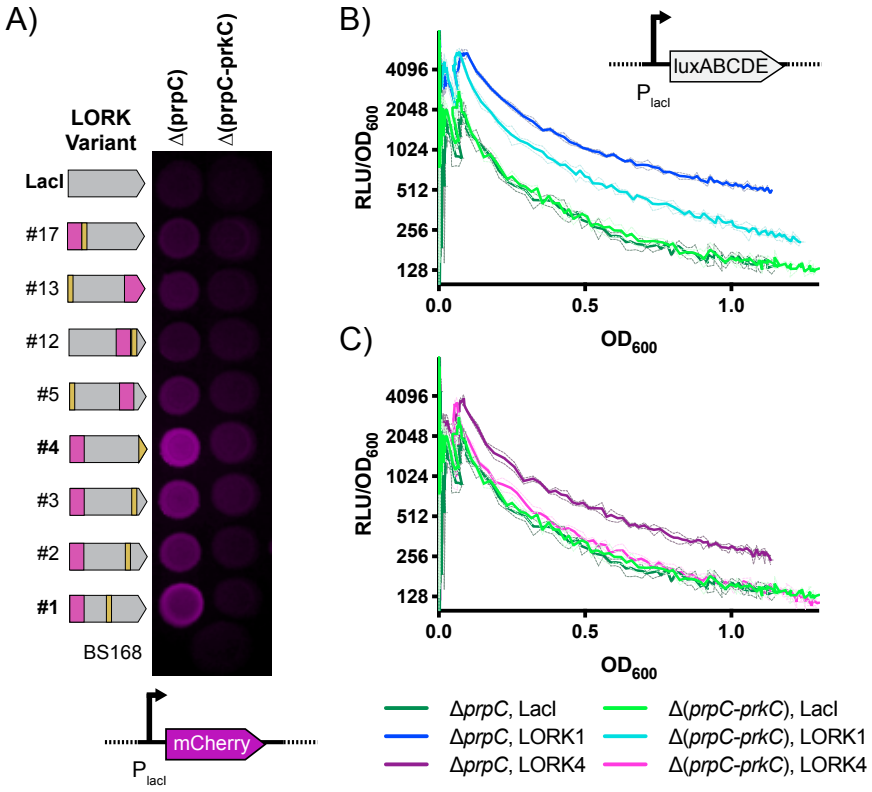


Figure 4.4: **Kinase response of LORK variants A)** The LORK variants and LacI regulating mCherry in the Δ*prpC* and Δ(*prpC-prkC*) genetic backgrounds (IP_453-494) were spotted on an LB agar plate, along with *B. subtilis* 168 (non-fluorescent control). Increased kinase activity should increase mCherry signal in a functioning sensor strain. **B)** LORK1 and **C)** LORK4, paired with LacI, were tested for their response to kinase activity in the Δ*prpC* and Δ(*prpC-prkC*) genetic backgrounds. LORK1 (IP_545 and 546), LORK4 (IP_547 and 548), and LacI (IP_549 and 550) regulated *luxABCDE*, which is reported as Relative Luminescence Units (RLU) normalized by OD_{600} (RLU/OD_{600}). RLU/OD_{600} is plotted on a $\log_2$ scale. Cultures were grown in triplicate in a plate reader. The means are displayed as bold lines and standard deviation as dashed lines. Representative experiments are displayed.

LORK4 can sense PrkC activity. A number of the LORK variants respond to PrkC

activity. While LacI and *B. subtilis* 168 are not fluorescent, a number of the LORK variants show a response to the presence of PrkC when spotted on agar (Fig. 4.4A). LORK1 and LORK4 showed the most robust response to kinase activity. These strains were also grown overnight in a plate reader and showed similar responses (Supp. Fig. C.3). LORK1, LORK4, and LacI, were individually combined with the $P_{hyperspank}$ promoter and various reporters as single constructs and studied further. Both LORK1 (Fig. 4.4B) and LORK4 (Fig. 4.4C) respond to kinase activity across the range of measured ODs, while LacI shows no difference between Δ*prpC* and Δ*(prpC-prkC)*. LORK1 show poor repression in the OFF state relative to LacI, with the Δ*(prpC-prkC)* running significantly above both LacI strains. LORK4, on the other hand, represses as well as LacI. This behavior is consistent across all observed ODs.

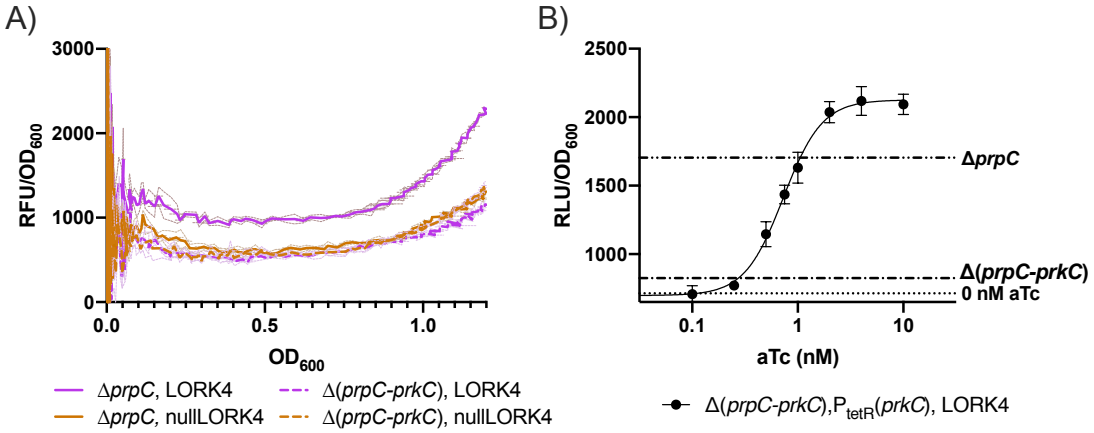


Figure 4.5: **LORK4 requires phosphorylatable substrate and responds to PrkC complementation** Data is from representative experiments. **A)** LORK4 (IP_547 and 548) and nullLORK4 (IP_640 and 641) in Δ*prpC* and Δ(*prpC-prkC*) backgrounds. LORK regulates YFP and is reported as Relative Fluorescence Units (RFU) normalized by OD_{600} (RFU/OD_{600}). The means are displayed as bold lines and standard deviation as faint dashed lines. **B)** LORK4 in Δ(*prpC-prkC*)P_{tetR}(*prkC*) background (IP_632). LORK4 regulates *luxABCDE* and is reported as Relative Luminescence Units (RLU) normalized by OD_{600} (RLU/OD_{600}). Dots are the mean of three technical replicates, error bars are standard deviation, and a four parameter logistics fit is displayed. The dotted line is the mean of the uninduced (0 nM aTc) condition, and the two dashed lines are the means of LORK4 in the Δ*prpC* (IP_547) and Δ(*prpC-prkC*) (IP_548) backgrounds grown at 10 nM aTc. The RLU/OD_{600} values reported are from OD_{600}=0.8.

LORK4 signal can be titrated by PrkC produced *in trans*. To confirm that LORK4

was responding to PrkC and not a polar effect of deletion, we induced PrkC *in trans*. We ran a number induction levels to test the dose response of LORK4 to PrkC. LORK4 shows a titratable response to PrkC with an EC50 of 0.75 nM and maximum response at 2 nM Maximum induction showed higher signal than that of Δ*prpC*. Also, absence of induction showed a slightly reduced signal relative to a Δ*(prpC-prkC)* strain. This suggests that normal levels of kinase are well within the dynamic range of LORK4.

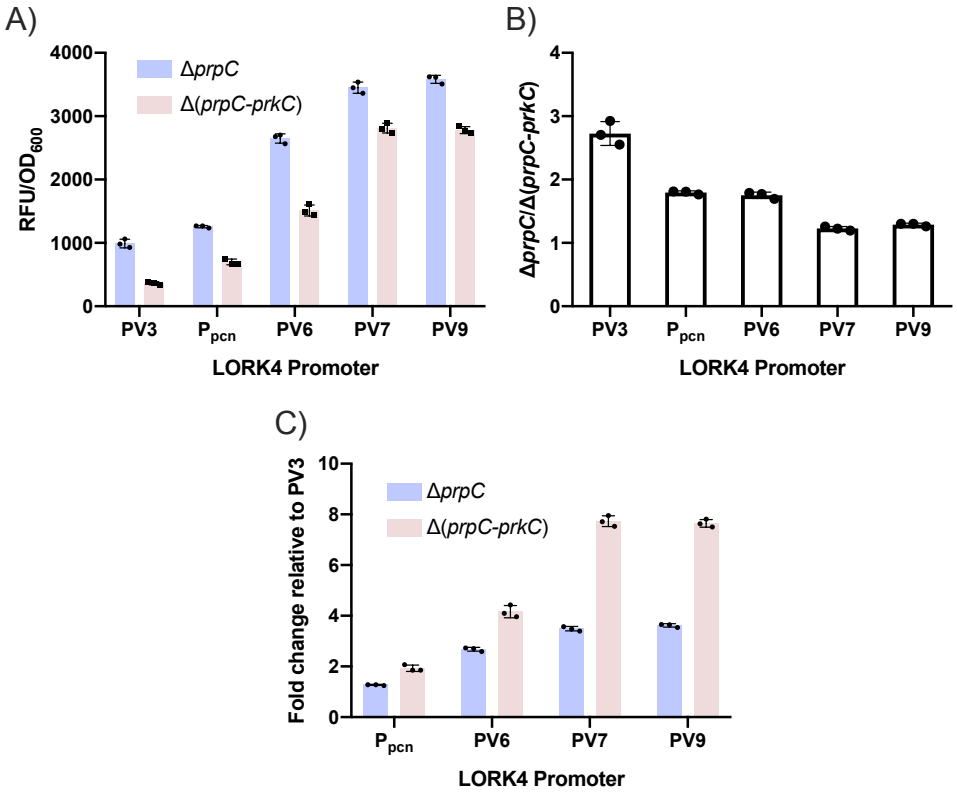


Figure 4.6: **LORK4 expression modulates sensitivity to kinase activity** LORK4 was expressed from P_{pcn} (IP_517 and 518), the promoter used in the other experiments, as well as PV3, PV6, PV7, and PV9, in the Δ*prpC* (CZ_12-15) and Δ*(prpC-prkC)* (CZ_16-19) backgrounds. LORK4 regulated YFP and values are reported when OD_{600}=0.8. Data is from a representative experiment. Bars are the means and error bars represent standard deviations. **A)** RFU normalized by OD_{600} (RFU/OD_{600}) for each promoter variant. **B)** Dynamic range of the sensors, reported as the Δ*prpC* signal divided by the Δ*(prpC-prkC)* signal. **C)** Fold change for each promoter variant relative to PV3 in the respective genetic background.

The sensitivity of LORK4 to kinase activity is affected by the expression of LORK4.

Our original *B. subtilis* LacI system, modeled after pDR111[135,136], expressed LacI from the promoter P_{pcn} (the penicillinase promoter, derived from *B. licheniformis*[137]), and so for the majority of testing in this paper we expressed our LORKs from P_{pcn}. To determine whether the expression level of LORK4 determined its sensitivity to kinase, we expressed LORK4 from four P_{veg} derived promoters. While all expression levels of LORK4 responded to kinase activity (Fig. 4.6A), the highest levels of LORK4 showed the greatest dynamic range (Fig. 4.6B). The decrease in dynamic range at lower expression levels appears to be due to the leakiness of the off state increasing more quickly than the signal from the ON state, as compared to PV1 (Fig. 4.6C).

4.3.3 PrkC Kinase Activity *In Vivo*

In vivo Phosphorylation of PrkC T290

The FRET sensor and LORK4 can be used to measure the phosphorylation state of PrkC T290 *in vivo*. Both the FRET sensor and LORK4 measure the phosphorylation state of a given substrate. To determine if our sensors worked, we have focused on the bulk response in the Δ*prpC* and Δ*(prpC-prkC)* backgrounds in minimal media. Quantitative proteomics work has suggested that in WT cells, the T290 PrkC autophosphorylation site is rarely phosphorylated[51]. That work was in bulk culture, while we can interrogate the pseudo-single cell phosphorylation state of T290 in WT, Δ*prpC*, and Δ*(prpC-prkC)* backgrounds (the data is pseudo-single cell because the image processing pipeline will occasionally count multiple cells as one cell, see Supp. Fig. C.1).

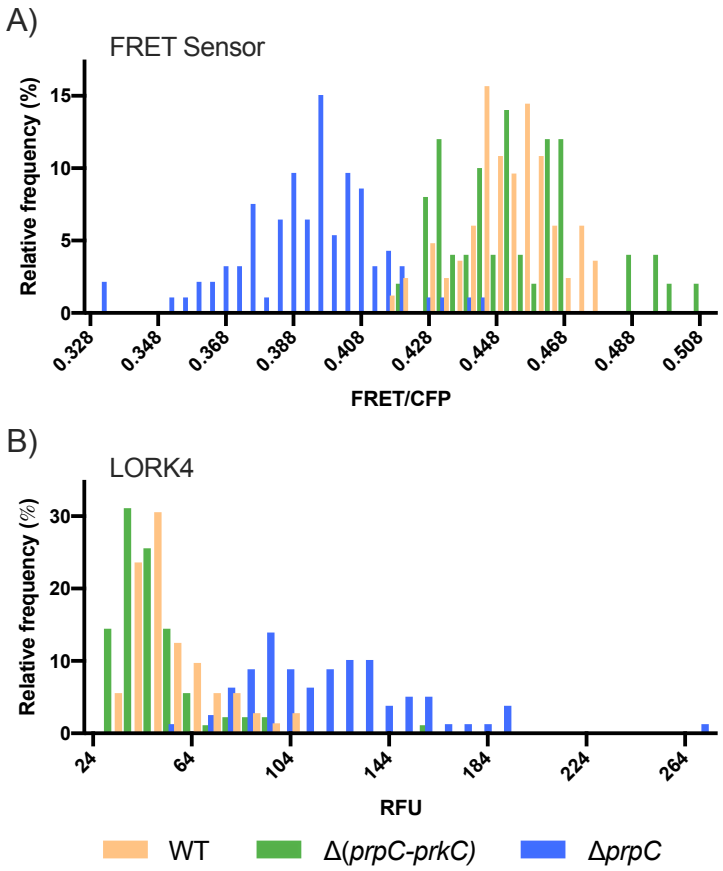


Figure 4.7: **FRET sensor and LORK4 show single cell kinase activity in Δ*prpC* but not in WT A)** Relative frequency of pseudo-single cell FRET signal normalized to CFP (FRET/CFP). The FRET sensor was grown in WT (IP_352), Δ*prpC* (IP_356) and Δ(*prpC-prkC*) (IP_360) backgrounds in MOPS minimal media. Data is of a single representative experiment. **B)** Relative frequency of pseudo-single cell RFU signal. LORK4 regulating YFP was grown in WT (IP_509), Δ*prpC* (IP_517) and Δ(*prpC-prkC*) (IP_518) backgrounds in MOPS Minimal media. Data is of a single representative experiment.

Both the FRET sensor and LORK4 show no difference between WT and the Δ(*prpC-prkC*) background. The FRET sensor shows that the Δ*prpC* and Δ(*prpC-prkC*) strains are completely separated between the 5th to 95th percentiles (Fig. 4.7A). LORK4 shows slightly worse separation, but the Δ*prpC* and Δ(*prpC-prkC*) strains are still completely separated between the 10th to 90th percentiles. For both the FRET sensor and LORK4, the WT strain is virtually identical to the Δ(*prpC-prkC*) strain (Figs. 4.7A and B). This suggests that the PrkC substrate is not maintained in a phosphorylated state under the conditions we observed. Consistent with this result, LORK4 signal is reduced to Δ(*prpC-prkC*) even

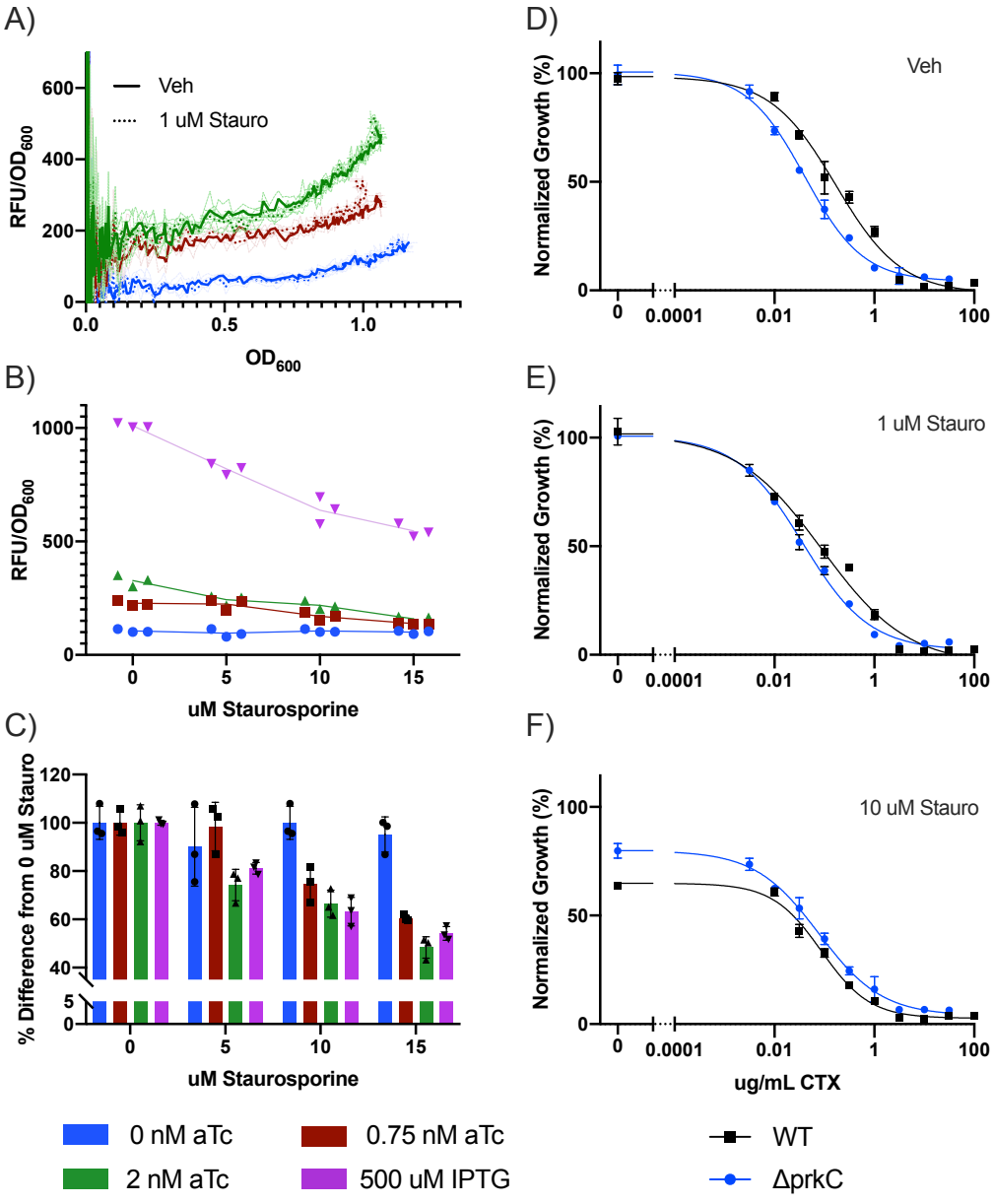
when PrpC is expressed minimally (such as from a leaky promoter) in a $\Delta prpC$ strain (see Supplemental Fig. C.4), which suggests that very low levels of PrpC will de-phosphorylate the prkC substrate.

Staurosporine's Effect on *B. subtilis*

LORK4 can be used to investigate the effects of putative STK inhibitors on *B. subtilis*. Staurosporine is the canonical Hanks-type STK inhibitor, and is routinely used in mammalian cells to inhibit STKs[138,139]. The active residues of bacterial Hanks-type STKs are well conserved, and it has been assumed that staurosporine functions as a kinase inhibitor in bacteria. While there is strong evidence that staurosporine affects bacterial growth[140–142], the available *in vivo* and *in vitro* data is not sufficient for determining whether the effect is due to kinase inhibition. The LORK4 system can indicate both changes in kinase activity as well as indicate changes in cellular physiology that affect the promoter it represses. To measure the LORK4 controlled promoter in a kinase independent manner, LORK4 can be induced with IPTG. As such, LORK4 is ideal for studying whether staurosporine is affecting bacteria through inhibiting STKs, or if staurosporine is causing widespread physiological changes.

We first tested staurosporine's ability to inhibit PrkC using LORK4 regulating mCherry with PrkC produced *in trans* from $P_{tetR}(prkC)$. LORK4 was chosen over the FRET sensor as it allows us to easily observe kinase activity across an entire growth curve. mCherry was chosen as there were concerns that staurosporine, which is a competitive ATP inhibitor, might interfere with luxABCDE (as it uses ATP to generate luminescence). YFP was not used because staurosporine significantly reduces the auto-fluorescence of LB in the YFP channel (data not shown).

Figure 4.8: **Staurosporine's effect on *B. subtilis* is not PrkC Dependent** In these experiments Vehicle (Veh) is DMSO. **A)**, **B)**, and **C)** LORK4 in the $\Delta(prpC\text{-}prkC)\mathrm{P}_{\mathrm{tetR}}(prkC)$ (IP_637) background was grown with different levels of staurosporine with no kinase induction (0 nM aTc), half-maximum kinase induction (0.75 nM aTc), maximum kinase induction (2 nM aTc), and LORK4 kinase independent induction (500 µM IPTG). LORK4 regulated mCherry. **A)** Veh and 1 µM staurosporine with different levels of kinase induction. Kinase independent induction is not shown. Signal is OD_{600} normalized fluorescence (RFU/OD_{600}). **B)** Four different levels of staurosporine (0 is the Veh controll) with the different aTc and IPTG concentrations. Values are taken at OD_{600}=0.8. **C)** Percent change by staurosporine for each aTc and IPTG induction condition relative to 0 µM. **D)**, **E)**, and **F)** WT and $\Delta prkC$ (JDB_1774) were grown with different levels of cefotaxime (CTX) and staurosporine. Percent growth after 4 hours relative to the 0 µg/mL CTX with Veh condition is reported. Dots are mean of (technical) triplicate and error bars are standard deviation. A 4-parameter logistics curve was fit using Prism. Representative experiment is shown. **D** Vehicle, **E** 1 µM staurosporine and **F** 10 µM staurosporine.


A)
Veh
1 uM Stauro
RFU/OD600
OD600
B)
RFU/OD600
uM Staurosporine
C)
% Difference from 0 uM Stauro
uM Staurosporine
D)
Veh
Normalized Growth (%)
E)
1 uM Stauro
Normalized Growth (%)
F)
10 uM Stauro
Normalized Growth (%)
ug/mL CTX
0 nM aTc
0.75 nM aTc
2 nM aTc
500 uM IPTG
WT
ΔprkC

Staurosporine reduces LORK4 regulated mCherry signal in a kinase independent manner. At 1 µM, staurosporine does not cause reduced fluorescence at any OD for both maximal levels of kinase (2 nM aTc) and half maximal levels of kinase (0.75 nM aTc) (Fig. 4.8A). At higher concentrations of staurosporine (5, 10, and 15 uM), staurosporine reduces mCherry fluorescence, but does so in a kinase independent manner (Fig. 4.8B). Both kinase induced and IPTG induced mCherry expression was reduced at high levels of Staurosporine, and there was no significant difference in the fractional reduction between the two methods of inducing mCherry (Fig. 4.8C).

To confirm that LORK4 was reporting a real result, we tested the cefotaxime sensitivity of the WT and Δ*prkC* strains with different levels of staurosporine. Previous studies have reported that PrkC homologs mediate beta-lactam sensitivity in other bacteria[47,141], and we demonstrate that this is also true in *B. subtilis*. We then used the kinase mediated beta-lactam sensitivity to examine the ability of staurosporine to inhibit PrkC. We tested growth inhibition by growing a micro-dilution assay in a plate reader and measuring growth after four hours (Fig. 4.8D-F). After the plate finished its growth in the plate reader, we read off the minimal inhibitory concentration (MIC) based on lowest drug concentration that showed no observable growth in all three wells (Table 4.1).

Table 4.1: **MIC of Cefotaxime** (µg/mL)

	Veh	1 µM Stauro	10 µM Stauro
Δ*prkC*	0.316	0.316	0.316
WT	3.16	3.16	0.316

MICs reported as µg/mL and based on micro-dilution assay read out after the experiment in Fig. 4.8 D-F.

Δ*prkC* is more sensitive than WT to CTX when grown with no staurosporine. When assayed by growth inhibition, Δ*prkC* shows about a 3.16 fold increased sensitivity to CTX relative to WT (Fig. 4.8D). This corresponded to a 10 fold decreased MIC of Δ*prkC* relative

to WT, with MICs of 0.316 µg/mL for Δ*prkC* and 3.16 µg/mL for WT (Table 4.1).

Staurosporine causes a growth defect in *B. subtilis* in a non-PrkC dependent manner. At low concentrations of staurosporine (1 µM), neither the WT nor Δ*prkC* strains show growth defects when grown with no CTX (Fig. 4.8E). Compared to growth with no staurosporine (Veh), in 1 µM staurosporine the WT strain may be slightly more sensitive to CTX concentrations of 0.01 µg/mL to 3.16 µg/mL. However, the growth defect is minimal and it does not affect the WT's MIC (Table 4.1). This is consistent with Fig. 4.8A, which shows no impact of 1 µM staurosporine on LORK4. In contrast, 10 µM staurosporine causes growth defects in both WT and Δ*prkC* regardless of the presence of CTX (Fig. 4.8F). Interestingly, 10 µM staurosporine reduces growth of WT below that of Δ*prkC*, but only reduces the MIC of WT to that of Δ*prkC* (Table 4.1). Staurosporine may inhibit PrkC, but only at levels where it also has non-specific impacts on *B. subtilis* physiology.

4.4 Discussion

In this work, we sought to develop the first (potentially) host-physiology independent, genetically encoded bacterial STK sensors. Both of the sensors we developed take advantage of the ability of an FHA domain to selectively bind a phosphorylated substrate and cause a conformational change. The first of the sensors, the FRET sensor, makes use of this conformational change to affect the distance between a FRET pair, and therefore produces a kinase dependent FRET signal. The second type of sensor, the LORKs, use the phosphorylation-dependent conformational change to "break" LacI, therefore tying LacI repression to kinase activity. The FRET sensors appear to be highly host-organism agnostic, since they were originally developed for mammalian cells, while the host agnosticism of the LORK sensors likely depends on the promoter that LORK regulates.

The FRET sensors we used was similar to a construct that had been designed to sense AuroraB in mammalian cells. We chose not to use a number of the more recent optimizations

of these mammalian sensors, due to concerns about their ability to respond dynamically to changes in signal. For *B. subtilis*, minor modifications were made to allow for high expression of the sensor. These modifications included: a strong promoter (P_{veg}), a canonical *B. subtilis* RBS, and codon optimization of the coding sequence. Three different 15 amino acid substrates were tested, of which only one worked. The substrate that worked (prkC) is a known PrkC autophosphorylation site and the only substrate that contained the sequence motif favored by FHA2. The other two sequences either did not have the FHA2 required TXXI motif (walR) or had not been validated as a PrkC substrate (optwalR).

FHA2 functions in bacterial cells, and should be considered as a modular tool for use in engineering. Fork Head Associated domains are common sensors of serine and threonine phosphorylations in eukaryotic systems[143]. However, as a result of the historical lack of work on bacterial STKs, there is little reported use of FHAs in bacterial systems[144]. The functioning of the FRET sensor argues that FHA2, and possibly other FHAs, can be treated as modular components in bacterial engineering projects.

Taking advantage of this modularity, we asked if FHA2 and a substrate can be added to LacI to make it sensitive to kinase activity. Previous work had shown that LacI is amenable to fusions to both the N- and C- termini as well as small insertions at a number of interior points (Fig. 4.1D). As there had been little work done to determine how close together the FHA2 and substrate needed to be to modify a host proteins conformation, we made a small library of all possible combinations (Supp. File C.1). A number of these constructs worked, representing a variety of distances between the FHA2 and substrate sequence, suggesting that neither physical nor sequence distance between these two components is the primary determinant of a functioning sensor (see Fig. 4.4A). The location of the FHA2 and substrate, besides determining the sensitivity to kinase activity, also determined the sensors ability to repress under non-phosphorylated conditions. Fortunately, LORK4 proved to be both sensitive to kinase activity as well as on par with LacI's repressive ability.

Both sensors respond to complemented PrkC, and respond to kinase activity in a sub-

strate dependent manner. When PrkC is produced at full induction (10 nM aTc) both the FRET sensor and LORK4 can be turned on "more" than in the Δ*prpC* background. This suggests that Δ*prpC* is within the dynamic range of both sensors and we would expect to observe both increases and decreases in kinase activity (if they occurred). Interestingly, LORK4 reaches full signal at 2 nM aTc, which is below full expression from the P_{tetR} promoter (see Supp. Fig. C.5). The LORK4 signal produced at 2 nM aTc is well below that of LORK4 induced with IPTG. This may suggest that all of the LORK4 proteins are phosphorylated (otherwise, 10 nM aTc would show an increased signal over 2 nM) but that they are only partially broken (since fully broken should look like IPTG induction). Partially broken LORK4 could be a result of the FHA2 and substrate varying between bound and unbound states.

The expression level of LORK4 can be used to modulate sensitivity to kinase activity. While it may be desirable to have the largest possible dynamic range between Δ*prpC* and Δ(*prpC-prkC*), other engineering applications may allow for a smaller dynamic range if higher signal can be achieved. By altering the expression level of LORK4, we demonstrate that higher levels of LORK4 increase the sensors dynamic range by reducing leaky expression from LORK4 controlled promoters in the OFF state, while reducing the signal in the ON state less. Conversely, reducing the amount of LORK4 reduces the sensors dynamic range, but results in a higher ON signal.

In concordance with previous research, the sensors show robust phosphorylation of the prkC substrate in the Δ*prpC* background, but no phosphorylation in the WT background. On a pseudo-single cell level, the sensors show almost complete separation between the Δ*prpC* and Δ(*prpC-prkC*) genotypes, although the FRET sensor shows better resolution of the two strains. However, no difference is observed between the Δ(*prpC-prkC*) and WT strains for either sensor. While the kinase is likely active in the WT strain, these sensors measure the phosphorylation state of their substrate. The substrate used here is a PrkC autophosphorylation site that is known to be targeted much more actively by PrpC than

PrkC. These sensors confirm this previous result, suggesting that under the conditions used in this work, the PrkC autophosphorylation site has very low levels of phosphorylation.

Staurosporine is used in the literature as a STK inhibitor, but we show that it has kinase independent effects at levels for which activity can be observed. While widely studied and verified in mammalian cell systems, less work has been done to verify staurosporine activity in bacterial systems. However, a number of papers have used staurosporine under the assumption that it inhibits bacterial STKs without affecting the rest of cell physiology[43,140–142]. LORK4, paired with $P_{tetR}(prkC)$, allowed us to use one strain to investigate the effect of staurosporine on both kinase activity (by inducing the kinase) as well as other aspects of the cells physiology (by inducing with IPTG and producing a kinase independent signal). We were able to show that at 1 µM, staurosporine had no effect on PrkC. At higher levels, staurosporine began to effect LORK4 signal, but it did so in a kinase independent manner. These results were confirmed by growth with both cefotaxime and staurosporine. Together, the results suggest that the kinase independent effects of staurosporine may not be separable from the kinase dependent effects in bacteria. As such, careful consideration should be given as to whether it is a good choice for bacterial experiments.

To our knowledge, this is the first description of (potentially) host-physiology independent phosphorylation sensors in bacteria. Both sensors work robustly in *B. subtilis* and allow us to ask questions that were previously off limits. While the two sensors work through the same mechanism, they lend themselves to different applications. Because the FRET sensor is internally controlled (the FRET signal is normalized to the CFP signal) it produces a consistent absolute signal across different experiments. This is particularly useful for looking at single cells as well changes in the shape of the signal distribution. In contrast, the LORK sensors have worse single cell resolution, but can regulate virtually any transcriptional signal. This makes it useful in anaerobic settings where fluorescent proteins are not usable. Further, the OR nature of LORK allows for robust internal controls as well as some potentially interesting synthetic biology applications. The ease of use of both sensors, and the organism

agnostic nature of the genetic parts, will hopefully encourage researches to port these systems to their favorite organisms.

4.5 Materials and Methods

Strains and Growth Conditions

For a full list of strains, see Supplemental Table C.2. *Bacillus subtilis* 168 *trpC2*, NC_000964, is the parent strain of the described *B. subtilis* genotypes and is considered wild type in this work. *B. subtilis* 168 Δ*prpC*, Δ*prkC*, and Δ(*prpC-prkC*) were gifts from Chet Price (University of California, Davis). For routine propagation, *B. subtilis* was grown in Lysogeny Broth (LB) Lennox and on LB Lennox agar plates. Only during initial strain construction were selective plates used, with antibiotics supplemented as follows: 5 µg/mL chloramphenicol, 10 µg/mL kanamycin, and MLS composed of 1 µg/mL erythromycin and 25 µg/mL lincomycin. For minimal media experiments (all experiments except where the use of LB is noted), MOPS Minimal Medium (Teknova) supplemented with 0.1% W/W glutamic acid and 40 µg/mL L-tryptophan was used. MOPS Minimal media was prepared fresh within 24 hours of use.

DH10β was used as the default cloning strain, and was routinely grown in LB Miller and on LB Miller agar plates. For *E. coli*, broth was supplemented with 50 µg/mL carbenicillin and agar plates were supplemented with 100 µg/mL ampicillin.

Cloning and Plasmid Construction

A complete list of plasmids can be found in Supplmental Table C.1 and plasmid maps can be found in Supplemental File C.2. All plasmids were designed for double crossover integration at standard insertion sites (i.e. *sacA*, *ganA*, and *amyE*). Golden Gate Assembly was used for most cloning. Q5 Hot Start polymerase was used for all PCR amplification. DNA for the coding sequences of CFP, FHA2, and YFP were ordered as gBlocks from IDT. For Golden

Gate reactions, 10xT4 Ligase Buffer (Promega), T4 Ligase (2,000,000 units/mL, NEB), and BSA (10 mg/mL, NEB) were used in all reactions. The appropriate restriction enzyme, either Eco31I, Esp3I, or SapI (Thermo FastDigest), was added. Substrates were added to plasmids by Golden Gate after annealing and phosphorylating pairs of oligos. For detailed information, see Appendix A. Briefly, complementary oligos were incubated with T4 Ligase Buffer (NEB) and T4 Poly Nucleotide Kinase (NEB), heated to boiling, and then slowly cooled to room temperature. Annealed oligos were then added to plasmids using Eco31I.

Golden Gate reactions were desalinated using drop dialysis (for a minimum of 10 minutes) and electroporated into DH10β Electrocompetent Cells (Thermo Fischer).

For point mutagenesis, a variation of NEBs Q5 Site-Directed Mutagenesis protocol was used. Briefly, primers were designed using NEBaseChanger, and then used to amplify the desired template. The amplicon was purified using the Zymo Clean and Concentrate kit and then 100 ng was incubated at room temperature for between five and 20 minutes in a 20 µL reaction comprised of 5x Quick Ligase Buffer (Promega) 0.5 µL T4 Ligase (2,000,000 units/mL, NEB), 0.5 µL T4 Poly Nucleotide Kinase, and 0.5 µL FastDigest DpnI (ThermoFisher). The reaction was purified using Zymo Clean and Concentrate then electroporated into DH10β Electrocompetent Cells (Thermo Fischer).

B. subtilis Strains Construction

B. subtilis was made competent using a modified two step competency protocol[145]. Briefly, *B. subtilis* was struck out on agar and grown overnight. The following morning, a colony was picked into 2 mL SpC media supplemented with L-tryptophan. The culture was incubated at 37 °C in a roller drum and after ∼3.5-4 hours the cells were diluted 1:10 into a baffled flask with SpII media supplemented with L-tryptophan. The culture was grown for 1.5 hours, spun down at 4000xG at room temperature for 10 minutes, and resuspended to 0.075x volume using the supernatant. If cells were too be stored for later use, 0.015x volume of 50% glycerol was added. Cells were aliquoted in microcentrifuge tubes, frozen in a dry ice-ethanol bath,

and stored at −80 °C for up to 12 months (although longer storage is likely possible). To use, aliquots were thawed in a 37 °C water bath.

To transform the cells, 6 µL of a plasmid miniprep or 1.2 µL of genomic DNA were added to 100 µL of competent cells plus 100 µL of SpIIE (the reaction can be halved if necessary). The culture was incubated in a roller drum at 37 °C for between one and two hours, and then the entire culture was plated on the appropriate antibiotic. Plates were incubated at 37 °C overnight, and then the colonies were restruck on fresh antibiotic plates and grown again overnight. The next day, single colonies were picked from the restreaks into 3 mL of LB Lennox. The cultures were grown for between four and six hours and used to make glycerol stocks (650 µL culture and 350 µL 50% glyecerol). After a plasmid transformation, 1.5 mL of culture was collected for genomic DNA isolation (to be used to confirm integration and crosses into desired genotypes).

Genomic DNA was isolated using a slightly modified Promega Genomic DNA Isolation kit. In short, the initial cell pellet was resuspended in 595 µL 50 mM EDTA and 5 µL of 20 mg/mL lysozyme was added. This was incubated at 37 °C for 5 to 10 minutes. The total time from addition of lysozyme to beginning the centrifuge step did not exceed 10 minutes. The RNAse step, and following 37 °C incubation step, were skipped.

Either phenotype tests (e.g. starch and iodine for *amyE*) or PCRs were used to confirm correct integrations of transformed plasmids.

Microscopy Experiments

For microscopy, strains were struck out from glycerol stocks and grown overnight. Individual colonies were grown in 3 mL of media. Cultures were grown at 37 °C in a roller drum until desired OD_{600} was reached. Unless otherwise noted, this meant an $OD_{600} \approx 0.1$ as measured by cuvette. 3 µL of culture (concentrated if necessary) was placed on a cover slip and then covered with an agar pad and slide (agar pad was composed of 50 µL of 1-1.2% agar in MOPS Minimal media). WT *B. subtilis* was imaged in all experiments as a non-fluorescent control.

Microscopy was performed on a Nikon TE2000 with a 100x Phase objective. A Lumencor SOLA light source was used. YFP (500/20x excitation, 535/30m emission), CFP (436/20x excitation, 480/40m emission), mCherry (560/40x excitation, 630/75m emission), and FRET (430/24x excitation, 535/30m emission) filters were used, as appropriate. NIS-Elements AR and a Hamamatsu ORCA-ERA camera were used to capture images. Phase contrast images were routinely imaged using a 150 ms exposure, while fluorescent channels used 400 ms

Images were analyzed using a MatLab script written by Alex Libby. For detailed description of the image analysis pipeline, see Supplemental Material and Methods. The Matlab script is Supplemental File C.3. Briefly, the script processes a multipage TIFF comprised of a phase contrast image (the first page) and any other channels that were imaged (subsequent pages). Bacteria in phase images have a characteristic white halo (see Supplemental Material), which can be easily detected with automated image analysis. The matlab script uses the halo to define a Region Of Interest (ROI) and create a mask for each bacteria. Each mask is then applied to the subsequent channels. For each mask, the mask's size, the total intensity for each channel, and the average intensity for each channel, are reported. The average intensity is used for all calculations.

For FRET experiments, each strain was imaged in the phase, FRET, CFP, and YFP channels. For each of the fluorescent channels, the background fluorescence was determined by averaging the values of all observed WT cells. The background fluorescence was then subtracted from the experimental fluorescent values. For each mask, the FRET/CFP signal was calculated. To determine the rawFRET/CFP signal, background adjusted FRET signal is divided by background adjusted CFP signal. A channel crosstalk constant is then subtracted from the rawFRET/CFP signal to get the final FRET/CFP signal for each mask. The channel crosstalk constant is a feature of the microscope's optical system and accounts for how signals from the individuals fluorophores (CFP and YFP) show up in the FRET signal when they should not. This constant is experimentally measured by performing the described experiment with strains containing either just CFP or YFP. For our system the constant is

0.5701 FRET/CFP. Details of how this is calculated can be found in Supplemental Materials and Methods.

LORK microscope experiments were imaged in the phase, and YFP channels. The YFP signal was background adjusted using the signal from the non-fluorescent control (WT).

Note that the above analysis pipeline frequently treats multiple linked (or spatially very close) cells as a single cell. As such, while much of the reported data is of single cells, some of it is also the average of multiple cells reported as a single cell, and so should be treated as "pseudo" single-cell.

Agar Plate Imaging

A custom built macroscope, courtesy of the Kishony Lab and the Harvard Department of Systems biology (see: `https://openwetware.org/wiki/Macroscope`), was used to take pictures of agar plates. Relevant strains were struck out, individual colonies were grown in 3 mL LB, and then 5 µL of culture was spotted on an LB plate. When the spots had grown to appropriate densities, brightfield and mCherry images were taken. To account for directionality of the light source, two differently orientated images were taken (although only one is shown). The image in Fig. 4.4 was edited in Fiji. First, red was replaced with magenta. Finally, the color thresholds were adjusted to increase contrast.

Plate Reader Experiments

For plate reader experiments, strains were struck out from glycerol stocks and grown overnight. In the morning, individual colonies were grown in 3 mL of media. Cultures were grown at 37 °C in a roller drum for 6 hours (MOPS Minimal media) or 4 hours (LB). For each well of a 96-well plate, 5 µL of culture was added to 145 µL of media. For experiments where both OD_{600} and another measurement were being taken, only columns 4 through 9 were used for cultures to avoid overly long read times (which in turn limits the shaking time and results in poor growth). The rest of the wells were filled with media, but not measured. For

experiments only measuring OD_{600}, the entire plate was used. Sterile, tissue culture treated, clear bottom, white plates with tops were used (VWR). The tops were taped down for the experiment.

Three plate readers were used: a Biotek H1 Synergy, a Biotek Neo, and a Biotek Neo2. All produce similar experimental results. Experiments were run for 12 hours and 10 minutes at 37 °C (1 °C gradient, top to bottom, for the H1 Synergy and 3 °C gradient for the Neo and Neo2) with orbital shaking (425 cpm, 3 mm). Luminescence was measured using a 1 second integration time. YFP and mCherry were measured using monochromators set to (485 nm excitation, 516 nm emission) and (579 nm excitation, 616 nm emission) respectively. All OD_{600} values were background adjusted using a media only control. Luminescent or fluoresecent signal was background adjusted using the WT as the non-luminescent/fluorescent control. Note that OD_{600} measured by the plate readers and through a cuvette are not directly comparable. All OD_{600} values reported here, unless otherwise noted, are from plate reader experiments.

Complementation Assays

For FRET sensor complementation assays, experiments were performed as described above, except the kinase complementation strain was grown in MOPS Minimal Media for 1 hour then split and supplemented with aTc (or lack thereof) as appropriate.

LORK kinase complentation assays were performed like the plate reader experiments, except that the relevant dosage of aTc was supplemented when the culture was added to the 96-well plate. For the dose response curve, the OD_{600} normalized fluorescence (RFU) signal was taken at OD_{600}=0.8 and a 4-paramater logistics curve was fit using Prism.

Staurosporine and Cefotaxime Assays

For staurosporine and cefotaxime, plate reader experiments were performed as described above, with relevant drugs being added to the 96-well plate at the same time as the cultures.

To measure growth inhibition due to cefotaxime and staurosporine (Fig. 4.8D-F), after 4 hours the OD_{600} of each strain was taken and normalized to the OD_{600} of the strains no cefotaxime, no staurosporine (Veh) control. The MIC of cefotaxime was taken by observing the lowest concentration for which cells had not grown during the plate reader experiment.

Chapter 5

Conclusion

Conclusion

Biology is conveyed through stories, and I have presented three such stories here. Each story was about a tool that dealt with some aspect of how biological stories are told. The first tool, a Potentially Organism Agnostic Knockout (POAK) system, expands the type of bacteria that we can work with, and therefore the settings of our biological stories. The second tool, DEcreasing the Selective Pressure Of phage Therapy (DESPOT), is used to attenuate pathogenic bacteria without selecting for resistant strains and provides a new type of ending for the stories we tell about bacterial infections. Finally, the third tool, a potentially Host Organism Agnostic Kinase Sensor (HOAKS), directly senses phosphorylation in bacteria and increases the ways we can witness biological stories.

Studying biology is often less about studying life, than it is about studying a few instantiations of life. The type of life we study is limited by our ability to manipulate organisms. In microbiology, one of the most significant limitations we face is the ability to make sequence specific, markerless knockouts in organisms other than those most commonly studied. Homologous recombination (HR), historically, has been the only method to make knockouts in most bacteria (Fig. 1.1C and D). Yet, HR is frequently inefficient to the point of uselessness, and requires additional strategies to remove the integrated antibiotic resistance[146,147]. In Chapter 2, I presented another approach to making knockouts: using a sequence-programmable endonuclease, Cas9, with an error-prone DNA repair system, non-homologous end-joining (NHEJ). I termed the combination of these two parts POAK.

To tests whether POAK was capable of making knockouts, I tested it in a well characterized organism (*Escherichia coli*) and a poorly characterized organism (*Weissella confusa*). POAK was able to make knockouts in both organisms, although it functioned better in *E. coli* (Fig. 2.5). The types of genetic edits were similar across multiple genes in both organisms (Fig. 2.6), and by making genetic alterations to a number of *W. confusa* sugar genes, we were able to gain insight into their function (Fig. 2.7). Additionally, POAK is unstable in both *E. coli* and *W. confusa* (Supp. Fig. A.7), meaning that POAK makes markerless,

sequence specific knockouts.

While POAK is a tool to expand the types of bacteria that biologists can engineer, we also aimed to change the act of engineering and use it for therapeutic ends. The general approach to bacterial infections is to kill the infecting bacteria. This has been the purview of antibiotics and, more recently, lytic phages. Unfortunately, the desired outcome (bacterial death) means that both of these therapies select for resistors and therefore future therapeutic failure. Non-bactericidal approaches would avoid this problem, yet of those being studied only fecal matter transplants have proven effective[35]. In Chapter 3, I described our use of temperate phages to engineer bacteria, in the hopes of eventually attenuating pathogenic bacterial virulence. This approach is called DESPOT.

DESPOT is a minimal CRISPRi system that uses sadCas9 to repress target genes (Fig. 3.2A). We showed that pDESPOT (the plasmid based system) effectively represses gene expression in both *E. coli* and *Salmonella typhimurium LT2*. We then integrated DESPOT into the *E. coli* temperate phage, lambda, resulting in lambda::DESPOT1. This modified temperate phage successfully repressed gene expression in *E. coli* (Fig. 3.3). Lambda::DESPOT1 was, however, unstable, as were two other variants with DESPOT integrated at slightly different locations (lambda::DESPOT2 and lambda::DESPOT3) (Fig. 3.4). Further, we have not yet been able to integrate DESPOT into the *S. typhimurium LT2* phage, P22. While we have been able to show that a temperate phage can be engineered to repress specific genes, significant work remains in developing DESPOT.

POAK and DESPOT are fundamentally genetic engineering technologies. Without being able to measure changes in phenotype, though, genetic engineering is worthless. There are countless phenotypes and many of them have measurement technologies. The technologies for measuring Serine/Threonine Kinase (STK) activity in bacteria, however, are lacking. STKs are implicated in a wide variety of cellular signaling and physiology regulation (Fig. 1.3C) yet the technologies for measuring phosphorylation activity in bacteria only work on bulk cultures (Fig. 1.3D) or they only indirectly observe kinase activity (Fig. 1.3E). In Chapter

4, I presented two bacterial phosphorylation sensors that were developed from modular parts first used in mammalian cells. The two sensors are called HOAKS.

HOAKS are based on the modularity of FHA2, which binds to phosphorylated substrates, and the substrate domain itself. By placing these two domains between a FRET pair, we were able to create a FRET sensor that responded to kinase activity in *Bacillus subtilis* in a substrate dependent manner (Fig. 4.2). The two parts were also able to regulate LacI's ability to repress transcription, as we showed with our complementary transcriptional Lactose OR Kinase (LORK) sensor. LORK4, in particular, showed robust response to kinase activity (Fig. 4.4C) while still remaining sensitive to lactose induction (Fig. 4.5A). The FRET sensor and LORK4 indicate that the substrate we used, the autophosphorylation site of *B. subtilis* STK PrkC, T290, is not phosphorylated in WT cells containing both kinase (PrkC) and phosphatase (PrpC) (Fig. 4.7). Finally, we used LORK4 to show that the canonical Hanks-type STK inhibitor, staurosporine, does not have a kinase dependent effect in *B. subtilis* (Fig. 4.8).

These tools have, hopefully, slightly expanded the biology we can study and the bacteria we can engineer. For each of these tools, I told a story. Yet, neither POAK nor HOAKS naturally fits into a story format. For all but the final piece of data (Fig. 2.7 and 4.8) those tools are not asking any questions. Instead, the rest of the data is describing the tools. In contrast, the most incomplete of the stories, DESPOT, naturally exists as one. DESPOT is the hero of the story, and we know the conflict that will occur: DESPOT will fight an infection. There does not need to be data on the conflict for the story to seem real, because the conflict and the outcomes are so vividly obvious. Yet, what does it mean if most of the data in this dissertation does not work to further biological stories?

Biology is defined by stories, but maybe it should not be. To an extent, it is necessary. Positioning our bio-molecule of interest as a hero in a story allows us to make a model of how the world works. That the antibiotic causes our hero protein to react and increase the cell's defenses makes sense. Without the ending, though, the story becomes meaningless. What

happens if the protein does not sense the antibiotic, or if it does not respond? Then, it is a negative result. In a similar vein, it is not a story if the entire time was spent describing the protein or the antibiotic, or anything else, for that matter. It is only a story if there is a protagonist, a conflict, and an ending.

The story format, therefore, excludes a significant amount of science. Both POAK and HOAKS would be more coherent accounts if they did not have their final data figures. The chapters would be descriptions of tools, and the stories would be left for people who want to understand the causal stories inside of cells. This type of publishing is not unknown to biology. Both crystal structures[117] and genomes[148] are reported as is. Yet, outside of those two areas, stories are required, and so tools must always have their endings. Even more of a problem than limiting tool development, though, is that the desire for stories precludes negative results. By requiring a causal story about what a protein does, it prevents biologists from saying that a protein does nothing under the observed conditions. That type of data only gets communicated between the lines of successful results.

There are not enough biologists for the absence of non-stories and negative data to be unproblematic. If hundreds of people worked on every problem, we would be able to read reality from between the lines of the positive results. But biologists rarely replicate each other's work, so each publication becomes the one true story. Publishing negative results, reporting tools without novel results, and replicating studies, are all ways to deal with this problem. At the moment, however, we are all selecting for stories, and so we get stories. Always.

Appendix A

Chapter 2 Supplemental Material

A.1 Supplemental Figures

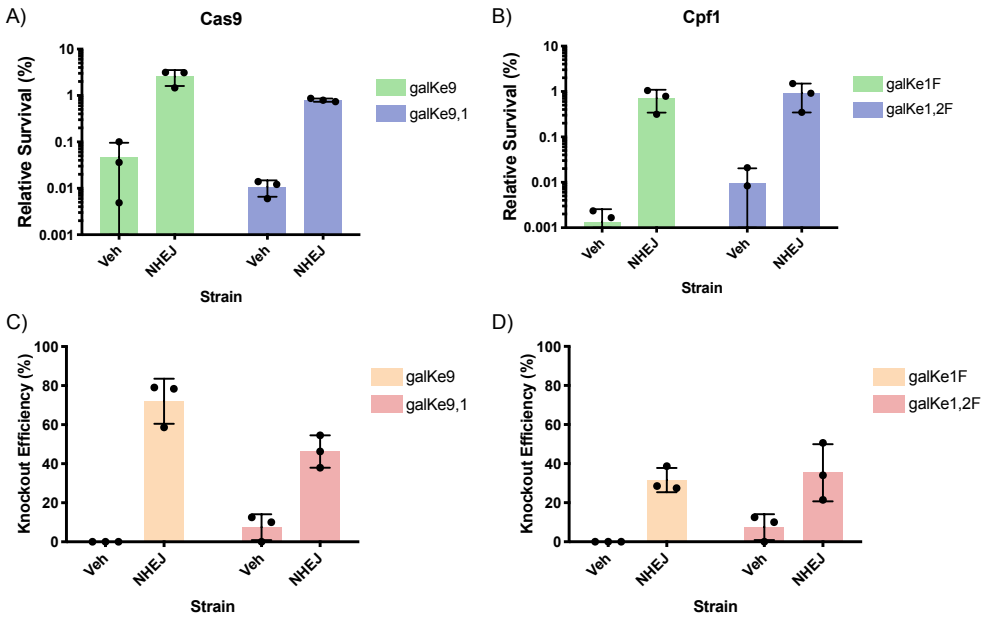


Figure A.1: **Functioning of a dual plasmid Cas9/Cpf1 and NHEJ system in *E. coli*** **A)** Relative survival when *E. coli*, either with a blank vehicle plasmid (Veh) or an NHEJ expressing plasmid (NHEJ), is transformed with Cas9 with a single gRNA targeting *galK* (galKe9) or two gRNAs (galKe9,1). **B)** Relative survival when *E. coli*, either with a blank plasmid (Veh) or an NHEJ expressing plasmid (NHEJ), is transformed with Cpf1 with a single gRNA targeting *galK* (galKe1F) or two gRNAs targeting (galKe1,2F). **C)** Knockout efficiency from **A**. **D)** Knockout efficiency from **B**.

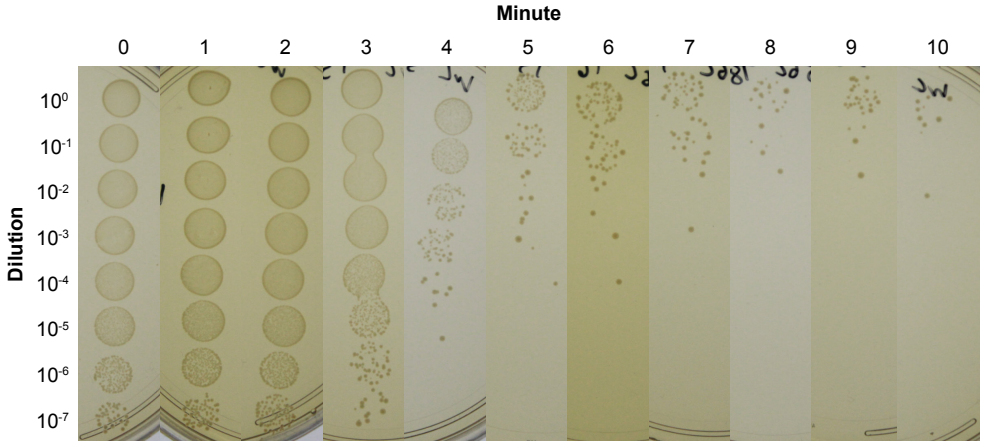


Figure A.2: ***W. confusa* dies rapidly at** 56 °C A *W. confusa* culture was incubated at 56 °C for the number of minutes indicated, then serially diluted and spotted on MRS.

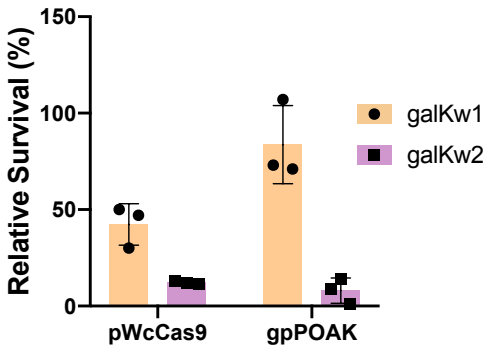


Figure A.3: **Cas9 mediated death is gRNA dependent in *W. confusa*** Relative survival of *W. confusa* after transformation with pWcCas9 and gpPOAK with either the galKw1 gRNA or the galKw2 gRNA. The bar is the mean of three transformations, and error bars are standard deviation.

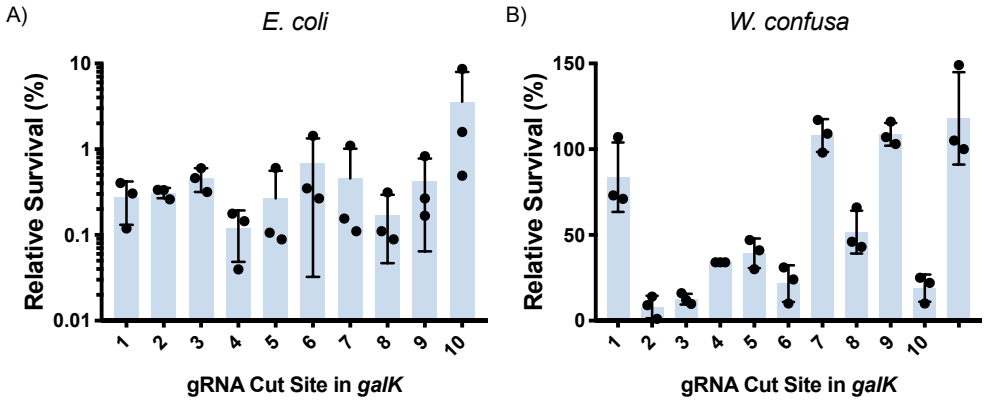


Figure A.4: **Effect of cut position on relative survival in POAK knockouts in *E. coli* and *W. confusa*** **A)** Relative survival when *E. coli* is transformed with gnPOAK containing a gRNA that cuts at the indicated position of galK. Bars are the mean of three transformations, error bars represent the standard deviation. **B)** Relative survival when *W. confusa* is transformed with gpPOAK containing a gRNA that cuts at the indicated position of galK. Cut sites 115 bp and 213 bp are gRNAs galk1 and galk2 respectively. Bars are the mean of three transformations, error bars represent the standard deviation.

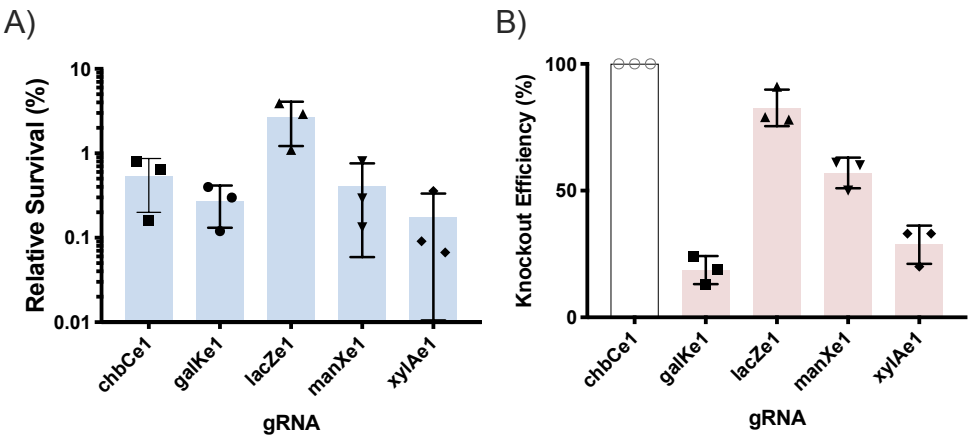


Figure A.5: **Survival and knockout efficiency of gnPOAK for five *E. coli* genes A)** Relative survival when *E. coli* is transformed with gnPOAK containing gRNAs targeting 5 different genes. Bars are mean of three transformations, error bars are standard deviation. **B)** Knockout efficiency for the same five gRNAs. *chbC* knockouts could not be assayed, as the gene does not confer a measurable phenotype in MG1655[149]. Bars are mean of three transformations, error bars are standard deviation.

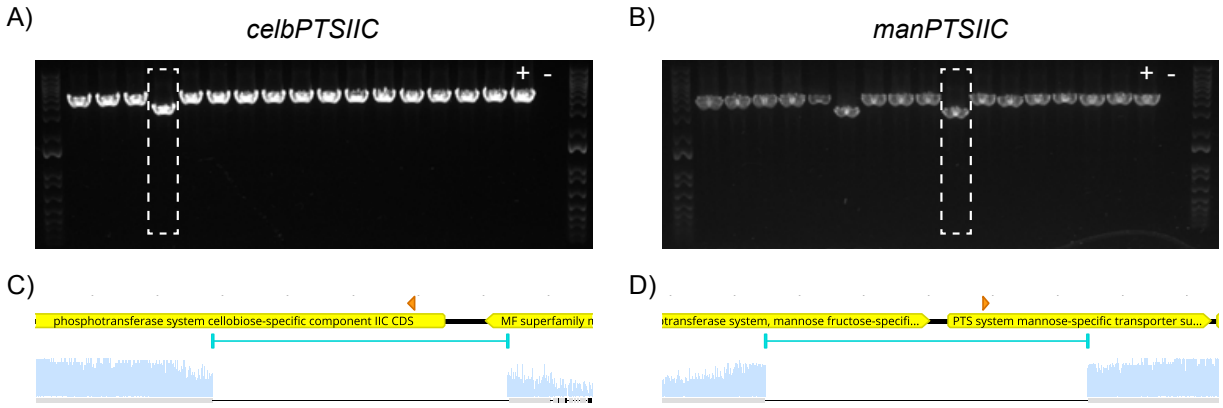


Figure A.6: **Knockouts of *celbPTSIIC* and *manPTSIIC* in *W. confusa* A)** and **B)** PCRs to isolate knockouts in **A)** *celbPTSIIC* and **B)** *manPTSIIC* from 16 colonies transformed with the celbPTSIICw1 gRNA and the manPTSIICw1 gRNA respectively. Positive and negative template controls are shown in the last lanes. **C)** and **D)** Sequencing of highlighted PCRS from **A** and **B** respectively. Orange arrows represent the cut sites, and the regions with no light blue sequencing coverage delineate the deleted sequences.

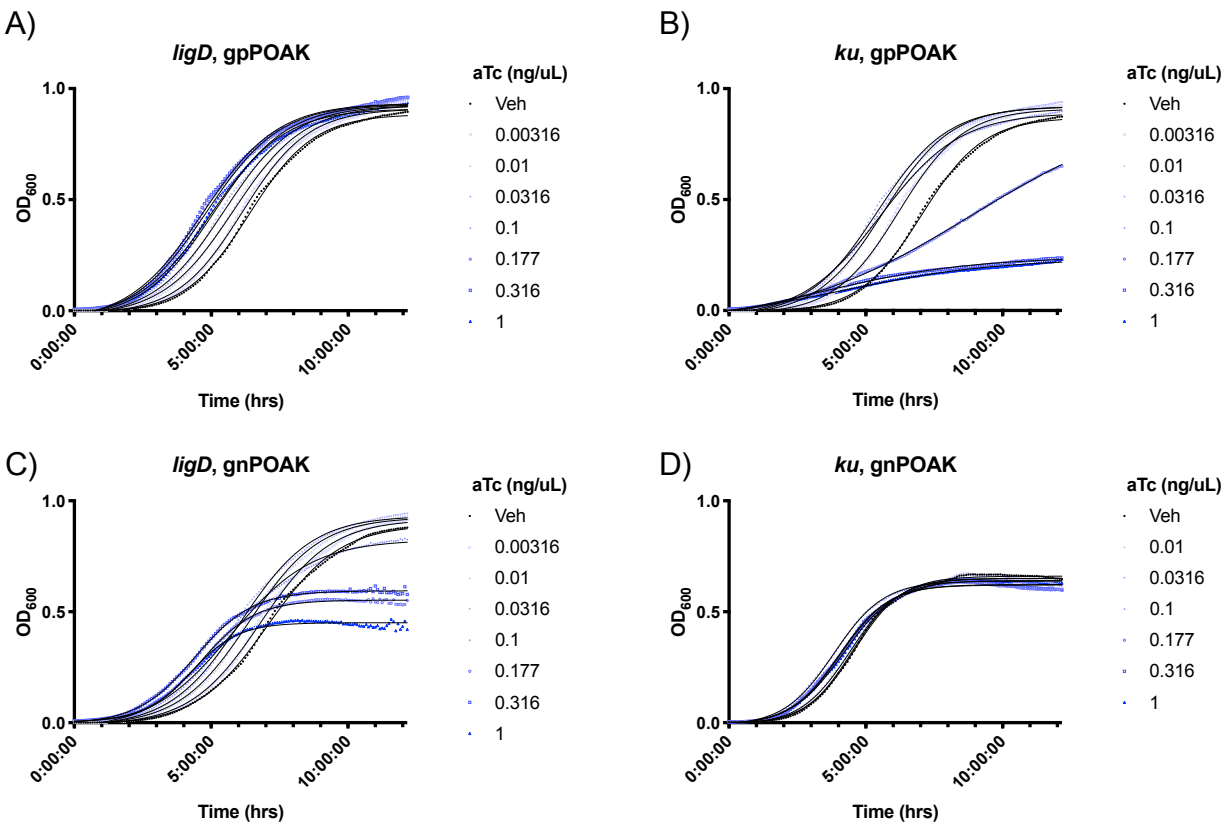


Figure A.7: **Ku expression reduces growth in *E. coli*** *E. coli* was grown with either a gpPOAK or gnPOAK plasmid that had an incomplete NHEJ system, such that one gene was deleted and the other (*ku* or *ligD*) was under P_{tetA} regulation. The *E. coli* was grown with selection as well as different concentrations of inducer (aTc). **A)** gpPOAK with only *ligD*. **B)** gpPOAK with only *ku*. **C)** gnPOAK with only *ligD*. **D)** gnPOAK with only *ku*.

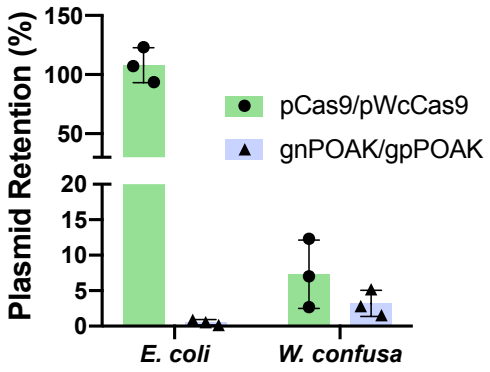


Figure A.8: **gnPOAK is unstable in *E. coli* and both pWcCas9 and gpPOAK are unstable in *W. confusa*** *E. coli* and *W. confusa* were grown with aTc and the relevant plasmid (pCas9 and gnPOAK for *E. coli*, pWcCas9 and gpPOAK for *W. confusa*). Cultures were plated on selective (for plasmid) and non-selective plates, and the percentage of colonies that grew on the selective places relative to non-selective is reported. Bars are the mean of three different cultures, and error bars represent the standard deviation.

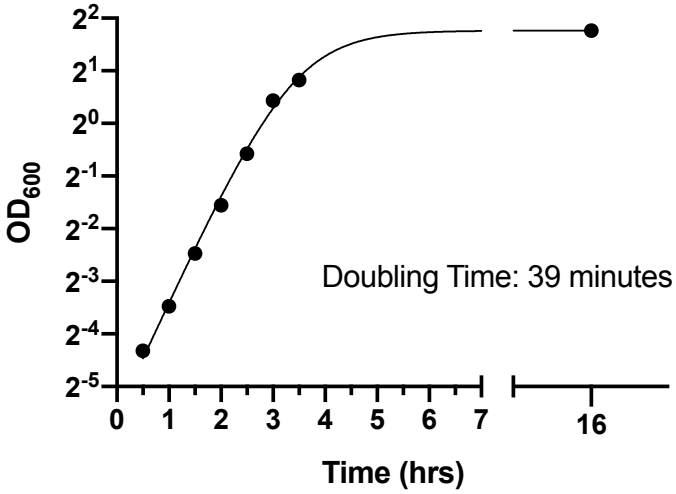


Figure A.9: ***W. confusa* doubles every 40 minutes** Representative growth curve of *W. confusa* when grown in MRSG at 37 °C. A four-parameter logistics curve is shown and was fit using Prism.

A.2 Strains, Plasmids, and gRNAs

Table A.1: **Strains**

Strain	Organism	Genotype
MG1655	*Escherichia coli*	Wild Type
DSM 20196	*Weissella Confusa*	Wild Type
BW25113	*Escherichia coli*	Δ(*araD-araB*)567, Δ*lacZ*4787 (::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514
JW1726-1Δ*kanR*	*Escherichia coli*	BW25113 Δ*chbC*513
JW0740-3Δ*kanR*	*Escherichia coli*	BW25113 Δ*galk*729
JW1806-1Δ*kanR*	*Escherichia coli*	BW25113 Δ*manX*741
JW3537-1Δ*kanR*	*Escherichia coli*	BW25113 Δ*xylA*748

Table A.2: **Plasmids**

Plasmid	Description	Strain	Resistance	Host
pVeh	Empty vector (i.e. Veh)	IP_8	Kn	Turbo
pNHEJ	pVeh+Ptrc(NHEJ)	IP_22	Kn	Turbo
pCas9sp	PproC(Cas9)	IP_23	Sp	Δgalk
pCas9spgalK9,1	PproC(Cas9), galk9,1	IP_24	Sp	Δgalk
pCas9spgalK9	PproC(Cas9), galk9	IP_25	Sp	Δgalk
pCpf1sp	PproC(Cpf1)	IP_34	Sp	Δgalk
pCpf1spgalK1F	PproC(Cpf1), galk1F	IP_36	Sp	Δgalk
pCpf1spgalK1,2F	PproC(Cpf1), galk1,2F	IP_37	Sp	Δgalk
gpPOAKtemp	PproC(Cas9), Pteta(ligd, ku), pBAV1K backbone	IP_104	Erm	DH5α
pCas9temp	PproC(Cas9), pBAV1K backbone	IP_109	Erm	DH10ß
pWcCas9	Pwc-eno(Cas9), pBAV1K backbone	IP_128	Erm	DH10ß
gpPOAK	Pwc-eno(Cas9), Pteta(ligd, ku), pBAV1K backbone	IP_132	Erm	DH10ß
pCas9	PproC(Cas9), pBBR1 origin, KnR	IP_245	Kn	ΔgalK
gnPOAK	PproC(Cas9), Pteta(ligd, ku), pBBR1 origin, KnR	IP_246	Kn	ΔgalK
pCpf1	PproC(Cpf1), pBBR1 origin, KnR	IP_247	Kn	ΔgalK
gnPOAK_Cpf1	PproC(Cpf1), Pteta(ligd, ku), pBBR1 origin, KnR	IP_248	Kn	Δgalk
pCas9galK9		IP_250	Kn	ΔgalK
pCas9galK9,1		IP_251	Kn	Δgalk
gnPOAKgalK9		IP_252	Kn	ΔgalK
gnPOAKgalK9,1		IP_253	Kn	Δgalk
pCpf1galK1F		IP_255	Kn	ΔgalK

Table A.2: (continued)

Plasmid	Description	Strain	Resistance	Host
pCpf1galK1,2F		IP_256	Kn	Δgalk
gnPOAK_Cpf1galK1F		IP_259	Kn	ΔgalK
gnPOAK_Cpf1galK1,2F		IP_260	Kn	Δgalk
gpPOAKcelbPTSIICw1		IP_281	Erm	DH10ß
gpPOAKgalKw1		IP_282	Erm	DH10ß
gpPOAKgalKw2		IP_283	Erm	DH10ß
gpPOAKgalKw3		IP_284	Erm	DH10ß
gpPOAKgalKw4		IP_285	Erm	DH10ß
gpPOAKgalKw5		IP_286	Erm	DH10ß
gpPOAKgalKw6		IP_287	Erm	DH10ß
gpPOAKgalKw7		IP_288	Erm	DH10ß
gpPOAKgalKw8		IP_289	Erm	DH10ß
gpPOAKgalKw9		IP_290	Erm	DH10ß
gpPOAKgalKw10		IP_291	Erm	DH10ß
gpPOAKgalKw11		IP_292	Erm	DH10ß
gpPOAKmaltPw1		IP_293	Erm	DH10ß
gpPOAKmanPTSIICw1		IP_294	Erm	DH10ß
gpPOAKxylAw1		IP_295	Erm	DH10ß
gnPOAKchbCe1		IP_296	Kn	ΔcelB
gnPOAKgalKe1		IP_297	Kn	Δgalk
gnPOAKgalKe2		IP_298	Kn	Δgalk
gnPOAKgalKe3		IP_299	Kn	Δgalk
gnPOAKgalKe4		IP_300	Kn	Δgalk
gnPOAKgalKe5		IP_301	Kn	Δgalk
gnPOAKgalKe6		IP_302	Kn	Δgalk
gnPOAKgalKe7		IP_303	Kn	Δgalk
gnPOAKgalKe8		IP_304	Kn	Δgalk

Table A.2: (continued)

Plasmid	Description	Strain	Resistance	Host
gnPOAKgalKe9		IP_305	Kn	Δgalk
gnPOAKgalKe10		IP_306	Kn	Δgalk
gnPOAKlacZe1		IP_307	Kn	Δgalk
gnPOAKmanXe1		IP_308	Kn	ΔmanY
gnPOAKxylAe1		IP_309	Kn	ΔxylA
pWcCas9galKw1		IP_315	Erm	DH10ß
pCas9galKe1		IP_317	Kn	Δgalk
pWcCas9galKw2		IP_336	Erm	DH10ß

Table A.3: **gRNAs**

Name	Sequence	Cut Site from ATG (bp)
galK1F	CCAACGCATTTGGCTACCCTGC	34
galK2F	TAAACCATCACAAGGAGCAGGA	1118
celbPTSIICw1	AATTGCGTATGCGATGCCAT	112
galKw1	GAGCACACCGATTATAATGG	115
galKw2	CTATTCAGCGAACTTCCCAG	213
galKw3	TTGGCTACCCCGTAACAAAA	332
galKw4	AATATGCTAGCTGACTTGAT	433
galKw5	GGCCTGTTCGTCTTCACCAA	557
galKw6	AAAGTACAACGAACGTCGTG	684
galKw7	GGCAGCGTCAGCGTCATTAA	786
galKw8	GCCCAAACGTTTCCAAATCG	897
galKw9	GGTGCTCGTATGACTGGTGC	1021
galKw10	ATTTCGGCGACGAAAAATGA	1144
galKw11	AGATTCACATAGTAGTCAAT	1208
maltPw1	AATGAGTACATGGGTATGCG	116
manPTSIICw1	ATTGGCAACTGGTCACCTAA	127
xylAw1	CCTAAGGTAGAATTTATCGG	119
chbCe1	AATGCCGTTAACCCTTGCGG	124
galKe1	GTTCACCAATCAAATTCACG	93
galKe2	CTGCCATCACGCGAACTTTA	195
galKe3	GGCTAACTACGTTCGTGGCG	285
galKe4	GCTTCACTGGAAGTCGCGGT	400
galKe5	GATCAGCTAATTTCCGCGCT	529
galKe6	CAGTAACTTCAAACGTACCC	642
galKe7	AACAGCGTTGAACTCTTCAA	749
galKe8	GCAAGGCGACCTGAAACGTA	858
galKe9	GGTGGCGTACGCATGACCGG	988

Table A.3: (continued)

Name	Sequence	Cut Site from ATG (bp)
galKe10	GAACAATATGAAGCAAAAAC	1081
lacZe1	TTACGCCAGCTGGCGAAAGG	96
manXe1	GTACTTTTCAATCAGCGTTT	127
xylAe1	CTTACCCAACACCAGTTCGT	88

A.3 gRNA Design and Cloning (for spCas9)

1. Use a computer to design gRNAs. PAM: 5' NGG. The length of the gRNAs doesn't really matter. I have used 22 bp before, but now usually use 20 bp because I arbitrarily decided to.

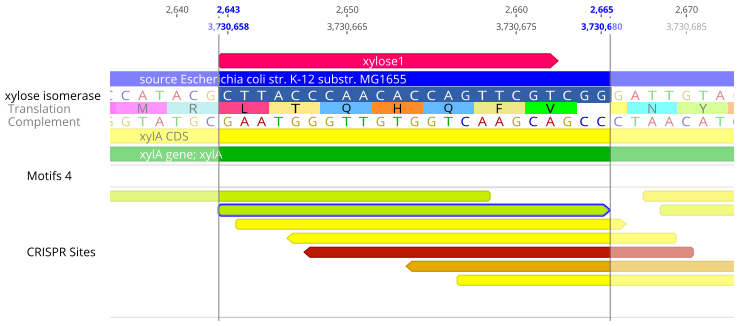


Figure A.10: gRNA of interest (in green)

2. Select the targeting region of the gRNA (everything besides the NGG)

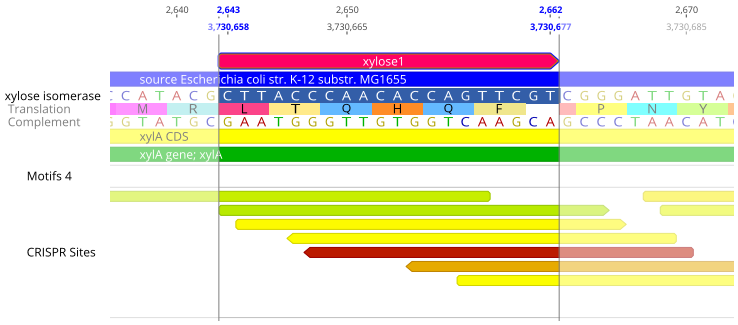


Figure A.11: Portion of desired gRNA to be copied (in pink)

3. Copy it into the gRNA template such that the 5' side of the gRNA is followed by a GTTT.

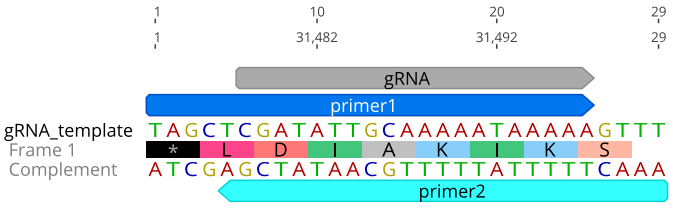


Figure A.12: gRNA template

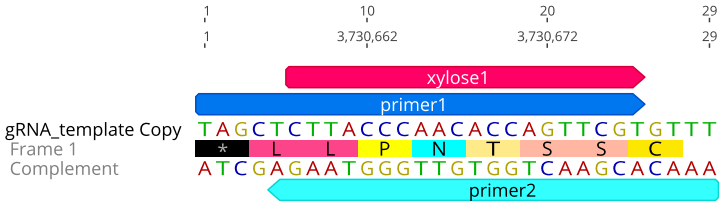


Figure A.13: Example gRNA copied into template

4. Order the oligos labeled as "primer 1" and "primer 2"

5. A day or two before you are going to be doing the cloning, make the following reaction:

Table A.4: Reaction for Oligo Phosphorylation and Annealing

1 µL	T4 Polynucleotide Kinase
2 µL	100 µM Oligo 1
2 µL	100 µM Oligo 2
4 µL	10x T4 Ligase Buffer
31 µL	ddH_2O

6. And use the following conditions

Table A.5: Thermocycle for Oligo Phosphorylation and Annealing

37 °C	30 min.
95 °C	5 min.
−1 °C/min	
25 °C	1 min.
4 °C	∞

7. Dilute reaction 1:50 in water, and use 1 µL of dilution per 20 µL golden gate reaction.

Appendix B

Chapter 3 Supplemental Material

B.1 Strains and Plasmids

Table B.1: **Strains**

Strain	Organism	Genotype
MG1655	*Escherichia coli*	Wild Type
C600	*Escherichia coli*	F-, thr-1, leuB6(Am), fhuA21, cyn-101, lacY1, glnX44(AS), λ-, e14-, rfbC1, glpR200(glpc), thiE1
EC001	*Escherichia coli*	MG1655 glmS::rfp,gfp
PAS709	Salmonella typhimurium LT2	ΔSP-1 ΔSP-2
PAS710	Salmonella typhimurium LT2	metA22 metE551 trpD2 ilv-452 leu- pro-(leaky) hsdLT6 hsdSA29 hsdB strA120 hsdR::pKD46
DH10ß	*Escherichia coli*	F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZM15 ΔlacX74 recA1 endA1 araD139 Δ(ara-leu)7697 galU galK λ– rpsL(StrR) nupG tonA
λ	Lamda Phage	Wild Type
P22	P22 Phage	Wild Type
DESPOT1	*Escherichia coli*	MG1655 glmS::rfp,gfp λ::DESPOT1
DESPOT2	*Escherichia coli*	MG1655 glmS::rfp,gfp λ::DESPOT2
DESPOT3	*Escherichia coli*	MG1655 glmS::rfp,gfp λ::DESPOT3

Table B.2: **Plasmids**

Name	Description	Resistance	Strain
pDESPOT	Pproc(sadCas9)	Cm	pLG16
pDESPOTrfp1	Pproc(sadCas9), rfp1	Cm	pLG18
pDESPOT1	Pproc(sadCas9), homology arms#1	Cm	IP_397
pDESPOT1galk1	Pproc(sadCas9), galk1	Cm	IP_398
pDESPOT1rfp1	Pproc(sadCas9), rfp1, homology arms#1	Cm	IP_406
pDESPOT2	Pproc(sadCas9), homology arms#2	Cm	IP_565
pDESPOT2rfp1	Pproc(sadCas9), rfp1, homology arms#2	Cm	IP_566
pDESPOT3	Pproc(sadCas9), homology arms#3	Cm	IP_567
pDESPOT3rfp1	Pproc(sadCas9), rfp1, homology arms#3	Cm	IP_568
pDESPOTp22	Pproc(sadCas9), *gtrA-C* homology, temperature sensitive	Cm	IP_646

Appendix C

Chapter 4 Supplemental Material

C.1 Supplemental Figures

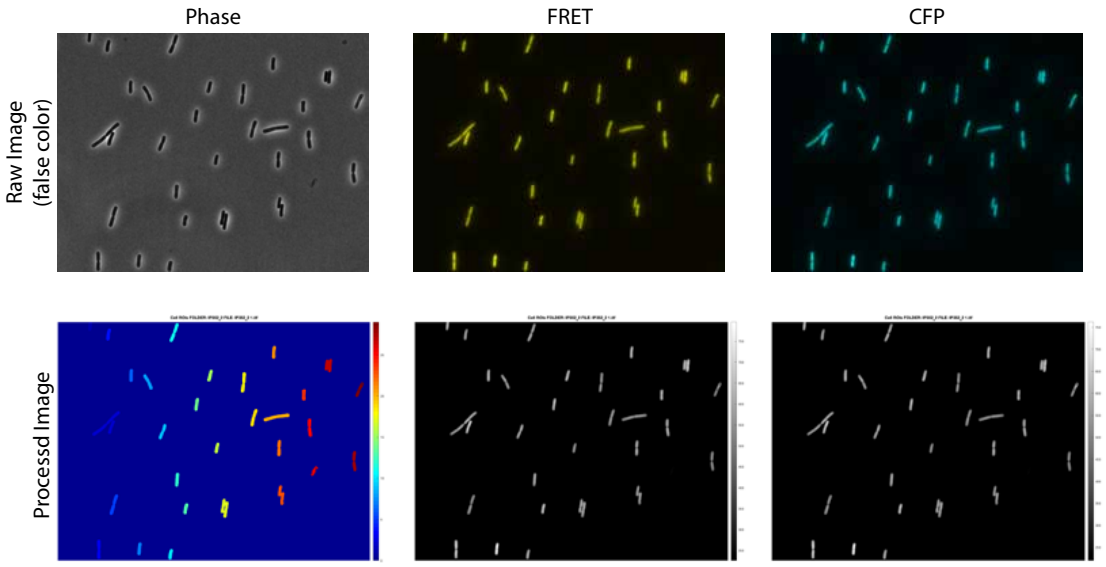


Figure C.1: **Image processing pipeline** Images at top are raw images (except for false coloring) in the channels used for the FRET sensor. Bottom row phase image shows the identified ROI for each cell and the arbitrary number given (rainbow scale bar). Note that if cells are close together (e.g. some of the yellow cells) they will be counted as one cell. The FRET and CFP lower row images shows the applied ROI from the phase image.

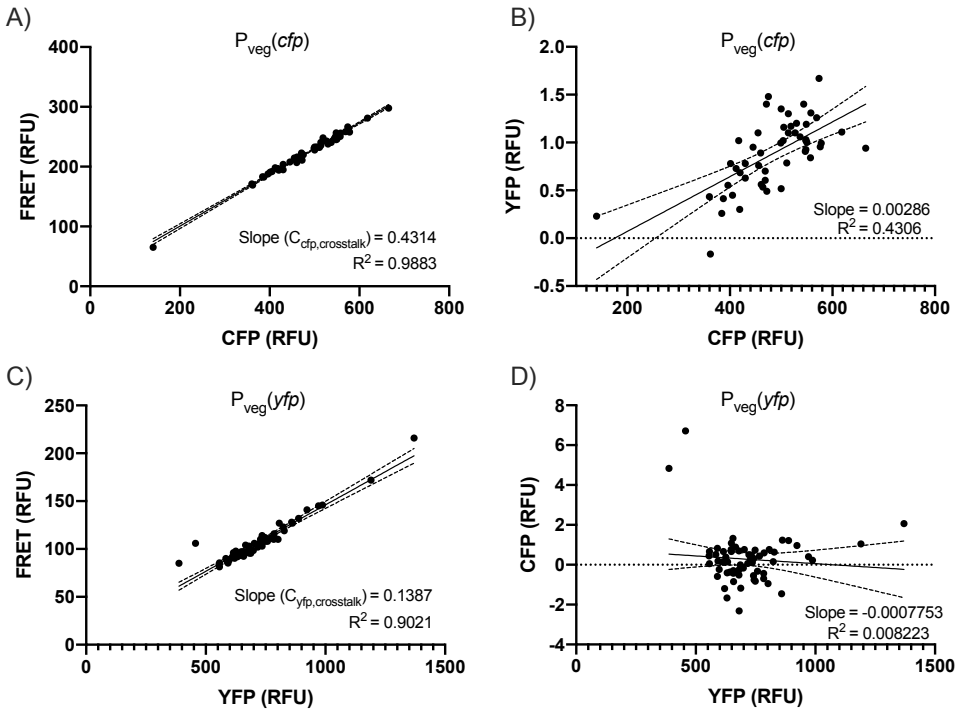


Figure C.2: **CFP to FRET channel crosstalk from CFP and YFP** CFP and YFP were expressed under the P_{veg} promoter in *B. subtillis* and fluorescence in all three channels (FRET, CFP, and YFP) was measured. Linear regressions were run using Prism, and representative experiments are shown. **A)** CFP and FRET signal from $P_{veg}(cfp)$. **B)** CFP and YFP signal from $P_{veg}(cfp)$. **C)** YFP and FRET signal from $P_{veg}(yfp)$. **D)** YFP and CFP signal from $P_{veg}(yfp)$.

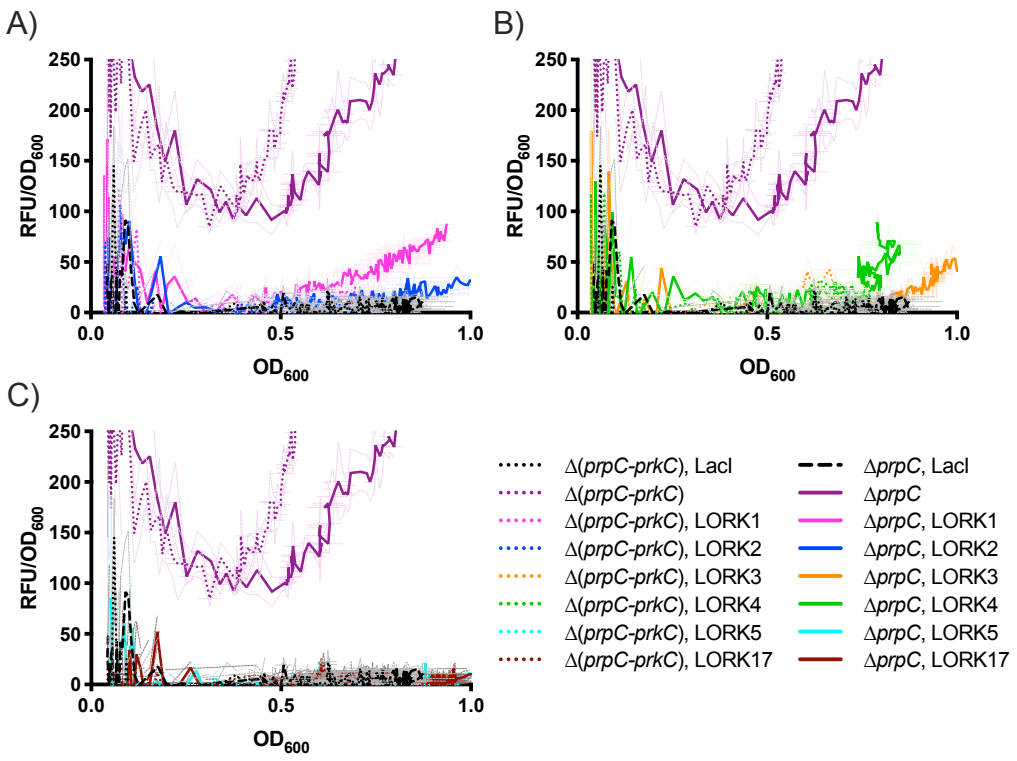


Figure C.3: **Initial LORK1 and 4 constructs respond to PrkC activity when grown in the plate reader** A selection of LORK variants, and LacI, expressed from P_{veg} regulating mCherry in the Δ*prpC* and Δ(*prpC-prkC*) genetic backgrounds. mCherry signal is reported as Relative Fluorescence Units (RFU) normalized by OD_{600} (RFU/OD_{600}). Bold lines are mean of three technical replicates, and faint lines are standard deviation. Representative experiments are shown. **A)** LORK1 and 2, plus controls. **B)** LORK3 and 4, plus controls. **C)** LORK5 and 17, plus controls.

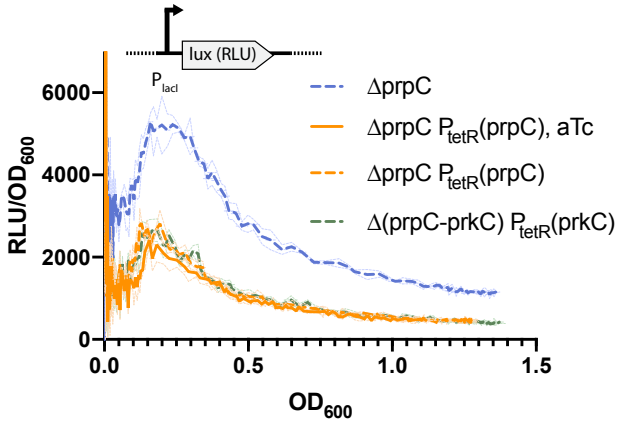


Figure C.4: **Leaky $\mathbf{P_{tetR}}$(*prpC*) reduces LORK4 signal in Δ*prpC* background** LORK4 in the Δ(*prpC*)P_{tetR}(*prpC*) background. LORK4 in the Δ(*prpC*) and Δ(*prpC-prkC*)P_{tetR}(*prkC*) backgrounds are shown as controls. LORK4 regulates *luxABCDE* and is reported as Relative Luminescence Units (RLU) normalized by OD_{600} (RLU/OD_{600}). All samples are uninduced, except the solid orange line which is induced with 10 nM aTc. Bold lines are mean of three technical replicates, and faint lines are standard deviation. Representative experiment is shown.

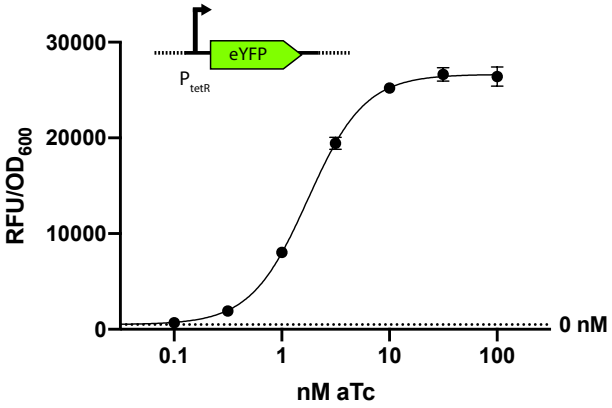


Figure C.5: **TetR system induction curve in *B. subtilis*** *B. subtilis* with TetR regulating YFP was grown with various levels of aTc in MOPS Minimal media. YFP is reported as Relative Fluorescence Units (RFU) normalized by OD_{600} (RFU/OD_{600}). Dots are the mean of three technical replicates, error bars are standard deviation, and a four parameter logistics fit is displayed. The dotted line is the mean of the uninduced (0 nM aTc) condition. The RLU/OD_{600} is reported at OD_{600}=0.8.

C.2 FRET Crosstalk Calculations

Raw fluorescence data for FRET experiments is processed (Fig. C.1) and then background subtracted using a non-fluorescent control (BS168). The background subtracted signal, $\frac{FRET_{raw}}{CFP_{raw}}$, is then corrected for signal into the FRET channel from YFP and CFP fluorophores that is not due to FRET (channel crosstalk).

$$\frac{FRET}{CFP} = \frac{FRET_{raw}}{CFP_{raw}} - \left(\frac{FRET}{CFP}\right)_{\text{due to CFP}} - \frac{YFP}{CFP}\left(\frac{FRET}{YFP}\right)_{\text{due to YFP}} \tag{C.1}$$

Eq. C.1 shows the general equation for correcting $\frac{FRET_{raw}}{CFP_{raw}}$ signal.

$$\left(\frac{FRET}{CFP}\right)_{\text{due to CFP}} = C_{cfp,crosstalk} \tag{C.2}$$

$$\left(\frac{FRET}{YFP}\right)_{\text{due to YFP}} = C_{yfp,crosstalk} \tag{C.3}$$

Eq. C.2 and C.3 show the crosstalk terms. Further, since our CFP and YFP are parts of the same protein, we assume that they are always equimolar. Therefore:

$$\frac{YFP}{CFP} = 1 \tag{C.4}$$

Eq. C.1 can be simplified to:

$$\frac{FRET}{CFP} = \frac{FRET_{raw}}{CFP_{raw}} - C_{cfp,crosstalk} - C_{yfp,crosstalk} \tag{C.5}$$

And the empirical values from Fig. C.2 can then be substituted,

$$\frac{FRET}{CFP} = \frac{FRET_{raw}}{CFP_{raw}} - 0.5701 \tag{C.6}$$

Eq. C.6 was used to adjust the data after background subtraction. No further manipulations were performed.

C.3 Plasmids and Strains

Table C.1: **Plasmids**

Strain	Description	Host
IP_318	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-bsaI cloning sites-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_319	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-iqedeemakaipii-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_320	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-igaddyvakpfstr-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_321	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-Igaddyvakpistr-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_323	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_324	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-stop codon-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_325	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-igaddyvtkpfstr-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_326	E. Coli[AmpR] B. Subtilis[sacA::Pveg(cfp-fha2-igaddyvtkpistr-yfp-stop) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_411	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK1) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß

Table C.1: (continued)

Strain	Description	Host
IP_412	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK2) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_413	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK3) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_414	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_415	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK5) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_422	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK12) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_423	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK13) Cm] shuttle vector for double crossover integration into B. Subtilis	BW25113 ΔpcnB
IP_427	E. Coli[AmpR] B. Subtilis[sacA::Pveg(LORK117) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_431	E. Coli[AmpR] B. Subtilis[sacA::Pveg(lacI) Cm] shuttle vector for double crossover integration into B. Subtilis	BW25113 ΔpcnB
IP_498	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(bsaI cloning sites) Ppcn(LORK1) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_499	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(bsaI cloning sites) Ppcn(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_500	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(bsaI cloning sites) Ppcn(lacI) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_501	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) Ppcn(LORK1) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß

Table C.1: (continued)

Strain	Description	Host
IP_502	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) Ppcn(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_503	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) Ppcn(lacI) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_505	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(mCherry) Ppcn(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_527	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(luxABCDE) Ppcn(LORK1) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_528	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(luxABCDE) Ppcn(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_529	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(luxABCDE) Ppcn(lacI) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_570	E. Coli[AmpR] B. Subtilis[sacA::PtetR(yfp) Ppcn(tetR) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_625	E. Coli[AmpR] B. Subtilis[ganA::PtetR(prkC) Ppcn(tetR) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_626	E. Coli[AmpR] B. Subtilis[ganA::PtetR(prpC) Ppcn(tetR) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
IP_638	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) Ppcn(nullLORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
CZ_8	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) PV3(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß

Table C.1: (continued)

Strain	Description	Host
CZ_9	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) PV6(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
CZ_10	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) PV7(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß
CZ_11	E. Coli[AmpR] B. Subtilis[sacA::Phyperspank(yfp) PV9(LORK4) Cm] shuttle vector for double crossover integration into B. Subtilis	DH10ß

Table C.2: **Strains**

Strain	Genotype	Cross
BS168 trpC2		
JBD_1773	BS168 trpC2 ΔprpC	
JBD_1774	BS168 trpC2 ΔprkC	
JBD_1775	BS168 trpC2 Δ(prpC-prkC)	
ELB_415	Bs168 sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm.	168xIP_323
ELB_416	Bs168 sacA::Pveg(cfp-fha2-iqedeemakaipii-yfp-stop) cm.	168xIP_319
ELB_419	Bs168 sacA::Pveg(cfp-fha2-igaddyvtkpistr-yfp-stop) cm.	168xIP_326
ELB_422	Bs168 sacA::Pveg(cfp-fha2-igaddyvtkpfstr-yfp-stop) cm.	168xIP_321
ELB_423	Bs168 sacA::Pveg(cfp-fha2-Igaddyvakpistr-yfp-stop) cm.	168xIP_321
ELB_424	Bs168 sacA::Pveg(cfp-fha2-igaddyvakpfstr-yfp-stop) cm.	168xIP_320
IP_352	Bs168 sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm.	168xELB_415
IP_353	Bs168 sacA::Pveg(cfp-fha2-iqedeemakaipii-yfp-stop) cm.	168xELB_416
IP_354	Bs168 sacA::Pveg(cfp-fha2-igaddyvtkpistr-yfp-stop) cm.	168xELB_419
IP_355	Bs168 sacA::Pveg(cfp-fha2-Igaddyvakpistr-yfp-stop) cm.	168xELB_423
IP_356	Bs168 d(prpC) sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm.	1773xELB_415
IP_357	Bs168 d(prpC) sacA::Pveg(cfp-fha2-iqedeemakaipii-yfp-stop) cm.	1773xELB_416
IP_358	Bs168 d(prpC) sacA::Pveg(cfp-fha2-igaddyvtkpistr-yfp-stop) cm.	1773xELB_419
IP_359	Bs168 d(prpC) sacA::Pveg(cfp-fha2-Igaddyvakpistr-yfp-stop) cm.	1773xELB_423
IP_360	Bs168 d(prpC-prkC) sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm.	1775xELB_415
IP_361	Bs168 d(prpC-prkC) sacA::Pveg(cfp-fha2-iqedeemakaipii-yfp-stop) cm.	1775xELB_416

Table C.2: (continued)

Strain	Genotype	Cross
IP_362	Bs168 d(prpC-prkC) sacA::Pveg(cfp-fha2-igaddyvtkpistr-yfp-stop) cm.	1775xELB_419
IP_363	Bs168 d(prpC-prkC) sacA::Pveg(cfp-fha2-Igaddyvakpistr-yfp-stop) cm.	1775xELB_423
IP_376	BS168 amyE::Phyperspank(mcherry) Sp	168xIP_377
IP_403	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp.	1773xIP_376
IP_404	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp.	1775xIP_376
IP_432	BS168 sacA::Pveg(LORK1) Cm	168x411
IP_433	BS168 sacA::Pveg(LORK2) Cm	168x412
IP_434	BS168 sacA::Pveg(LORK3) Cm	168x413
IP_435	BS168 sacA::Pveg(LORK4) Cm	168x414
IP_436	BS168 sacA::Pveg(LORK5) Cm	168x415
IP_443	BS168 sacA::Pveg(LORK12) Cm	168x422
IP_444	BS168 sacA::Pveg(LORK13) Cm	168x423
IP_448	BS168 sacA::Pveg(LORK17) Cm	168x427
IP_452	BS168 sacA::Pveg(lacI) Cm	168x431
IP_453	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK1) Cm	403x432
IP_454	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK1) Cm	404x432
IP_455	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK2) Cm	403x433
IP_456	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK2) Cm	404x433
IP_457	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK3) Cm	403x434
IP_458	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK3) Cm	404x434

Table C.2: (continued)

Strain	Genotype	Cross
IP_459	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK4) Cm	403x435
IP_460	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK4) Cm	404x435
IP_461	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK5) Cm	403x436
IP_462	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK5) Cm	404x436
IP_473	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK12) Cm	403x442
IP_474	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK12) Cm	404x442
IP_475	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK13) Cm	403x443
IP_476	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK13) Cm	404x443
IP_485	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK17) Cm	403x448
IP_486	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(LORK17) Cm	404x448
IP_493	BS168 Δ(prpC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(lacI) Cm	403x452
IP_494	BS168 Δ(prpC-prkC) amyE::Phyperspank(mcherry) Sp, sacA::Pveg(lacI) Cm	404x452
IP_508	BS168 sacA::Phyperspank(yfp) Ppcn(lacI) cm	168x503
IP_509	BS168 sacA::Phyperspank(yfp) Ppcn(LORK4) cm	168x502
IP_511	BS168 sacA::Phyperspank(mCherry) Ppcn(LORK4) cm	168x505
IP_515	BS168 ΔprpC sacA::Phyperspank(yfp) Ppcn(lacI) cm	1773x508
IP_516	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) Ppcn(lacI) cm	1775x508
IP_517	BS168 ΔprpC sacA::Phyperspank(yfp) Ppcn(LORK4) cm	1773x509

Table C.2: (continued)

Strain	Genotype	Cross
IP_518	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) Ppcn(LORK4) cm	1775x509
IP_521	BS168 ΔprpC sacA::Phyperspank(mCherry) Ppcn(LORK4) cm	1773x511
IP_522	BS168 Δ(prpC-prkC) sacA::Phyperspank(mCherry) Ppcn(LORK4) cm	1775x511
IP_530	BS168 sacA::Phyperspank(luxABCDE) Ppcn(LORK1) cm	168x527
IP_531	BS168 sacA::Phyperspank(luxABCDE) Ppcn(LORK4) cm	168x528
IP_532	BS168 sacA::Phyperspank(luxABCDE) Ppcn(lacI) cm	168x529
IP_545	BS168 ΔprpC sacA::Phyperspank(luxABCDE) Ppcn(LORK1) cm	1773x530
IP_546	BS168 Δ(prpC-prkC) sacA::Phyperspank(luxABCDE) Ppcn(LORK1) cm	1775x530
IP_547	BS168 ΔprpC sacA::Phyperspank(luxABCDE) Ppcn(LORK4) cm	1773x531
IP_548	BS168 Δ(prpC-prkC) sacA::Phyperspank(luxABCDE) Ppcn(LORK4) cm	1775x531
IP_549	BS168 ΔprpC sacA::Phyperspank(luxABCDE) Ppcn(lacI) cm	1773x532
IP_550	BS168 Δ(prpC-prkC) sacA::Phyperspank(luxABCDE) Ppcn(lacI) cm	1775x532
IP_572	BS168 sacA::Phyperspank(yfp) Ppcn(tetR) cm	168x570
IP_588	BS168 sacA::Phyperspank(yfp) PV3(LORK4) cm	168xCZ8
IP_589	BS168 sacA::Phyperspank(yfp) PV6(LORK4) cm	168xCZ9
IP_590	BS168 sacA::Phyperspank(yfp) PV7(LORK4) cm	168xCZ10
IP_591	BS168 sacA::Phyperspank(yfp) PV9(LORK4) cm	168xCZ11
IP_599	BS168 sacA::Pveg(cfp-fha2-igaddyvtkpfstr-yfp-stop) cm.	168xELB422

Table C.2: (continued)

Strain	Genotype	Cross
IP_600	BS168 ΔprpC sacA::Pveg(cfp-fha2-igaddyvtkpfstr-yfp-stop) cm.	1773xELB422
IP_601	BS168 Δ(prpC-prkC) sacA::Pveg(cfp-fha2-igaddyvtkpfstr-yfp-stop) cm.	1775xELB422
IP_602	BS168 sacA::Pveg(cfp-fha2-igaddyvakpfstr-yfp-stop) cm.	168xELB424
IP_603	BS168 ΔprpC sacA::Pveg(cfp-fha2-igaddyvakpfstr-yfp-stop) cm.	1773xELB424
IP_604	BS168 Δ(prpC-prkC) sacA::Pveg(cfp-fha2-igaddyvakpfstr-yfp-stop) cm.	1775xELB424
IP_627	BS168 ganA::PtetR(prkC) Ppcn(tetR) Erm	168x625
IP_628	BS168 ganA::PtetR(prpC) Ppcn(tetR) Erm	168x626
IP_630	BS168 Δ(prpC-prkC) sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm., ganA::PtetR(prkC) Ppcn(tetR) Erm	360x627
IP_631	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) Ppcn(LORK4) cm, ganA::PtetR(prkC) Ppcn(tetR) Erm	518x627
IP_632	BS168 Δ(prpC-prkC) sacA::Phyperspank(luxABCDE) Ppcn(LORK4) cm, ganA::PtetR(prkC) Ppcn(tetR) Erm	548x627
IP_634	BS168 ΔprpC sacA::Pveg(cfp-fha2-iqedeemtkaipii-yfp-stop) cm., ganA::PtetR(prpC) Ppcn(tetR) Erm	356x628
IP_635	BS168 ΔprpC sacA::Phyperspank(yfp) Ppcn(LORK4) cm, ganA::PtetR(prpC) Ppcn(tetR) Erm	517x628
IP_636	BS168 ΔprpC sacA::Phyperspank(luxABCDE) Ppcn(LORK4) cm, ganA::PtetR(prpC) Ppcn(tetR) Erm	547x628
IP_637	BS168 Δ(prpC-prkC) sacA::Phyperspank(mCherry) Ppcn(LORK4) cm, ganA::PtetR(prkC) Ppcn(tetR) Erm	522x627
IP_639	BS168 sacA::Phyperspank(yfp) Ppcn(nullLORK4) cm	168x638

Table C.2: (continued)

Strain	Genotype	Cross
IP_640	BS168 ΔprpC sacA::Phyperspank(yfp) Ppcn(nullLORK4) cm	1773x639
IP_641	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) Ppcn(nullLORK4) cm	1775x639
CZ_12	BS168 ΔprpC sacA::Phyperspank(yfp) PV3(LORK4) cm	1773x588
CZ_13	BS168 ΔprpC sacA::Phyperspank(yfp) PV6(LORK4) cm	1773x589
CZ_14	BS168 ΔprpC sacA::Phyperspank(yfp) PV7(LORK4) cm	1773x590
CZ_15	BS168 ΔprpC sacA::Phyperspank(yfp) PV9(LORK4) cm	1773x591
CZ_16	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) PV3(LORK4) cm	1775x588
CZ_17	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) PV6(LORK4) cm	1775x589
CZ_18	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) PV7(LORK4) cm	1775x590
CZ_19	BS168 Δ(prpC-prkC) sacA::Phyperspank(yfp) PV9(LORK4) cm	1775x591

Bibliography

1. Krzywinski, M. & Cairo, A. Points of view: Storytelling. eng. *Nat Methods* **10,** 687. ISSN: 1548-7105 (Electronic); 1548-7091 (Linking) (Aug. 2013).
2. Katz, Y. Against storytelling of scientific results. eng. *Nat Methods* **10,** 1045. ISSN: 1548-7105 (Electronic); 1548-7091 (Linking) (Nov. 2013).
3. Paulsen, I. T. *et al.* Role of Mobile DNA in the Evolution of Vancomycin-Resistant Enterococcus faecalis. *Science* **299,** 2071–2074. ISSN: 0036-8075 (2003).
4. Dixon, S. J., Costanzo, M., Baryshnikova, A., Andrews, B. & Boone, C. Systematic mapping of genetic interaction networks. eng. *Annu Rev Genet* **43,** 601–625. ISSN: 1545-2948 (Electronic); 0066-4197 (Linking) (2009).
5. Haas, F. L., Clark, J. B., Wyss, O. & Stone, W. S. Mutations and Mutagenic Agents in Bacteria. *The American Naturalist* **84,** 261–274. ISSN: 00030147, 15375323 (1950).
6. Alper, M. D. & Ames, B. N. Positive selection of mutants with deletions of the gal-chl region of the Salmonella chromosome as a screening procedure for mutagens that cause deletions. eng. *J Bacteriol* **121,** 259–266. ISSN: 0021-9193 (Print); 0021-9193 (Linking) (Jan. 1975).
7. Reyrat, J. M., Pelicic, V., Gicquel, B. & Rappuoli, R. Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infection and immunity* **66,** 4011–4017 (Sept. 1998).
8. Muñoz-López, M. & Garcıa-Pérez, J. DNA transposons: nature and applications in genomics. *Current genomics* **11,** 115–128 (Apr. 2010).
9. Rebollo, R., Romanish, M. T. & Mager, D. L. Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. *Annual Review of Genetics* **46.** PMID: 22905872, 21–42 (2012).
10. Kulasekara, H. D. Transposon mutagenesis. eng. *Methods Mol Biol* **1149,** 501–519. ISSN: 1940-6029 (Electronic); 1064-3745 (Linking) (2014).
11. Picardeau, M. Transposition of fly mariner elements into bacteria as a genetic tool for mutagenesis. eng. *Genetica* **138,** 551–558. ISSN: 1573-6857 (Electronic); 0016-6707 (Linking) (May 2010).
12. Wetmore, K. M. *et al.* Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. *mBio* **6** (ed Moran, M. A.) doi:10.1128/mBio.00306-15. eprint: https://mbio.asm.org/content/6/3/e00306-15.full.pdf. <https://mbio.asm.org/content/6/3/e00306-15> (2015).

13. Baym, M., Shaket, L., Anzai, I. A., Adesina, O. & Barstow, B. Rapid construction of a whole-genome transposon insertion collection for Shewanella oneidensis by Knockout Sudoku. *Nature Communications* **7,** (Nov. 2016).

14. Court, D. L., Sawitzke, J. A. & Thomason, L. C. Genetic Engineering Using Homologous Recombination. *Annual Review of Genetics* **36.** PMID: 12429697, 361–388 (2002).

15. Yu, D. *et al.* An efficient recombination system for chromosome engineering in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 5978–5983 (May 2000).

16. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97,** 6640–6645. ISSN: 0027-8424 (2000).

17. Van Pijkeren, J.-P. & Britton, R. A. High efficiency recombineering in lactic acid bacteria. *Nucleic acids research* **40,** e76–e76 (May 2012).

18. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173–1183. ISSN: 0092-8674 (2013).

19. Peters, J. M. *et al.* Mobile-CRISPRi: Enabling Genetic Analysis of Diverse Bacteria. *bioRxiv.* doi:10.1101/315499. eprint: https://www.biorxiv.org/content/early/2018/05/05/315499.full.pdf. <https://www.biorxiv.org/content/early/2018/05/05/315499> (2018).

20. Aminov, R. I. A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers in microbiology* **1,** 134, 134–134 (Dec. 2010).

21. Hughes, D. & Andersson, D. I. Evolutionary Trajectories to Antibiotic Resistance. *Annual Review of Microbiology* **71.** PMID: 28697667, 579–596 (2017).

22. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of anaesthesiology, clinical pharmacology* **33,** 300–305 (July 2017).

23. Kohanski, M. A., Dwyer, D. J. & Collins, J. J. How antibiotics kill bacteria: from targets to networks. *Nature reviews. Microbiology* **8,** 423–435 (June 2010).

24. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P & T : a peer-reviewed journal for formulary management* **40,** 277–283 (Apr. 2015).

25. Dever, L. A. & Dermody, T. S. Mechanisms of bacterial resistance to antibiotics. eng. *Arch Intern Med* **151,** 886–895. ISSN: 0003-9926 (Print); 0003-9926 (Linking) (May 1991).

26. Munita, J. M. & Arias, C. A. Mechanisms of Antibiotic Resistance. *Microbiology spectrum* **4,** (Apr. 2016).

27. Normark, B. H. & Normark, S. Evolution and spread of antibiotic resistance. eng. *J Intern Med* **252,** 91–106. ISSN: 0954-6820 (Print); 0954-6820 (Linking) (Aug. 2002).

28. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochimica et biophysica acta* **1794,** 808–816 (May 2009).

29. Lin, D. M., Koskella, B. & Lin, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World journal of gastrointestinal pharmacology and therapeutics* **8,** 162–173 (Aug. 2017).

30. Viertel, T. M., Ritter, K. & Horz, H.-P. Viruses versus bacteria-novel approaches to phage therapy as a tool against multidrug-resistant pathogens. eng. *J Antimicrob Chemother* **69,** 2326–2336. ISSN: 1460-2091 (Electronic); 0305-7453 (Linking) (Sept. 2014).

31. Duerkop, B. A., Huo, W., Bhardwaj, P., Palmer, K. L. & Hooper, L. V. Molecular Basis for Lytic Bacteriophage Resistance in Enterococci. *mBio* **7** (ed Miller, J. F.) doi:10.1128/mBio.01304-16. eprint: https://mbio.asm.org/content/7/4/e01304-16.full.pdf. <https://mbio.asm.org/content/7/4/e01304-16> (2016).

32. Horvath, P. & Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science* **327,** 167–170. ISSN: 0036-8075 (2010).

33. Opal, S. M. Non-antibiotic treatments for bacterial diseases in an era of progressive antibiotic resistance. *Critical care (London, England)* **20,** 397, 397–397 (Dec. 2016).

34. Sperandio, V. Novel approaches to bacterial infection therapy by interfering with bacteria-to-bacteria signaling. *Expert review of anti-infective therapy* **5,** 271–276 (Apr. 2007).

35. Brandt, L. J. & Reddy, S. S. Fecal microbiota transplantation for recurrent clostridium difficile infection. eng. *J Clin Gastroenterol* **45 Suppl,** S159–67. ISSN: 1539-2031 (Electronic); 0192-0790 (Linking) (Nov. 2011).

36. Mattila, E. *et al.* Fecal transplantation, through colonoscopy, is effective therapy for recurrent Clostridium difficile infection. eng. *Gastroenterology* **142,** 490–496. ISSN: 1528-0012 (Electronic); 0016-5085 (Linking) (Mar. 2012).

37. Hauser, A. R., Mecsas, J. & Moir, D. T. Beyond Antibiotics: New Therapeutic Approaches for Bacterial Infections. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **63,** 89–95 (July 2016).

38. Johnson, B. K. & Abramovitch, R. B. Small Molecules That Sabotage Bacterial Virulence. eng. *Trends Pharmacol Sci* **38,** 339–362. ISSN: 1873-3735 (Electronic); 0165-6147 (Linking) (Apr. 2017).

39. Stancik, I. A. *et al.* Serine/Threonine Protein Kinases from Bacteria, Archaea and Eukarya Share a Common Evolutionary Origin Deeply Rooted in the Tree of Life. eng. *J Mol Biol* **430,** 27–32. ISSN: 1089-8638 (Electronic); 0022-2836 (Linking) (Jan. 2018).

40. Hanks, S. K. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome biology* **4,** 111–111 (2003).

41. Munoz-Dorado, J., Inouye, S. & Inouye, M. A gene encoding a protein serine/threonine kinase is required for normal development of M. xanthus, a gram-negative bacterium. eng. *Cell* **67,** 995–1006. ISSN: 0092-8674 (Print); 0092-8674 (Linking) (Nov. 1991).

42. Libby, E. A., Goss, L. A. & Dworkin, J. The Eukaryotic-Like Ser/Thr Kinase PrkC Regulates the Essential WalRK Two-Component System in Bacillus subtilis. *PLOS Genetics* **11,** 1–22 (June 2015).

43. Shah, I. M., Laaberki, M.-H., Popham, D. L. & Dworkin, J. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell* **135,** 486–496 (Oct. 2008).

44. Nagarajan, S. N. *et al.* Protein kinase A (PknA) of Mycobacterium tuberculosis is independently activated and is critical for growth in vitro and survival of the pathogen in the host. *The Journal of biological chemistry* **290,** 9626–9645 (Apr. 2015).

45. Pensinger, D. A. *et al.* The Listeria monocytogenes PASTA Kinase PrkA and Its Substrate YvcK Are Required for Cell Wall Homeostasis, Metabolism, and Virulence. *PLOS Pathogens* **12,** 1–28 (Nov. 2016).

46. Jarick, M. *et al.* The serine/threonine kinase Stk and the phosphatase Stp regulate cell wall synthesis in Staphylococcus aureus. *Scientific reports* **8,** 13693, 13693–13693 (Sept. 2018).

47. Beltramini, A. M., Mukhopadhyay, C. D. & Pancholi, V. Modulation of cell wall structure and antimicrobial susceptibility by a Staphylococcus aureus eukaryote-like serine/threonine kinase and phosphatase. *Infection and immunity* **77,** 1406–1416 (Apr. 2009).

48. White, F. M. & Wolf-Yadlin, A. Methods for the Analysis of Protein Phosphorylation–Mediated Cellular Signaling Networks. *Annual Review of Analytical Chemistry* **9.** PMID: 27049636, 295–315 (2016).

49. Lin, M.-H. *et al.* Phosphoproteomics of Klebsiella pneumoniae NTUH-K2044 reveals a tight link between tyrosine phosphorylation and virulence. eng. *Mol Cell Proteomics* **8,** 2613–2623. ISSN: 1535-9484 (Electronic); 1535-9476 (Linking) (Dec. 2009).

50. Lin, M.-H., Sugiyama, N. & Ishihama, Y. Systematic profiling of the bacterial phosphoproteome reveals bacterium-specific features of phosphorylation. *Science Signaling* **8,** rs10–rs10. ISSN: 1945-0877 (2015).

51. Ravikumar, V. *et al.* Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. eng. *Mol Cell Proteomics* **13,** 1965–1978. ISSN: 1535-9484 (Electronic); 1535-9476 (Linking) (Aug. 2014).

52. Mijakovic, I. & Macek, B. Impact of phosphoproteomics on studies of bacterial physiology. *FEMS Microbiology Reviews* **36,** 877–892 (2012).

53. Buddelmeijer, N., Aarsman, M., Den Blaauwen, T., *et al.* Immunolabeling of Proteins in situ in Escherichia coli K12 Strains. *Molecular Microbiology* (2013).

54. Kosako, H. Phos-tag Western blotting for detecting stoichiometric protein phosphorylation in cells. <`http://dx.doi.org/10.1038/nprot.2009.170`> (Sept. 2009).

55. Zhang, C. *et al.* The Eukaryote-Like Serine/Threonine Kinase STK Regulates the Growth and Metabolism of Zoonotic Streptococcus suis. *Frontiers in Cellular and Infection Microbiology* **7,** 66. ISSN: 2235-2988 (2017).

56. Bertolin, G. *et al.* A FRET biosensor reveals spatiotemporal activation and functions of aurora kinase A in living cells. *Nature communications* **7,** 12674, 12674–12674 (Sept. 2016).

57. Fuller, B. G. *et al.* Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature* **453,** 1132–1136 (June 2008).

58. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. eng. *Mol Syst Biol* **2,** 2006.0008. ISSN: 1744-4292 (Electronic); 1744-4292 (Linking) (2006).

59. Goodall, E. C. A. *et al.* The Essential Genome of Escherichia coli K-12. *mBio* **9,** e02096–17 (Feb. 2018).

60. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science (New York, N.Y.)* **343,** 84–87 (Jan. 2014).

61. Liu, H. & Deutschbauer, A. M. Rapidly moving new bacteria to model-organism status. eng. *Curr Opin Biotechnol* **51,** 116–122. ISSN: 1879-0429 (Electronic); 0958-1669 (Linking) (June 2018).

62. Hiom, K. DNA Repair: Common Approaches to Fixing Double-Strand Breaks. *Current Biology* **19,** R523–R525. ISSN: 0960-9822 (2009).

63. Mladenov, E., Magin, S., Soni, A. & Iliakis, G. DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. eng. *Semin Cancer Biol* **37-38,** 51–64. ISSN: 1096-3650 (Electronic); 1044-579X (Linking) (June 2016).

64. Su, T. *et al.* A CRISPR-Cas9 Assisted Non-Homologous End-Joining Strategy for One-step Engineering of Bacterial Genome. *Scientific Reports* **6,** (Nov. 2016).

65. Bortesi, L. *et al.* Patterns of CRISPR/Cas9 activity in plants, animals and microbes. *Plant biotechnology journal* **14,** 2203–2216 (Dec. 2016).

66. Bowater, R. & Doherty, A. J. Making Ends Meet: Repairing Breaks in Bacterial DNA by Non-Homologous End-Joining. *PLOS Genetics* **2,** 1–7 (Feb. 2006).

67. Matthews, L. A. & Simmons, L. A. Bacterial Nonhomologous End Joining Requires Teamwork. *Journal of Bacteriology* **196,** 3363–3365. ISSN: 0021-9193 (2014).

68. Pitcher, R. S., Brissett, N. C. & Doherty, A. J. Nonhomologous end-joining in bacteria: a microbial perspective. eng. *Annu Rev Microbiol* **61,** 259–282. ISSN: 0066-4227 (Print); 0066-4227 (Linking) (2007).

69. Lu, S. *et al.* Single-Homology-Arm Linear DNA Recombination by the Nonhomologous End Joining Pathway as a Novel and Simple Gene Inactivation Method: a Proof-of-Concept Study in Dietzia sp. Strain DQ12-45-1b. eng. *Appl Environ Microbiol* **84.** ISSN: 1098-5336 (Electronic); 0099-2240 (Linking). doi:10.1128/AEM.00795-18 (Oct. 2018).

70. Tong, Y., Charusanti, P., Zhang, L., Weber, T. & Lee, S. Y. CRISPR-Cas9 Based Engineering of Actinomycetal Genomes. *ACS Synthetic Biology* **4.** PMID: 25806970, 1020–1029 (2015).

71. Malyarchuk, S. *et al.* Expression of Mycobacterium tuberculosis Ku and Ligase D in Escherichia coli results in RecA and RecB-independent DNA end-joining at regions of microhomology. *DNA repair* **6,** 1413–1424 (Oct. 2007).

72. Adams, B. L. The Next Generation of Synthetic Biology Chassis: Moving Synthetic Biology from the Laboratory to the Field. *ACS Synthetic Biology* **5.** PMID: 27665861, 1328–1330 (2016).

73. Calero, P. & Nikel, P. I. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microbial Biotechnology* **12,** 98–124 (2019).

74. Rivas-Marın, E., Canosa, I., Santero, E. & Devos, D. P. Development of Genetic Tools for the Manipulation of the Planctomycetes. *Frontiers in microbiology* **7,** 914, 914–914 (June 2016).

75. Abriouel, H. *et al.* The controversial nature of the Weissella genus: technological and functional aspects versus whole genome analysis-based pathogenic potential for their application in food and health. *Frontiers in Microbiology* **6,** 1197. ISSN: 1664-302X (2015).

76. Fusco, V. *et al.* The genus Weissella: taxonomy, ecology and biotechnological potential. eng. *Front Microbiol* **6,** 155. ISSN: 1664-302X (Print); 1664-302X (Linking) (2015).

77. Ku, H.-J., Park, M. S. & Lee, J.-H. Characterization of a minimal pKW2124 replicon from Weissella cibaria KLC140 and its application for the construction of the Weissella expression vector pKUCm1. *Frontiers in microbiology* **6,** 35, 35–35 (Feb. 2015).

78. Antoine, R. & Locht, C. Isolation and molecular characterization of a novel broad-host-range plasmid from Bordetella bronchiseptica with sequence similarities to plasmids from gram-positive organisms. eng. *Mol Microbiol* **6,** 1785–1799. ISSN: 0950-382X (Print); 0950-382X (Linking) (July 1992).

79. Beck, C. F., Mutzel, R., Barbé, J. & Müller, W. A multifunctional gene (tetR) controls Tn10-encoded tetracycline resistance. *Journal of bacteriology* **150,** 633–642 (May 1982).

80. Poyart, C. & Trieu-Cuot, P. A broad-host-range mobilizable shuttle vector for the construction of transcriptional fusions to beta-galactosidase in gram-positive bacteria. eng. *FEMS Microbiol Lett* **156,** 193–198. ISSN: 0378-1097 (Print); 0378-1097 (Linking) (Nov. 1997).

81. Bryksin, A. V. & Matsumura, I. Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. eng. *PLoS One* **5,** e13244. ISSN: 1932-6203 (Electronic); 1932-6203 (Linking) (Oct. 2010).

82. Esvelt, K. M. *et al.* Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. eng. *Nat Methods* **10,** 1116–1121. ISSN: 1548-7105 (Electronic); 1548-7091 (Linking) (Nov. 2013).

83. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. eng. *Mol Syst Biol* **2,** 2006.0008. ISSN: 1744-4292 (Electronic); 1744-4292 (Linking) (2006).

84. De Man, J., Rogosa, d. & Sharpe, M. E. A medium for the cultivation of lactobacilli. *Journal of applied Bacteriology* **23,** 130–135 (1960).

85. MacConkey, A. Lactose-Fermenting Bacteria in Faeces. *Journal of Hygiene* **5,** 333–379 (1905).

86. Engler, C., Kandzia, R. & Marillonnet, S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLOS ONE* **3,** 1–7 (Nov. 2008).

87. Tu, Q. *et al.* Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Scientific Reports* **6,** (Apr. 2016).

88. Choudhry, P. High-Throughput Method for Automated Colony and Cell Counting by Digital Image Analysis Based on Edge Detection. *PloS one* **11,** e0148469, e0148469–e0148469 (Feb. 2016).

89. Doss, J., Culbertson, K., Hahn, D., Camacho, J. & Barekzi, N. A Review of Phage Therapy against Bacterial Pathogens of Aquatic and Terrestrial Organisms. *Viruses* **9,** 50 (Mar. 2017).

90. Clokie, M. R., Millard, A. D., Letarov, A. V. & Heaphy, S. Phages in nature. *Bacteriophage* **1,** 31–45 (Jan. 2011).

91. Hobbs, Z. & Abedon, S. T. Diversity of phage infection types and associated terminology: the problem with 'Lytic or lysogenic'. *FEMS Microbiology Letters* **363,** fnw047 (2016).

92. Miller, E. S. *et al.* Bacteriophage T4 genome. *Microbiology and molecular biology reviews : MMBR* **67,** 86–156 (Mar. 2003).

93. Echols, H. DEVELOPMENTAL PATHWAYS FOR THE TEMPERATE PHAGE: LYSIS VS LYSOGENY. *Annual Review of Genetics* **6.** PMID: 4604314, 157–190 (1972).

94. Oppenheim, A. B., Kobiler, O., Stavans, J., Court, D. L. & Adhya, S. Switches in Bacteriophage Lambda Development. *Annual Review of Genetics* **39.** PMID: 16285866, 409–429 (2005).

95. Chaudhry, W. N. *et al.* Leaky resistance and the conditions for the existence of lytic bacteriophage. *PLOS Biology* **16,** 1–23 (Aug. 2018).

96. Kim, S. *et al.* Quorum Sensing Can Be Repurposed To Promote Information Transfer between Bacteria in the Mammalian Gut. *ACS Synthetic Biology* **7.** PMID: 30125499, 2270–2281 (2018).

97. Palmer, A. G., Streng, E. & Blackwell, H. E. Attenuation of virulence in pathogenic bacteria using synthetic quorum-sensing modulators under native conditions on plant hosts. *ACS chemical biology* **6,** 1348–1356 (Dec. 2011).

98. Sha, J. *et al.* Deletion of the Braun Lipoprotein-Encoding Gene and Altering the Function of Lipopolysaccharide Attenuate the Plague Bacterium. *Infection and Immunity* **81** (ed McCormick, B. A.) 815–828. ISSN: 0019-9567 (2013).

99. Hsu, B. B. *et al.* Bacteriophages dynamically modulate the gut microbiota and metabolome. *bioRxiv.* doi:10.1101/454579. eprint: https://www.biorxiv.org/content/early/2018/10/26/454579.full.pdf. <https://www.biorxiv.org/content/early/2018/10/26/454579> (2018).

100. Kim, M.-S. & Bae, J.-W. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *The ISME Journal* **12,** 1127–1141 (2018).

101. Kim, E. *et al.* In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nature Communications* **8,** (Feb. 2017).

102. Nishimasu, H. *et al.* Crystal Structure of Staphylococcus aureus Cas9. *Cell* **162,** 1113–1126 (Aug. 2015).

103. Thakore, P. I. *et al.* RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors. *Nature Communications* **9,** 1674 (2018).

104. Epp, C., Pearson, M. L. & Enquist, L. Downstream regulation of int gene expression by the b2 region in phage lambda. *Gene* **13,** 327–337. ISSN: 0378-1119 (1981).

105. Rosenvold, E. C., Calva, E., Burgess, R. R. & Szybalski, W. In Vitro transcription from the b2 region of bacteriophage. *Virology* **107,** 476–487. ISSN: 0042-6822 (1980).

106. Davies, M. R., Broadbent, S. E., Harris, S. R., Thomson, N. R. & van der Woude, M. W. Horizontally Acquired Glycosyltransferase Operons Drive Salmonellae Lipopolysaccharide Diversity. *PLOS Genetics* **9,** 1–13 (June 2013).

107. Susskind, M. M. & Botstein, D. Molecular genetics of bacteriophage P22. *Microbiological reviews* **42,** 385–413 (June 1978).

108. Vander Byl, C. & Kropinski, A. M. Sequence of the genome of Salmonella bacteriophage P22. *Journal of bacteriology* **182,** 6472–6481 (Nov. 2000).

109. Hanks, S., Quinn, A. & Hunter, T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241,** 42–52. ISSN: 0036-8075 (1988).

110. Gao, R. & Stock, A. M. Biological Insights from Structures of Two-Component Proteins. *Annual Review of Microbiology* **63.** PMID: 19575571, 133–154 (2009).

111. Pearson, R. B. & Kemp, B. E. in *Protein Phosphorylation Part A: Protein Kinases: Assays, Purification, Antibodies, Functional Analysis, Cloning, and Expression* 62–81 (Academic Press, 1991). doi:https://doi.org/10.1016/0076-6879(91)00127-I. <http://www.sciencedirect.com/science/article/pii/007668799100127I>.

112. Shi, Y. Serine/Threonine Phosphatases: Mechanism through Structure. *Cell* **139,** 468–484. ISSN: 0092-8674 (2009).

113. Xiao, Q., Luechapanichkul, R., Zhai, Y. & Pei, D. Specificity profiling of protein phosphatases toward phosphoseryl and phosphothreonyl peptides. *Journal of the American Chemical Society* **135,** 9760–9767 (July 2013).

114. Hunter, T. Why nature chose phosphate to modify proteins. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367,** 2513–2516 (Sept. 2012).

115. Janczarek, M., Vinardell, J.-M., Lipa, P. & Karaś, M. Hanks-Type Serine/Threonine Protein Kinases and Phosphatases in Bacteria: Roles in Signaling and Adaptation to Various Environments. *International journal of molecular sciences* **19,** 2872 (Sept. 2018).

116. Wang, Z. & Cole, P. A. Catalytic mechanisms and regulation of protein kinases. *Methods in enzymology* **548,** 1–21 (2014).

117. Ortiz-Lombardia, M., Pompeo, F., Boitel, B. & Alzari, P. M. Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. eng. *J Biol Chem* **278,** 13094–13100. ISSN: 0021-9258 (Print); 0021-9258 (Linking) (Apr. 2003).

118. Peck, S. C. Analysis of protein phosphorylation: methods and strategies for studying kinases and substrates. eng. *Plant J* **45,** 512–522. ISSN: 0960-7412 (Print); 0960-7412 (Linking) (Feb. 2006).

119. Ferreira, A. S. *et al.* Comparative transcriptomic analysis of the Burkholderia cepacia tyrosine kinase bceF mutant reveals a role in tolerance to stress, biofilm formation, and virulence. *Applied and environmental microbiology* **79,** 3009–3020 (May 2013).

120. Switzer, A. *et al.* A role for a conserved kinase in the transcriptional control of methionine biosynthesis in Escherichia coli experiencing sustained nitrogen starvation. *bioRxiv.* doi:10.1101/254623. eprint: https://www.biorxiv.org/content/early/2018/01/27/254623.full.pdf. <https://www.biorxiv.org/content/early/2018/01/27/254623> (2018).

121. Fridman, M. *et al.* Two Unique Phosphorylation-Driven Signaling Pathways Crosstalk in Staphylococcus aureus to Modulate the Cell-Wall Charge: Stk1/Stp1 Meets GraSR. *Biochemistry* **52.** PMID: 24102310, 7975–7986 (2013).

122. Kalantari, A., Derouiche, A., Shi, L. & Mijakovic, I. Serine/threonine/tyrosine phosphorylation regulates DNA binding of bacterial transcriptional regulators. *Microbiology* **161,** 1720–1729 (2015).

123. Piñas, G. E. *et al.* Crosstalk between the serine/threonine kinase StkP and the response regulator ComE controls the stress response and intracellular survival of Streptococcus pneumoniae. *PLOS Pathogens* **14,** 1–33 (June 2018).

124. Byeon, I. J., Yongkiettrakul, S. & Tsai, M. D. Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes. eng. *J Mol Biol* **314,** 577–588. ISSN: 0022-2836 (Print); 0022-2836 (Linking) (Nov. 2001).

125. Liao, H., Byeon, I. J. & Tsai, M. D. Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53. eng. *J Mol Biol* **294,** 1041–1049. ISSN: 0022-2836 (Print); 0022-2836 (Linking) (Dec. 1999).

126. Durocher, D. *et al.* The Molecular Basis of FHA Domain:Phosphopeptide Binding Specificity and Implications for Phospho-Dependent Signaling Mechanisms. *Molecular Cell* **6,** 1169–1182. ISSN: 1097-2765 (2000).

127. Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z. & Chu, J. A Guide to Fluorescent Protein FRET Pairs. *Sensors (Basel, Switzerland)* **16,** 1488 (Sept. 2016).

128. Nelson, B. D., Manoil, C. & Traxler, B. Insertion mutagenesis of the lac repressor and its implications for structure-function analysis. eng. *J Bacteriol* **179,** 3721–3728. ISSN: 0021-9193 (Print); 0021-9193 (Linking) (June 1997).

129. Zhang, J. & Allen, M. D. FRET-based biosensors for protein kinases: illuminating the kinome. *Mol. BioSyst.* **3,** 759–765 (11 2007).

130. Komatsu, N. *et al.* Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Molecular biology of the cell* **22,** 4647–4656 (Dec. 2011).

131. Madec, E., Laszkiewicz, A., Iwanicki, A., Obuchowski, M. & Seror, S. Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes. eng. *Mol Microbiol* **46,** 571–586. ISSN: 0950-382X (Print); 0950-382X (Linking) (Oct. 2002).

132. Pietack, N. *et al.* In vitro phosphorylation of key metabolic enzymes from Bacillus subtilis: PrkC phosphorylates enzymes from different branches of basic metabolism. eng. *J Mol Microbiol Biotechnol* **18,** 129–140. ISSN: 1660-2412 (Electronic); 1464-1801 (Linking) (2010).

133. Newell, C. A. & Gray, J. C. Binding of lac repressor-GFP fusion protein to lac operator sites inserted in the tobacco chloroplast genome examined by chromatin immunoprecipitation. *Nucleic acids research* **38,** e145–e145 (Aug. 2010).

134. Ding, D.-Q. & Hiraoka, Y. Visualization of a Specific Genome Locus by the lacO/LacI-GFP System. eng. *Cold Spring Harb Protoc* **2017,** pdb.prot091934. ISSN: 1559-6095 (Electronic); 1559-6095 (Linking) (Oct. 2017).

135. Britton, R. A. *et al.* Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of Bacillus subtilis. *Journal of bacteriology* **184,** 4881–4890 (Sept. 2002).

136. Quisel, J. D., Burkholder, W. F. & Grossman, A. D. In vivo effects of sporulation kinases on mutant Spo0A proteins in Bacillus subtilis. *Journal of bacteriology* **183,** 6573–6578 (Nov. 2001).

137. Yansura, D. G. & Henner, D. J. Use of the Escherichia coli lac repressor and operator to control gene expression in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* **81,** 439–443 (Jan. 1984).

138. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. eng. *Nat Biotechnol* **26,** 127–132. ISSN: 1546-1696 (Electronic); 1087-0156 (Linking) (Jan. 2008).

139. Tamaoki, T. *et al.* Staurosporine, a potent inhibitor of phospholipidCa++dependent protein kinase. *Biochemical and Biophysical Research Communications* **135,** 397–402. ISSN: 0006-291X (1986).

140. Fernandez, P. *et al.* The Ser/Thr Protein Kinase PknB Is Essential for Sustaining Mycobacterial Growth. *Journal of Bacteriology* **188,** 7778–7784. ISSN: 0021-9193 (2006).

141. Pensinger, D. A. *et al.* Selective Pharmacologic Inhibition of a PASTA Kinase Increases Listeria monocytogenes Susceptibility to -Lactam Antibiotics. *Antimicrobial Agents and Chemotherapy* **58,** 4486–4494. ISSN: 0066-4804 (2014).

142. Vornhagen, J. *et al.* Kinase Inhibitors that Increase the Sensitivity of Methicillin Resistant Staphylococcus aureus to -Lactam Antibiotics. *Pathogens (Basel, Switzerland)* **4,** 708–721 (Oct. 2015).

143. Li, J., Lee, G. I., Van Doren, S. R. & Walker, J. C. The FHA domain mediates phosphoprotein interactions. eng. *J Cell Sci* **113 Pt 23,** 4143–4149. ISSN: 0021-9533 (Print); 0021-9533 (Linking) (Dec. 2000).

144. Pallen, M., Chaudhuri, R. & Khan, A. Bacterial FHA domains: neglected players in the phospho-threonine signalling game? eng. *Trends Microbiol* **10,** 556–563. ISSN: 0966-842X (Print); 0966-842X (Linking) (Dec. 2002).

145. Harwood, C. & Cutting, S. *Molecular Biological Methods for Bacillus* ISBN: 9780471923930. <https://books.google.com/books?id=-thiQgAACAAJ> (Wiley, 1991).

146. Oh, M. H., Lee, J. C., Kim, J., Choi, C. H. & Han, K. Simple Method for Markerless Gene Deletion in Multidrug-Resistant Acinetobacter baumannii. *Applied and Environmental Microbiology* **81** (ed Nojiri, H.) 3357–3368. ISSN: 0099-2240 (2015).

147. Wang, Y., Yau, Y.-Y., Perkins-Balding, D. & Thomson, J. G. Recombinase technology: applications and possibilities. *Plant cell reports* **30,** 267–285 (Mar. 2011).

148. Ziesack, M. *et al.* Escherichia coli NGF-1, a Genetically Tractable, Efficiently Colonizing Murine Gut Isolate. *Microbiology Resource Announcements* **7** (ed Stewart, F. J.) doi:10.1128/MRA.01416-18. eprint: https://mra.asm.org/content/7/22/e01416-18.full.pdf. <https://mra.asm.org/content/7/22/e01416-18> (2018).

149. Verma, S. C. & Mahadevan, S. The chbG Gene of the Chitobiose (chb) Operon of Escherichia coli Encodes a Chitooligosaccharide Deacetylase. *Journal of Bacteriology* **194,** 4959–4971. ISSN: 0021-9193 (2012).